

Durham E-Theses

Assessing options for the measurement of movement behaviours of pre-school children at scale

SOPHIE MICHELLE PHILLIPS

How to cite:

PHILLIPS, SOPHIE MICHELLE (2022) Assessing options for the measurement of movement behaviours of pre-school children at scale. Doctoral thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a <https://etheses.durham.ac.uk/id/eprint/14570/> is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Assessing options for the measurement of movement behaviours of pre-school children at scale

Sophie Michelle Phillips

Abstract

A recent paradigm shift suggests that the components of movement behaviour (physical activity, sedentary behaviour, and sleep) of pre-school children, which are important for child health, should be examined collectively rather than in isolation. Furthermore, there is evidence of inequalities of these behaviours, which in turn impacts on inequalities in child health. The validity, reliability, and feasibility of existing measurement tools for assessing movement behaviour in pre-school children is unclear. The aim of this doctoral research was to develop a measurement tool for public health purposes (thus a proxy reported tool) which assesses the movement behaviours of pre-school aged children.

Critical reviews of the existing literature found that there are currently no valid, reliable, and feasible measurement tools to assess movement behaviours in pre-school children. These reviews highlighted the need to address this gap in knowledge by developing a novel proxy reported measurement tool that is feasible and acceptable, as well as valid and reliable, for assessing movement behaviours of pre-school aged children.

Focus groups with parents and carers of pre-school children from deprived areas in the UK provided qualitative insights on the acceptability and feasibility of existing measurement tools, including their views on the presentation of proxy reported tools (e.g. age appropriate activities, preference for real-time recording over recall). These parents and carers also co-developed a novel measurement tool 'Movement Measurement in the Early Years' (MoveMEY), alongside input from topic expert researchers.

Working with a different group of pre-school children and their parents and carers, MoveMEY was pilot tested to determine the content validity of the tool. The tool was found to comprise of relevant items, be comprehensive of children's movement behaviours, and understandable to parents.

MoveMEY is a promising tool for the measurement of all components of movement behaviours in pre-school children and is ready for further validation.

**Assessing options for the measurement of movement
behaviours of pre-school children at scale**

Sophie Michelle Phillips

Thesis submitted to Durham University for the degree of Doctor of
Philosophy (PhD)

Department of Sport and Exercise Sciences
Durham University

2022

Table of Contents

Abstract.....	1
List of Tables.....	9
List of Figures.....	10
List of Appendices.....	11
List of Abbreviations	12
Declaration.....	14
Statement of Copyright.....	15
Acknowledgements.....	16
Published Peer Reviewed Journal Articles.....	18
Papers published from the PhD research outlined in this thesis.....	18
Additional Papers.....	18
1 Chapter One: Introduction	19
1.1 Rationale of thesis.....	19
1.2 Research aim and objectives	23
1.2.1 Aim	23
1.2.2 Objectives.....	23
1.3 Thesis Structure	24
1.4 PhD Story in a Picture	30
2 Chapter Two: Literature Review	31
2.1 Introduction	31
2.2 Definitions of key concepts.....	31
2.2.1 Early Years.....	31
2.2.2 Pre-school aged children.....	32
2.2.3 Physical activity	33
2.2.4 Sedentary behaviour.....	36
2.2.5 Sleep.....	37
2.2.6 Movement behaviours.....	38
2.3 Health behaviours in pre-school aged children	39
2.3.1 Physical activity and health and development in pre-school aged children	40
2.3.2 Sedentary behaviour and health and development in pre-school aged children	43
2.3.3 Sleep and health and development in pre-school aged children	46
2.3.4 Standing Time	49
2.4 Health Inequalities	51
2.4.1 Health Inequalities in the UK	51

2.4.2	Inequalities of movement behaviour opportunities for pre-school aged children	52
2.5	Movement behaviour approach	54
2.5.1	Overview of movement behaviour approach	54
2.5.2	Analysis of movement behaviour	57
2.5.3	Movement behaviour and health and development in pre-school aged children	58
2.6	Movement behaviour guidelines and recommendations for pre-school aged children	60
2.6.1	Global movement behaviour guidelines for pre-school aged children	61
2.6.2	History of UK physical activity and movement behaviour guidelines for pre-school aged children	63
2.6.3	Guideline adherence among pre-school aged children	68
2.7	Brief overview of physical activity and sedentary behaviour measurement tools	70
2.7.1	Calorimetry	71
2.7.2	Direct Observation	73
2.7.3	Wearable cameras	74
2.7.4	Heart Rate Monitors	75
2.7.5	Accelerometers	76
2.7.6	Combined heart rate and accelerometry	86
2.7.7	Pedometers	86
2.7.8	Consumer Wearables	87
2.7.9	Proxy-reported measurement tools	89
2.7.10	Strengths and limitations of tools	92
2.8	Brief overview of sleep measurement tools	99
2.8.1	Polysomnography	100
2.8.2	Electroencephalography	100
2.8.3	Videosomnography	101
2.8.4	Direct Observation	101
2.8.5	Accelerometers	102
2.8.6	Consumer wearables	107
2.8.7	Proxy reported measurement tools	109
2.8.8	Strengths and limitations of tools	112
2.9	Measurement of all movement behaviours simultaneously in pre-school aged children	118
2.10	Issues and error in movement behaviour measurement using proxy reported tools	122
2.11	Population level measurement of health behaviours	125
2.11.1	Importance of large scale surveillance of health behaviours	125
2.11.2	Challenges and feasibility concerns for large scale measurement	126
2.11.3	Current activity surveillance and monitoring systems	131

2.11.4	Social disparities of surveillance and measurement of health behaviours	133
2.12	Impact of the COVID 19 pandemic on children	136
2.12.1	COVID 19 and movement behaviours of pre-school aged children.....	137
3	Chapter Three: A systematic review of the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children.	140
3.1	Prelude.....	141
3.2	Introduction	141
3.3	Aims and objectives	143
3.4	Methods.....	143
3.4.1	Search Strategy	144
3.4.2	Eligibility Criteria	147
3.4.3	Screening for relevant studies	148
3.4.4	Data extraction.....	149
3.4.5	Risk of bias	149
3.4.6	Definitions of validity, reliability and feasibility.....	152
3.4.7	Interpretation and synthesis of results.....	156
3.4.8	Combining the results of individual studies for overall interpretation	159
3.5	Results.....	161
3.5.1	Study selection.....	161
3.5.2	Description of studies	162
3.5.3	Risk of bias assessment.....	167
3.5.4	Summary of measurement properties of tools	169
3.6	Discussion.....	293
3.6.1	Strengths and limitations of the review.....	299
3.6.2	Implications and recommendations for future research.....	301
3.7	Conclusion.....	304
3.8	Summary and link to next chapter.....	304
4	Chapter Four: The validity, reliability, and feasibility of measurement tools used to assess sleep of pre-school aged children: a rapid review.....	306
4.1	Prelude.....	307
4.2	Introduction	307
4.3	Aims and objectives	309
4.4	Methods.....	310
4.4.1	Search strategy.....	311
4.4.2	Eligibility criteria.....	315
4.4.3	Screening for relevant studies	316

4.4.4	Data extraction.....	316
4.4.5	Risk of bias	317
4.4.6	Definitions of validity, reliability and feasibility.....	318
4.4.7	Interpretation and synthesis of results.....	320
4.5	Results.....	322
4.5.1	Study selection.....	322
4.5.2	Description of studies	324
4.5.3	Risk of bias assessment.....	328
4.5.4	Summary of measurement properties of tools	329
4.5.5	<i>Validity</i>	329
4.5.6	<i>Reliability</i>	359
4.5.7	<i>Feasibility</i>	363
4.5.8	Generalisability of results and health inequality implications.....	369
4.6	Discussion.....	373
4.6.1	Strengths and limitations of review.....	378
4.6.2	Implications and recommendations for future research.....	380
4.7	Conclusion.....	381
4.8	Summary and link to next chapter.....	382
5	Chapter Five: Qualitative acceptability and feasibility of tools used to measure movement behaviours of pre-school aged children	383
5.1	Prelude.....	384
5.2	Introduction	384
5.3	Aims and objectives	386
5.4	Methods.....	387
5.4.1	Sampling and recruitment	387
5.4.2	Procedure.....	388
5.4.3	Materials	392
5.4.4	Data management	392
5.4.5	Data analysis	393
5.4.6	Insurance.....	395
5.4.7	Ethics.....	395
5.4.8	Risk assessment	395
5.5	Results.....	395
5.5.1	Theme 1: Reporting contextual information is important to parents and carers	398
5.5.2	Theme 2: Device based tools: practical issues.....	400
5.5.3	Theme 3: Encouraging children to wear a device	406

5.5.4	Theme 4: Presentation of diaries and questionnaires (proxy-reported tools)	408
5.6	Discussion.....	411
5.6.1	Principal findings	411
5.6.2	Strengths and limitations	416
5.6.3	Implications of findings and recommendations	419
5.7	Conclusion.....	422
5.8	Summary and link to next chapter.....	422
6	Chapter Six: Movement Measurement in the Early Years (MoveMEY): Development and content validity of a movement behaviour measurement tool for pre-school aged children.....	424
6.1	Prelude.....	425
6.2	Introduction	425
6.3	Aims and objectives	427
6.4	Methods.....	428
6.4.1	Conceptual Framework.....	428
6.4.2	General procedures	429
6.4.3	Sampling and recruitment	433
6.4.4	Remuneration for participation	435
6.4.5	Procedure.....	435
6.4.6	Materials	440
6.4.7	Data analysis	441
6.4.8	Data management	442
6.4.9	Insurance.....	443
6.4.10	Ethics.....	443
6.5	Results.....	445
6.5.1	Participants	445
6.5.2	Step 1: Provisional development of MoveMEY	448
6.5.3	Step 2: Focus groups with parents and carers of pre-school aged children.....	458
6.5.4	Step 3: Qualitative surveys with topic expert researchers	466
6.5.5	Step 4: Interviews with parents of pre-school aged children.....	473
6.5.6	Final version of MoveMEY following the full development process and pilot	487
6.6	Discussion.....	496
6.6.1	Principal findings.....	496
6.6.2	Strengths and limitations	501
6.6.3	Implications of findings and future research	505
6.7	Conclusion.....	507
7	Chapter Seven: Final discussion and conclusions.....	508

7.1	Introduction	508
7.2	Key Findings from Thesis.....	508
7.2.1	Chapter 3.....	508
7.2.2	Chapter 4.....	509
7.2.3	Chapter 5.....	511
7.2.4	Chapter 6.....	512
7.3	Limitations.....	513
7.4	Implications of Findings	517
7.5	Future Research	524
7.6	Overall conclusion.....	526
References.....		711

List of Tables

Table 2.1. Characteristics, definitions, and examples of the dimensions of physical activity.....	35
Table 2.2. Indicators of sleep quality in pre-school children (Ohayon et al., 2017).....	49
Table 2.3. Current global physical activity and 24 hour movement behaviour guidelines for pre-school children	66
Table 2.4. Sample of calibration studies of accelerometer cut points for determining physical activity and sedentary behaviour of pre-school children	84
Table 2.5. Strengths and limitations of measurement tools used to assess physical activity and sedentary behaviour.....	92
Table 2.6. Strengths and limitations of measurement tools used to assess sleep	112
Table 3.1. Search strategy and outcomes	145
Table 3.2. Definition of each of the measurement properties included in this review.....	153
Table 3.3. Level of evidence for validity studies included in this review	155
Table 3.4. Main statistical analyses and interpretation of statistics	158
Table 3.5. Overview of all measurement tools examined and measurement properties explored ..	164
Table 3.6. Detailed information of studies examining level 1 validity evidence	170
Table 3.7. Summary table of level 1 validity evidence of the measurement tools.....	190
Table 3.8. Detailed information of studies examining level 2 validity studies	194
Table 3.9. Summary table of level 2 validity evidence of the measurement tools.....	229
Table 3.10. Detailed information of studies of level 3 validity studies.....	235
Table 3.11. Summary table of level 3 validity evidence of the measurement tools	252
Table 3.12. Detailed information of studies of level 4 validity evidence.....	257
Table 3.13. Summary table of level 4 validity evidence of the measurement tools.....	263
Table 3.14. Detailed information of studies examining test-retest reliability of measurement tools.....	267
Table 3.15. Detailed information of studies examining intra-instrument reliability of measurement tools.	271
Table 3.16. Detailed information of studies with an element of feasibility of measurement tools ..	274
Table 3.17. Summary table of the reliability and feasibility of the measurement tools	288
Table 4.1. Full search strategy and outcomes	313
Table 4.2. Definition of each of the measurement properties included in this review.....	319
Table 4.3. Main statistical analyses and interpretation of statistics	321
Table 4.4. Overview of measurement properties examined in each study	326
Table 4.5. Detailed information of studies examining construct validity	331
Table 4.6. Detailed information of studies examining criterion validity	334
Table 4.7. Detailed information of studies examining convergent validity	339
Table 4.8. Detailed information of studies examining test-retest reliability	359

Table 4.9. Detailed information of studies examining internal consistency	360
Table 4.10. Detailed information on studies assessing feasibility of tools	364
Table 4.11. Summary table of the validity and reliability of measurement tools for assessing sleep of pre-school aged children.....	370
Table 5.1: Stages of thematic analysis	394
Table 5.2. Demographic characteristics of participants.....	396
Table 6.1. Demographic characteristics of participants.....	448
Table 6.2. Key findings on the format of MoveMEY from parents, carers and topic expert researchers	459
Table 6.3. Key findings on the content of MoveMEY from parents/carers and topic expert researchers	463
Table 6.4. Amendments to the tool.....	474

List of Figures

Figure 1.1. Overview of movement behaviours (Chaput et al., 2014a).....	20
Figure 1.2. Flow diagram displaying the structure of the thesis.....	29
Figure 1.3. PhD Story in a Picture	30
Figure 2.1. Behaviours examined in this thesis (WHO, 2019).....	39
Figure 2.2. Overview of impacts of physical activity on pre-school aged children.....	42
Figure 2.3. Overview of impacts of sedentary behaviour on pre-school aged children.....	45
Figure 2.4. Overview of impacts of sleep duration on pre-school aged children.....	47
Figure 2.5. Estimated distribution of movement behaviours in a 24 hour day (Chaput et al., 2014a)	55
Figure 2.6. Timeline of UK physical activity guidelines for pre-school children	65
Figure 3.1. PRISMA flow chart of included studies	161
Figure 4.1. PRISMA flow chart of included studies	323
Figure 5.1. Images of accelerometers (in order: Actigraph GT3X+, ActivPAL4 micro, Actical).....	390
Figure 5.2. Overview of key themes	397
Figure 5.3. Infographic of recommendations for measuring movement behaviours of pre-school aged children	421
Figure 6.1. Development and content validity assessment.....	431
Figure 6.2. Overview of steps taken in developing and determining content validity of MoveMEY.	432
Figure 6.3. Procedure for content validity study	440
Figure 6.4. Participants for each step of the development and content validity assessment of MoveMEY	445
Figure 6.5. Provisional tool questions mapped to guidelines.....	454
Figure 6.6. Final tool following rigorous development and pilot mapped to guidelines.....	490
Figure 7.1. Key findings from thesis: what is already known on this topic and what does this research add.	521

List of Appendices

Appendix A: Abstracts of published research.....	527
Appendix B: Training log	531
Appendix C: Knowledge exchange log	536
Appendix D: Additional projects worked on alongside PhD.....	540
Appendix E: Placement in practice	541
Chapter Appendices	542
Appendix 3.1: PRISMA Checklist for Systematic Review.....	542
Appendix 3.2: PROPSERO Systematic Review Protocol	545
Appendix 3.3: Inclusion checklist	546
Appendix 3.4: Data Extraction Form.....	548
Appendix 3.5: List of excluded studies with reasons.....	552
Appendix 3.6: Downs and Black risk of bias assessment on all studies.....	560
Appendix 3.7: COSMIN risk of bias assessment on proxy reported tools.....	561
Appendix 4.1: PRISMA Checklist for Rapid Review	562
Appendix 4.2: PROPSERO Rapid Review Protocol.....	566
Appendix 4.3: Excluded studies with reasons	567
Appendix 4.4: Risk of bias assessment of included studies.....	574
Appendix 5.1: Recruitment poster	575
Appendix 5.2: Parent/Carer Recruitment Letter.....	576
Appendix 5.3: Information sheet	577
Appendix 5.4: Privacy notice	580
Appendix 5.5: Consent form	584
Appendix 5.6: Demographic information questionnaire	586
Appendix 5.7: Accelerometer information.....	589
Appendix 5.8: Accelerometer instructions.....	590
Appendix 5.9: Accelerometer logs	593
Appendix 5.10: Sleep diary	596
Appendix 5.11: Measurement of Movement in the Early Years Provisional Questionnaire	598
Appendix 5.12: Focus group discussion guide	608
Appendix 5.13: Debrief	612
Appendix 5.14: Risk assessment	613
Appendix 6.1: COSMIN guidelines.....	616
Appendix 6.2 Studies examining movement behaviours of pre-school children at scale	622
Appendix 6.3: Current activity surveillance systems.....	635
Appendix 6.4: MoveMEY Scoring Sheet (Provisional)	645

Appendix 6.5: Face and content validity qualitative survey.....	650
Appendix 6.6: Recruitment poster	654
Appendix 6.7: Information sheet	655
Appendix 6.8: Privacy notice	657
Appendix 6.9: Consent form.....	662
Appendix 6.10: Interview discussion guide	663
Appendix 6.11: Demographic information questionnaire	666
Appendix 6.12: Movement Measurement in the Early Years measurement tool following co-development.....	671
Appendix 6.13: Study pack letter	686
Appendix 6.14: Debrief form	687
Appendix 6.15: MoveMEY Scoring Sheet (Tested in Pilot).....	688
Appendix 6.16: Movement Measurement in the Early Years measurement tool (Final version following development and pilot).....	692
Appendix 6.17: MoveMEY Scoring Sheet (Final).....	707

List of Abbreviations

BCSQ- Brief Child Sleep Questionnaire

BMI- Body mass index

CCTQ- Children's ChronoType Questionnaire

COSMIN – Consensus-based Standards for the selection of health Measurement Instruments

CSHQ- Children's Sleep Habits Questionnaire

DLW- Doubly labelled water

EMA- Ecological momentary assessment

EPHPP- Effective Public Health Practice Project

IMD- Index of multiple deprivation

LPA – Light physical activity

MET- Metabolic equivalent of task

MoveMEY- Movement Measurement in the Early Years

MVPA – Moderate to vigorous physical activity

NSF- National Sleep Foundation

PA – Physical activity

PRISMA- Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PSG- Polysomnography

SB- Sedentary behaviour

SES- Socioeconomic status

SSTQ- Sleep schedule time questionnaire

UK- United Kingdom

USA- United States of America

WHO- World Health Organization

Declaration

I confirm that the work in this thesis is my own and has not been submitted for any other degree at another university. Others have supported aspects of this research, I describe below what I did and how others supported me for each chapter where others were involved:

Chapter 3

I developed the protocol, with expert advice from my supervisors, Dr Frances Hillier-Brown and Professor Carolyn Summerbell. I conducted the searches, with help on the search strategy from Durham University librarians, Benjamin Taylorson and James Bissett. I conducted data screening, data extraction, risk of bias assessment, data synthesis and interpreted the results. Dr Matt Hobbs and Dr Frances Hillier-Brown assisted with independently double screening articles. Dr Frances Hillier-Brown and Cassey Muir assisted with independent double data extraction. Dr Frances Hillier-Brown assisted with independently conducting the risk of bias assessment and interpretation of the results.

Chapter 4

I developed the protocol, with expert advice from my supervisors: Professor Carolyn Summerbell, Dr Frances Hillier-Brown, Dr Kathryn Hesketh, and Professor Sonia Saxena, and advisor Professor Helen Ball. I conducted the searches, data screening, data extraction, risk of bias, data synthesis and interpretation. Dr Frances Hillier-Brown assisted with study eligibility. Professor Carolyn Summerbell assisted with double checking of the data extraction and risk of bias assessments for accuracy.

Chapter 5

I developed the protocol, recruitment, data collection materials, and wrote the ethics application with guidance from my supervisors. Dr Daniel Bingham assisted by providing some example accelerometer logs and instructions. I recruited the participating centres, whilst a gatekeeper at the centre facilitated recruitment of parents and carers at their setting, including providing a suitable space and time for the session. I gained consent from all participants, carried out the study sessions and completed the analysis. Transcription of the focus group discussions were conducted by a university approved provider.

Chapter 6

I developed the protocol, recruitment, data collection materials, and wrote the ethics application with guidance of my supervisors. Dr Rachel Mowbray and Joe Armstrong from County Durham Sport helped with ideas on how best to conduct this piece of research. I recruited some settings and was assisted by Jen McGuigan from Durham County Council who provided the information to additional settings. I was the primary contact for the recruitment of participants, gained consent from participants, was responsible for preparation and delivery of study materials, arranged and conducted all interview sessions, and carried out all analysis. Transcription of the interviews were conducted by a university approved provider.

Statement of Copyright

The copyright of this thesis rests with the author. No quotation from it should be published without prior written consent and information derived from it should be acknowledged. This thesis and associated research was funded through a National Institute for Health Research School for Public Health Research (NIHR SPHR) PhD studentship, Grant Reference Number PD-SPH-2015.

Acknowledgements

I would firstly like to give my deepest thank you to my supervisors, Professor Carolyn Summerbell, Dr Frances Hillier-Brown, Professor Sonia Saxena and Dr Kathryn Hesketh for their invaluable support, guidance, dedication, and positivity throughout this whole PhD process. I would like to give a special thank you to Frances, for always being there for me throughout the years of this PhD, both for academic and personal support. Thank you for continuously inspiring me through your patience, kindness, and understanding.

I am very grateful to the National Institute for Health Research, School for Public Health Research for funding my PhD studentship and for all the opportunities this has afforded me.

I am hugely grateful to all the early years settings, gatekeepers, parents and carers of pre-school children who were involved in this research, without you, none of this would have been possible. I would also like to extend my thanks to Professor Helen Ball, Dr Daniel Bingham, Dr Esther Van Sluijs, and Professor John Reilly for offering their support, guidance, and expert knowledge throughout my PhD.

My heartfelt thanks to all the amazing colleagues and friends I have met at Durham University. Ben, Amée, Verity, Alice, Carol, Gianmarco, Phill, and Naomi, thank you for being part of this journey with me and for your invaluable support and friendship throughout this process. Special thanks to Hester Hockin-Boyers, for your kind and compassionate support and friendship. You have inspired me in so many ways, both in my personal life and in my studies, for which I will be forever thankful. Karel Musílek, thank you for all the happiness, encouragement and support you have given me throughout my PhD life and for helping to make this such an enjoyable experience (even given the circumstances of a pandemic and

lockdowns!). Although there isn't enough space to write about you all here, I would like to thank all my friends further afield. I am so grateful to have such fun, caring, and supportive people in my life.

Finally, I would like to express my love and greatest thanks to my wonderful family. There will never be enough words to describe how amazing you all are and how much you mean to me. Thank you to my beautiful sister and best friend Emily, for bringing so much joy and light to my life. You are the most kind-hearted, courageous, and gorgeous person and I feel both proud and lucky that you are my sister. Josh, thank you for never failing to make me laugh, no matter what. But more so, thanks for always having my back. Daniel, thank you for being a calming influence (but never at the expense of fun!), for being a best friend and a 'go-to'- being in your company is a recipe for happiness. Mum and Dad- thank you for always believing in me, for being my biggest fans, and for being with me every step of the way, always.

Published Peer Reviewed Journal Articles

Papers published from the PhD research outlined in this thesis

Phillips, SM., Summerbell, C., Hobbs, M., Hesketh KR., Saxena S., Muir, C., and Hillier-Brown, FC. (2021). A systematic review of the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children. *International Journal of Behavioural Nutrition and Physical Activity*, 18(1), 141.

Phillips, SM., Summerbell C., Ball HL., Hesketh KR., Saxena S., and Hillier-Brown, FC. (2021) The Validity, Reliability, and Feasibility of Measurement Tools Used to Assess Sleep of Pre-school Aged Children: A Systematic Rapid Review. *Frontiers in Pediatrics*, 9:770262. doi: 10.3389/fped.2021.770262

Phillips, SM., Summerbell, C., Hesketh, K., Saxena S., and Hillier-Brown, FC. (2022). Parental views on the acceptability and feasibility of measurement tools used to assess movement behaviours of pre-school children: a qualitative study. *International Journal of Environmental Research and Public Health*, 19, 3733.

<https://doi.org/10.3390/ijerph19063733>

Additional Papers

Griffin, N., **Phillips, SM.,** Hillier-Brown, FC., Wistow J., Fairbrother, H., Holding, E., Powell, K., and Summerbell, C. (2021). A critique of the English national policy from a social determinants of health perspective using a realist and problem representation approach: the 'Childhood Obesity: a plan for action' (2016, 2018, 2019). *BMC Public Health* **21**, 2284. <https://doi.org/10.1186/s12889-021-12364->

1 [Chapter One: Introduction](#)

1.1 [Rationale of thesis](#)

Human movement occurs along a spectrum of intensity that can be classified into three distinct behaviours: physical activity, sedentary behaviour, and sleep (see figure 1.1, extracted from Chaput et al., 2014a). Previously, the three behaviours have been examined independently and demonstrate associations with numerous health and developmental indicators in pre-school aged children (Carson et al., 2017; Chaput et al., 2017; Poitras et al., 2017). However, recent evidence suggests that an integrated approach to viewing human movement behaviours, as opposed to viewing the behaviours in isolation, may be a stronger approach to improve the health of children and address current public health concerns (Chaput et al., 2014a; Chaput, Saunders and Carson, 2017; Okely et al., 2018; Tremblay, 2020; Tremblay and Ross, 2020). The rationale behind this is that time spent in one behaviour may modify the health-related influence of time spent in any other of the behaviours (Buman et al., 2014; Rao, Orpana and Krewski, 2016; Rosenberger et al., 2019). Alongside this, evidence suggests that movement (and other health) behaviours established during the early years may track through childhood and adolescence (Biddle et al., 2010; Jones et al., 2013; van Ekris et al., 2020) and into adulthood (Craigie et al., 2011; Telama et al., 2014). The early years are consequently an important group of focus to examine and establish healthy movement behaviours.

The combinations of physical activity, sedentary behaviour and sleep (hereon collectively termed ‘movement behaviours’), have become a major focus for health promotion and prevention of non-communicable diseases (World Health Organization, 2017). This is reflected in the change of international guidance. In 2019, the World Health

Organization (WHO) developed and published 24-hour movement behaviour guidelines for pre-school aged children that encompassed physical activity, sedentary behaviour and sleep recommendations (WHO, 2019), with similar guidelines implemented worldwide (Draper et al., 2020a; Ministry of Health New Zealand, 2017b; Okely et al., 2017; Tremblay et al., 2017b). The guidelines were developed following recommendations from two major WHO reports: *Report on the Commission of Ending Childhood Obesity* (WHO, 2017) and the *Global Action Plan on Physical Activity* (WHO, 2018), calling for leadership from government and actions from policy makers to devise and implement guidelines on all three behaviours, in an act towards improving child health. The current United Kingdom (UK) government guidelines for pre-school aged children focus on physical activity only (Chief Medical Officers, 2019), despite the expert committee recommending 24-hour movement behaviour guidelines to the UK Chief Medical Officers (Reilly et al., 2018; Reilly et al., 2020). The expert group stated that to appropriately tackle factors high on the public health agenda in the UK, such as obesity prevention, there is the need to shift the focus to understand how children in their early years spend the whole 24 hours of their day.

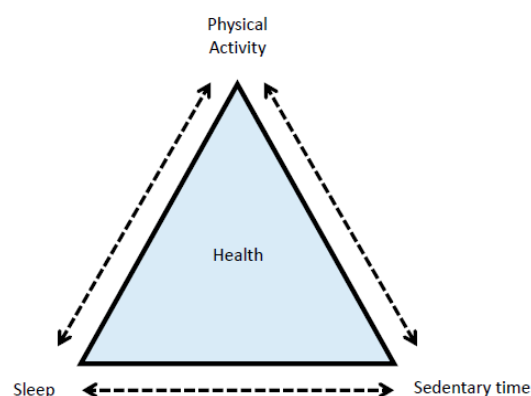


Figure 1.1. Overview of movement behaviours (Chaput et al., 2014a)

Limited evidence has been identified examining combinations of movement behaviours with health and development indicators in young children (Kuzik et al., 2017). It has been suggested that measurement issues, in particular a lack of valid, reliable, and feasible methods used to assess the whole spectrum of movement behaviours may explain this lack of evidence (Kuzik et al., 2017). At present, there are inconsistencies in measurement, with the use of multiple tools of unknown or poor quality (Chaput et al., 2017; Kuzik et al., 2017). The measurement tool of choice underpins movement behaviour research and is one of the first considerations in any study examining these behaviours. Methodological research, including measurement, surveillance, data processing and analysis, are the first step in the process towards progressing this research area (Pedišić, Dumuid and Olds, 2017). Appropriate methods of measurement are needed in order to assess the movement behaviours of pre-school aged children (Aubert et al., 2021; WHO, 2019), including to:

- Improve understanding of patterns of the movement behaviours, and determine how they relate to other health, well-being and developmental indicators
- Monitor and survey population level prevalence and trends in movement behaviours
- Determine if guidelines and recommendations are being met
- Advance the evidence base examining movement behaviours to help inform the development of further recommendations, interventions, policies and programmes
- Assess effectiveness of interventions and initiatives aimed at changing these behaviours

(Bauman et al., 2006; Chaput, 2019; Chen et al., 2021; Tremblay et al., 2016).

Despite this, there is currently a paucity of valid, reliable, and feasible methods available to measure these behaviours of pre-school children, particularly on a large scale. Measurement tools used at large scale need to be practical and feasible, and not intensive or burdensome for participants. It is important to ensure measurement tools are valid, acceptable, and feasible in lower socioeconomic status groups due to the strong gradient of poorer health as socioeconomic status decreases across the population. As such, this thesis aims to address this need, by examining ways of measuring the movement behaviours of pre-school children, suitable for use at scale. The most recent data collected at scale on physical activity of pre-school children in England was by the Health Survey for England in 2015, which included responses for 1276 children aged 2-4 years old (NHS Digital, 2016). Based on this, the work in this thesis examining and developing a tool for use 'at scale' is referring to a tool that could be used to measure movement behaviours of 1000's of children. There is a need for a tool that can be used at scale for the purpose of population surveillance of the movement behaviour guidelines of pre-school children (WHO, 2019). A tool used for this purpose needs to primarily be feasible for use by large population samples. In addition, it is important that the tool is valid in terms of appropriately detecting the behaviours that children of this age engage in, including the specific behaviours outlined in the guidelines (e.g. screen time) and can demonstrate good reliability in that the tool is stable at detecting behaviours at different time points. The tool needs to be able to determine all aspects of the guidelines, to quantify compliance with the guidelines, including:

- 1) Duration and frequency of physical activity, including being able to detect moderate to vigorous physical activity
- 2) Duration, frequency, and type of sedentary behaviour, specifically including screen time and time spent restrained (e.g. in a pram or car seat).
- 3) Duration and quality of sleep

For this PhD thesis, this included evaluating pre-existing tools and the co-development of a new tool used to examine movement behaviours of pre-school aged children, with a focus on targeting individuals from lower socioeconomic status communities.

1.2 Research aim and objectives

1.2.1 Aim

To address the need identified above, the aim of the present research was to:

- Develop a measurement tool suitable for assessing the movement behaviours of pre-school aged children that can be used at scale to survey international guidelines.

1.2.2 Objectives

In order to achieve this aim, the objectives of the thesis were:

- 1) To conduct a systematic review to identify the measurement properties (validity, reliability, and feasibility) of existing measurement tools used to assess physical activity and sedentary behaviour of pre-school aged children (covered in *Chapter 3*).
- 2) To conduct a rapid review to identify the measurement properties (validity, reliability, and feasibility) of existing measurement tools used to assess sleep of pre-school aged children (covered in *Chapter 4*).

- 3) To conduct focus groups to determine the acceptability and feasibility of measurement tools used to assess the movement behaviours (physical activity, sedentary behaviour, sleep) of pre-school aged children, from the perspective of parents and carers (covered in *Chapter 5*).
- 4) To develop a measurement tool used to assess the movement behaviours of pre-school aged children, together with parents and carers of pre-school aged children, and topic experts (covered in *Chapter 6*).
- 5) To pilot the devised measurement tool and conduct interviews with parents and carers of pre-school aged children following the pilot, to determine the content validity of the devised measurement tool for assessing the movement behaviours of pre-school aged children (covered in *Chapter 6*).

1.3 Thesis Structure

This thesis has started with the *Introduction (Chapter 1)*, outlining the purpose and rationale of the research that makes up this thesis, the aim and objectives, and here presenting the overall structure. Following this, *Chapter 2* introduces the thesis topic through a thorough critical review of the literature. The *Literature Review* presents relevant research on physical activity, sedentary behaviour, and sleep of pre-school aged children. This includes: definitions of key concepts, the links between movement behaviours and health and development outcomes, a discussion on health inequalities to demonstrate the importance of focusing on lower socioeconomic status communities within the empirical work within this thesis, the introduction of the movement behaviour approach (integration of physical activity, sedentary behaviour and sleep), and a substantial focus on the

measurement of movement behaviours (in terms of both available tools and measurement at scale through population level surveillance).

Chapters 3 to 6 of this thesis each flow on from one another as part of the overall 'story' of this thesis (as can be seen in Figure 1.2 and a picture version in Figure 1.3), however, each chapter can also be read as standalone pieces of research.

Chapter 3 *'A systematic review of the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children'* **addresses the first objective of this thesis.** In this chapter, I present a systematic review of the literature assessing the validity, reliability, and feasibility of measurement tools used to examine the physical activity and sedentary behaviour of pre-school aged children. Key findings from this review included that a range of accelerometers (in particular Actigraph GT3X, Actical, and ActivPAL) have good validity and reasonable feasibility for the assessment of physical activity and sedentary behaviour outcomes in young children. Few proxy measures were shown to provide valid data, alongside useful contextual information not captured by device-based measurement tools, however, the evidence was limited and of low quality. Some key gaps in the literature were identified through this piece of research, which informed the direction of the subsequent chapters (*Chapters 5 and 6*). This included: 1) More research on the feasibility and acceptability of measurement tools, particularly through qualitative research (*Chapter 5*) and 2) Further and more in-depth evaluation (primarily concerning examining face and content validity) of proxy reported measurement tools (*Chapter 6*).

Chapter 4 *'The validity, reliability, and feasibility of measurement tools used to assess sleep of pre-school aged children: a rapid review'* **addresses the second objective of**

this thesis. In this chapter, I present a systematically conducted rapid review of the literature assessing the validity, reliability, and feasibility of measurement tools used to examine sleep of pre-school aged children. I also present the rationale for sleep being examined in a separate review to the work on physical activity and sedentary behaviour tools (*Chapter 3*). Key findings from this review include highlighting a very limited evidence base in this area (n=16 studies), with few measurement tools evaluated, and most containing high risk of bias. One sleep diary, from a low risk of bias study, was shown to be valid for assessing sleep duration of pre-school aged children. No measurement tools were appropriate for determining sleep quality. Reporting of reliability and feasibility was minimal. Key gaps in the literature included the need to evaluate the validity and reliability of accelerometers for assessing sleep of pre-school aged children, which in turn will improve the quality of studies that assess questionnaires and diaries against accelerometers. Similar gaps in the literature were present to the review provided in *Chapter 3*, including: 1) A need for more qualitative work to examine the feasibility and acceptability of tools for assessing sleep of pre-school aged children (*Chapter 5*), and 2) Further work examining the content validity of proxy reported measurement tools (*Chapter 6*).

Chapter 5 '*Qualitative acceptability and feasibility of tools used to measure movement behaviours of pre-school aged children*' **addresses the third objective of this thesis.** In this chapter, I begin by providing a brief introduction on the importance of understanding the feasibility and acceptability of measurement tools used to measure movement of pre-school aged children, in particular the importance of assessing this using qualitative methodology. I then provide a detailed description of the qualitative methodology used, which included a demonstration of tools suitable for measuring movement behaviours of pre-school aged children (as identified through the reviews

presented in *Chapter 3* and *Chapter 4*), followed by focus group discussions to obtain the views of the parents and carers on the various tools. Results are presented according to the four overarching key themes, and verbatim quotes from the focus groups are used to support these themes. Key themes identified through this study include: 1) reporting contextual information being important to parents and carers, 2) practical issues when using device based tools, 3) encouraging children to wear a device, 4) preferences for presentation of diaries and questionnaires. This chapter discusses how these findings can contribute and inform the design, development, and implementation of future research projects measuring the movement behaviours of pre-school aged children.

Chapter 6 ‘Movement Measurement in the Early Years (MoveMEY): Development and content validity of a movement behaviour measurement tool for pre-school aged children’ addresses the final two objectives, objectives four and five, of this thesis. In this chapter, I present the full development process, including a pilot study and assessment of content validity of ‘Movement Measurement in the Early Years’ (MoveMEY), a parental reported measurement tool used to assess physical activity, sedentary behaviour, and sleep of pre-school aged children. This chapter starts by outlining the initial development process of MoveMEY and outcomes of these steps, which included:

- 1) Searching of the literature for pre-existing measurement tools, including:
 - The validity and reliability of measurement tools (namely proxy reported tools) identified through the reviews outlined in *Chapter 3* and *Chapter 4* of this thesis.
 - Thorough searches of the movement behaviour academic literature to identify questions used as part of large-scale studies.
 - Determining current surveillance systems.

- 2) Devising a provisional (to be used as a stimulus in focus group discussions) measurement tool based on the literature, by mapping questions to the current WHO 24-hour movement behaviour recommendations for pre-school aged children (WHO, 2019).
- 3) Co-development of the MoveMEY tool through focus group discussions with parents and carers of pre-school children and responses to qualitative surveys completed by a group of topic relevant researchers in the field of movement behaviours of pre-school children.

Following this, I present the results of a pilot study of the co-developed measurement tool, MoveMEY, where the content validity including comprehensiveness, comprehensibility and relevance of the tool was examined. This included interviews with parents and carers of pre-school aged children following completion of the MoveMEY measurement tool.

Chapter 7- The final chapter of this thesis discusses the key findings of the research conducted and presented within this thesis, the implications of the research, and directions for future research.

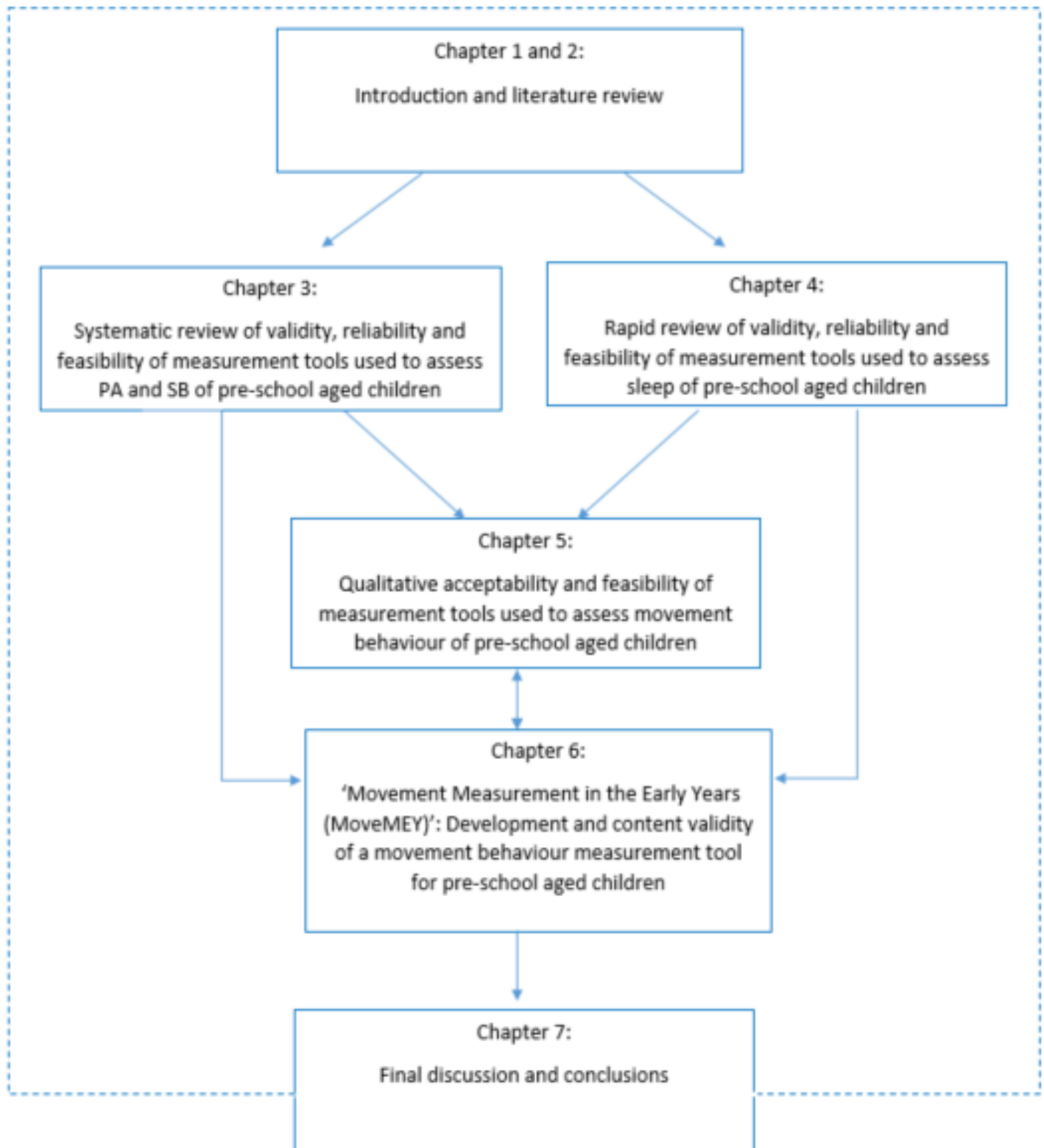


Figure 1.2. Flow diagram displaying the structure of the thesis

1.4 PhD Story in a Picture

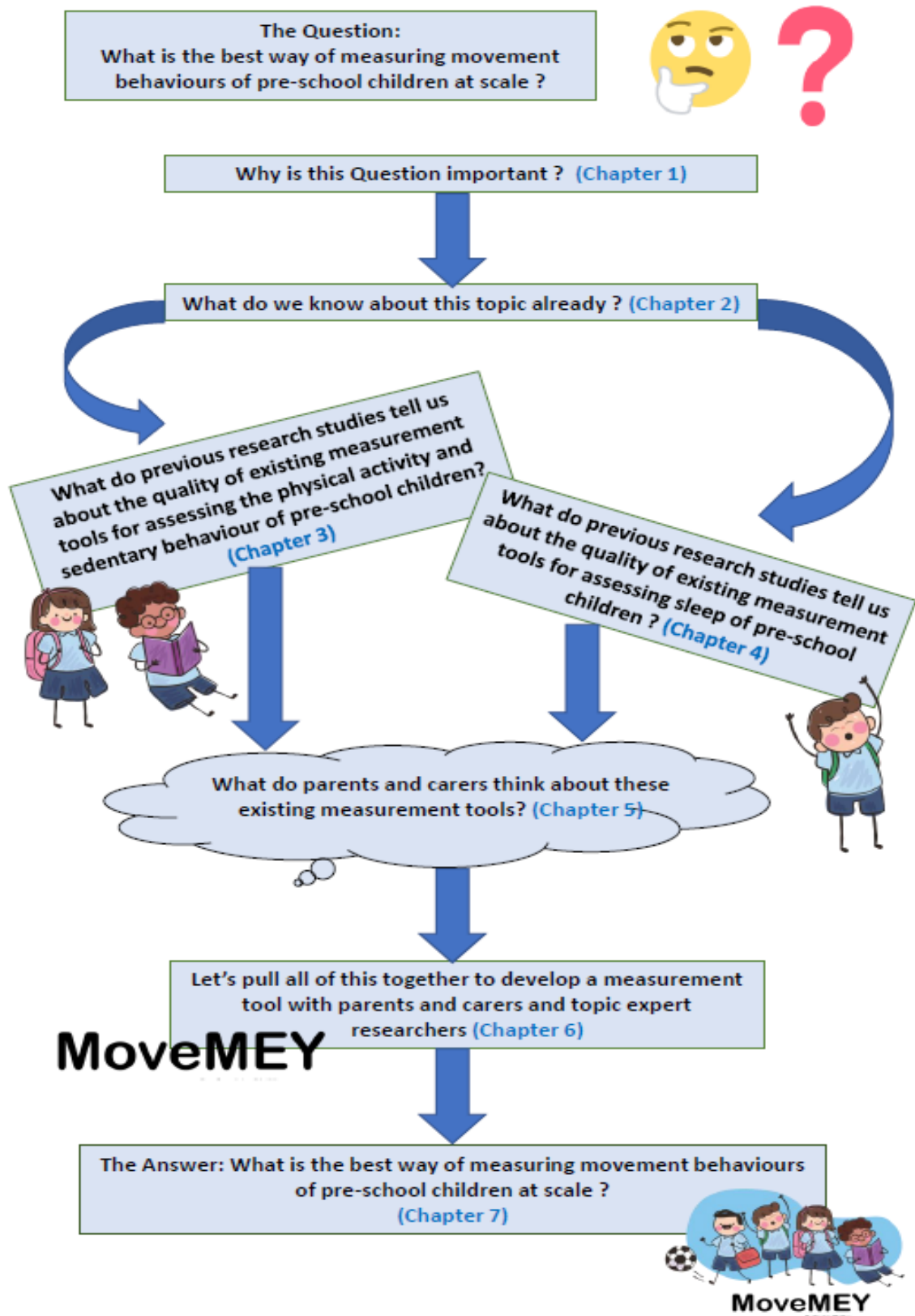


Figure 1.3. PhD Story in a Picture

2 [Chapter Two: Literature Review](#)

2.1 [Introduction](#)

This chapter outlines a critical review of the literature that presents the foundation for this thesis. The chapter starts by outlining the key concepts that underpin the thesis, followed by a discussion of the health behaviours in the early years and associations with health and developmental outcomes. The chapter then moves on to discussing health inequalities, broadly in the UK, followed by inequalities in movement behaviours in pre-school aged children. Following this, the chapter discusses the movement behaviour approach, including the evolution of guidelines and recommendations. As the focus of this thesis is measurement of movement behaviours, particularly at scale, I provide sections outlining overviews on the available tools to measure these behaviours, including strengths and limitations of each tool, tools applicable for measuring all movement behaviours, and a thorough discussion on population level measurement of movement behaviours. As a result of this PhD research taking place during the COVID-19 pandemic, I finish this chapter by briefly discussing some of the literature on movement behaviours during the pandemic, which has exacerbated the importance of this area.

2.2 [Definitions of key concepts](#)

2.2.1 [Early Years](#)

The 'early years' refers to children within the age range of 0-5 years old. The early years can be further divided into three distinct age groups: infants aged >1 year, toddlers aged 1-2 years, and pre-school children aged 3-5 years. The early years are a period of rapid growth and development, both cognitively and physically (Malina, Bouchard and Bar-Or, 2004). The difference in physical development for children in each of the early years age

groups, means that the concept of physical activity (PA) differs between these groups. For example, infants who are not at the stage of physical development to be ambulatory are encouraged to engage in tummy time; ambulatory toddlers should be active; whilst pre-school children should also take part in higher intensity activity (Chief Medical Officers, 2019). Similarly, the amount of sleep recommended for each of the age groups differs: 12-15 hours for infants; 11-14 hours for toddlers; and 10-13 hours for pre-schoolers (Hirshkowitz et al., 2015a; Hirshkowitz et al., 2015b). In line with these differences, infants and toddlers will not be the focus of the original research in this thesis. Pre-school children will be the primary focus, which I define and provide a rationale for in the next section (section 2.2.2).

2.2.2 Pre-school aged children

This thesis will focus specifically on the pre-school age group. A pre-school aged child is a child who has not yet reached the age of formal schooling. The age classification of pre-school children varies internationally based on compulsory school starting age (World Bank Data, 2020); this ranges from aged 4 in Northern Ireland (The Education Reform Order, 1989) to aged 8 in the states of Pennsylvania and Washington in the United States of America (USA) (National Center for Education Statistics, 2018). Pre-school age is when the movement behaviours (alongside eating habits) become an important area of focus, and where aspects of health, including childhood obesity, become an area of concern (NHS Digital, 2021). Recent evidence shows that health-related behaviours (including screen time and PA) at pre-school age influences increased engagement in health behaviours including PA in later childhood (Martin et al., 2022); making this period an important time to establish health behaviours.

The original research presented in this thesis will focus specifically on pre-school children aged 3 to 4 years old. There are several reasons for this:

- 1) Firstly, global PA and movement behaviour recommendations for pre-schoolers are targeted at children aged 3-4 years old (Chief Medical Officers, 2019; Ministry of Health New Zealand, 2017b; Tremblay et al. 2017b; WHO, 2019).
- 2) Secondly, the research within this thesis has all taken place within the UK where the compulsory school starting age is 5 years old (Department of Education, 2020).

The reviews presented in this thesis will include children aged 3-7 years, in order to capture studies with children classified as pre-school age internationally (World Bank Data, 2020).

2.2.3 Physical activity

A main focus of this thesis is the behaviour of physical activity (PA). PA is defined as *‘any bodily movement produced by skeletal muscles that results in energy expenditure’* (Bouchard, Blair and Haskell, 2012). Among young children, PA is often achieved through informal activity such as active play (Sims et al., 2021). Although there is currently no formal definition of active play, a review of the literature found that themes used to describe active play in pre-schoolers included: ‘increased energy’, ‘functional and dramatic play’, ‘unstructured’, ‘fun’, and ‘freely chosen’ (Truelove, Vanderloo and Tucker, 2017). PA is a complex and multifaceted behaviour that can be characterised by varying dimensions, including: 1) Frequency; 2) Intensity; 3) Duration; 4) Type; and 5) Domain. The research question being examined in any project, and the dimensions of PA being examined, will determine the most appropriate measurement tool to use. This will be discussed at length in

section 2.7 (starting on p.70). Table 2.1 displays each of these characteristics, with the definition and examples of what this means, specifically for the pre-school population.

Table 2.1. Characteristics, definitions, and examples of the dimensions of physical activity

Characteristic	Definition	Examples
Frequency	How often physical activity takes place during a period of time.	Number of times a child engages in moderate to vigorous PA (MVPA) during one day.
Intensity	Physiological energy cost of the physical activity, often defined in terms of metabolic equivalent of task (MET) values.	Light PA (LPA): 1.23-2.33 METs MVPA: 3.00-6.77 METs (Puyau et al., 2016)
Duration	How long physical activity takes place during a period of time.	30 minute walk to school; 20 minutes active outdoor play at break time. Also includes time spent in activities of specific intensities, for example minutes per day in MVPA
Type	What type of physical activity takes place.	Active play, sport.
Context	The situations in which physical activity takes place, which includes multiple different components: where, who with, why /purpose of activity and when (season, day of week, time of day).	The physical activity could be part of a structured session that parents have paid for their child to attend with other children, or a child could be jumping on a trampoline on their own.
Domain	Where the physical activity takes place.	Home, preschool, physical education lessons, through sports teams, active travel, habitual PA.

The term PA is often incorrectly used interchangeably with the terms physical fitness, exercise, and sport, within scientific literature, the media and wider society.

Definitions of these attributes are (Caspersen, Powell and Christenson, 1985):

- 1) Physical fitness: *'A set of attributes that people have or achieve that relates to the ability to perform PA'*
- 2) Exercise: A subcategory of PA *'planned, structured, and repetitive bodily movement performed to improve or maintain physical fitness'*
- 3) Sport: *'A game, competition, or activity needing physical effort and skill that is played or done according to rules, for enjoyment and/or as a job'* (Hardman and Stensel, 2009).

2.2.4 Sedentary behaviour

Sedentary behaviour (SB) is defined as *"any waking behaviour characterised by an energy expenditure ≤ 1.5 metabolic equivalents, while in a sitting, reclining or lying posture"* (Tremblay et al., 2017a, p.9). Energy expenditure when engaging in different activities varies by age, due to changes in muscle mass and physical capabilities, which is certainly the case for a range of physical activities (Puyau et al., 2016). However, the energy cost of SB is thought to be consistent in pre-school children to that outlined in the consensus definition outlined above which was not targeted at young children. Reilly and colleagues explored this, and found values ranged from 1.17 to 1.38 MET's for a range of SB (TV viewing, sitting at a table, and floor based play) (Reilly et al., 2015). Slightly more variable yet comparable MET values were observed in a different sample of pre-school aged children ranging from 0.97 to 2.10 MET's (Puyau et al., 2016).

An important distinction to make is that SB is a separate and distinct construct from physical inactivity (Chinapaw et al., 2011; van der Ploeg and Hillsdon, 2017). SB is not a mere absence of PA, but an independent group of behaviours done whilst sitting, lying or reclining (Tremblay et al., 2017a). Whilst physical inactivity is defined as a lack of PA, such as not meeting the guidelines (Tremblay et al., 2010). An individual can be sufficiently physically active and meet the guidelines, yet still spend large proportions of their day engaging in SB. In adults, research has demonstrated independent associations with health from both SB and PA (Knaeps et al., 2016; Santos et al., 2014). A large review of the evidence suggested that relationships between SB and mortality were present independent of PA (Biswas et al., 2015). However, they also noted that those engaging in higher levels of PA had decreased negative associations (Biswas et al., 2015). At present, there is not sufficient evidence to suggest these independent associations for young children due to lack of high quality, longitudinal studies (Carson et al., 2017; Poitras et al., 2017). However, in light of the evidence, SB will also be examined in this thesis, alongside PA but as a separate group of behaviours.

2.2.5 Sleep

Sleep is a complex phenomenon, entailing varying aspects, including behavioural, physiological and neural components. Sleep is often defined as a state of disengagement and loss of consciousness to the environment (Carskadon and Dement, 2005). Sleep can be characterised by multiple dimensions, including: duration, quality, sleep architecture (stages of sleep in terms of brain activity, proportion of slow-wave and rapid eye movement sleep) and schedule (sleep onset and offset, naps, circadian rhythm) (Sadeh, 2015). In this thesis, the main sleep dimensions of focus will be sleep duration and quality, as these are the

dimensions used within public health guidance (Hirshkowitz et al., 2015a; Hirshkowitz et al., 2015b; Ohayon et al., 2017; WHO, 2019). Sleep duration is the total amount of sleep obtained in a 24-hour period, including both day time and night time sleep. Sleep quality can be determined using multiple indicators, including: sleep latency, minimal night time awakenings and wake after sleep onset, and high sleep efficiency (Ohayon et al., 2017).

2.2.6 Movement behaviours

The research in this thesis is focusing on the main three movement behaviours that make up the 24 hours of the day, as just outlined in the sections above: PA, SB, and sleep (see figure 2.1, image extracted from WHO, 2019). There are multiple terms that have been used to define the behaviours that encompass the 24 hour cycle, including: movement behaviours, physical behaviours, time-use behaviours, time-use activity behaviours, 24 hour activity cycle (Falck et al., 2021). In this thesis, the term movement behaviours is used as this is the definition used in national public health guidelines (Falck et al., 2021) and most commonly used for research projects examining early years movement (Kuzik et al., 2017; SUNRISE, 2020). When I refer to 'movement behaviours' throughout this thesis, this includes the combination of PA, SB, and sleep.

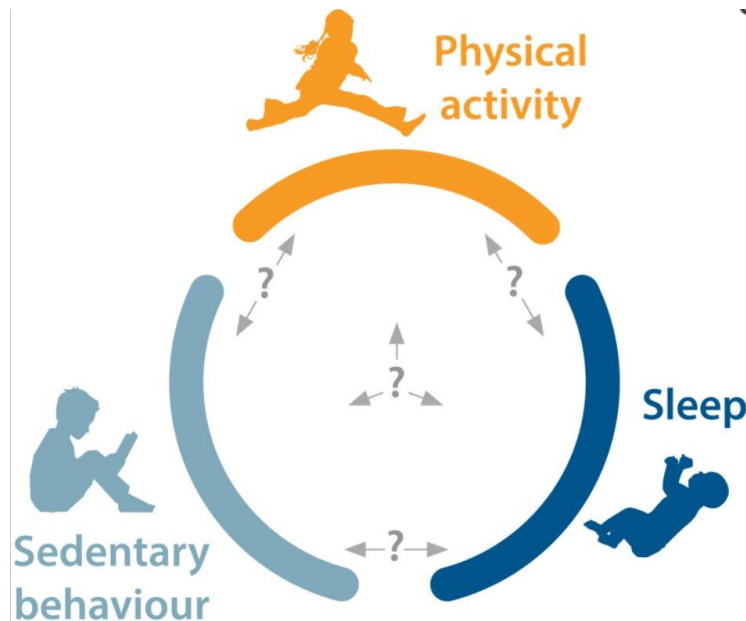


Figure 2.1. Behaviours examined in this thesis (WHO, 2019)

2.3 Health behaviours in pre-school aged children

PA, movement, and active play are important for healthy growth and development for children in their early years, as recently highlighted by the WHO in their publications: *'Report on the Commission on Ending Childhood Obesity'* (WHO, 2017) and *'A Global Action Plan on Physical Activity'* (WHO, 2018). The WHO emphasised the need to provide guidance on healthy sleep, PA, and reducing SB for children in their early years.

Thus far, the majority of the evidence on the movement behaviours for the general population (with most research focusing on adults) focuses on PA, or more so the risk factors of physical inactivity, which is said to be a leading risk factor for poor health and premature mortality (Forouzanfar et al., 2016; Katzmarzyk et al., 2021; Lee et al., 2012; WHO, 2021). Evidence suggests that physical inactivity causes 6-10% of the main non-communicable diseases identified as threats to global health: coronary heart disease, type 2 diabetes and breast and colon cancers (Lee et al., 2021). Recent and updated estimates

suggest that physical inactivity is strongly associated with increased risk of all outcomes previously found, as well as: cardiovascular disease, hypertension, dementia, depression, stroke, and seven different types of cancer (Katzmarzyk et al., 2021). In addition to health implications, globally, physical inactivity is estimated to have economic implications of 54 billion international dollars per year in direct health care costs (WHO, 2018). Although this evidence focuses on adult samples, in the WHO publications outlined above, and the wider literature, it has been suggested that to appropriately address public health crises impacting children, such as physical inactivity and childhood obesity, a broader and more integrated approach to viewing human movement, including understanding how children in their early years spend their 24 hours of the day, is required (Chaput et al., 2014; Reilly et al., 2020). Further, there is evidence to suggest that health behaviours, including PA and SB established during the early years may track (stability or relative ranking over time) through childhood and adolescence (Biddle et al., 2010; Jones et al., 2013; van Ekris et al., 2020) and into adulthood (Craigie et al., 2011; Telama et al., 2014); making the early years a critical area of focus to establish healthy movement behaviours.

2.3.1 Physical activity and health and development in pre-school aged children

PA, including active and outdoor play are associated with a host of health and developmental indicators in pre-school aged children. These include beneficial impacts on motor development, psychosocial health, cognitive development, bone and skeletal health, cardiometabolic health, physical fitness and adiposity (Carson et al., 2017; Timmons et al., 2012) (see figure 2.2). Although much of the evidence is based on randomised controlled trials and cross-sectional studies, a recent longitudinal analysis found that meeting the PA

guidelines as a pre-school aged child was associated with improved bone mineral density and content in later childhood (aged 9-11 years old) (Hinkley et al., 2020).

Figure 2.2. Overview of impacts of physical activity on pre-school aged children

Motor development	Psychosocial health	Cognitive development	Bone and skeletal health	Cardiometabolic health	Fitness	Adiposity
•Improved locomotor skills (e.g. walking, crawling, skipping) •Improved object control skills (e.g. throwing and catching) •Improved psychomotor skills (e.g. balance)	•Improved social competence (child's ability to handle social interactions such as relationships/get along with others) •Fewer behavioural problems (e.g. aggression)	•Improved creativity •Improved attention •Improved language (both word recall and ability to speak foreign language)	•Improved bone mineral content •Improved bone stiffness •Improved vitamin D levels (particularly for outdoor activity)	•Improved blood pressure •Improved cholesterol	•Improved cardiorespiratory fitness •Improved muscular fitness •Improved speed agility	•Lower body mass index (BMI)

2.3.2 Sedentary behaviour and health and development in pre-school aged children

In recent years, there has been an increasing awareness of the differential effects of SB and PA on health and developmental outcomes for young children, and of the risks associated with SB of pre-school children (LeBlanc et al., 2011; Poitras et al., 2017). This is in line with the rapid increase in SB research in the past 20 years (Biddle et al., 2019). Research demonstrates that SB shows unfavourable associations with adiposity, motor skill development, psychosocial health, some elements of cognitive development, and physical fitness (Poitras et al., 2017) (see figure 2.3).

Initially, much research on SB for all age groups, including pre-school children, was primarily through capturing TV viewing as a proxy for SB (Atkin et al., 2012; Bryant et al., 2007); demonstrating detrimental impacts of TV viewing and screen time for health and development (Dietz and Gortmaker, 1985; Poitras et al., 2017). TV viewing has been shown to be unfavourably associated with vocabulary in pre-school aged children (Carson, Rahman and Wiebe, 2017). However, TV viewing may not be representative of overall SB (Biddle, Gorely and Marshall, 2009). There are now a range of screen based activities included in measures of screen time for young children, for example, mobile phones, tablets and computers (Byrne, Terranova and Trost, 2021); however, screen time still accounts for only a proportion of the total time spent sedentary (Hoffmann et al., 2019). Although a child may engage in screen time whilst sedentary, these behaviours may be distinct constructs, as whether the impact of screen time operates through 'sitting' is unknown. Furthermore, there may be differences between active (e.g. video games) versus passive (e.g. TV viewing) screen time on health and developmental outcomes, and different types of screen based activities should be considered independently (Sweetser et al., 2012).

There is increasing acknowledgement of other activities that take place whilst sedentary, which for young children can include reading, storytelling, colouring, educational activities, and playing with toys, or sedentary whilst travelling (e.g in a car seat, on the bus). There is a small evidence base, but growing awareness around the differential impacts on health and development that varying SB may have (Carson et al., 2015). For example, reading and storytelling, even whilst sedentary, can be beneficial for the cognitive development of pre-school aged children (Carson et al., 2015) (see figure 2.3). Recent research has demonstrated that SB was a significant predictor of school readiness (whilst PA was not) (Jones et al., 2021). Although the type/context of SB was not examined, it is plausible that the type of activity conducted whilst sedentary may influence these outcomes, and are important to capture, to determine the full extent of which behaviours are beneficial for health and development of young children. Much more research is needed in this area, to understand the distinctions between different types of SB and screen time on the health and development of pre-school aged children; which requires measurement tools able to detect and make distinctions between type of activity.

Figure 2.3. Overview of impacts of sedentary behaviour on pre-school aged children

Adiposity	Motor skill development	Psychosocial health	Cognitive development	Fitness
•Higher weight status •Higher body mass index (BMI)	•Adverse locomotor skills	•Increased emotional problems •Increased behavioural problems including aggression	•Adverse communication and language development •Adverse auditory development •Increased attentional problems •*Improved language development* (beneficial when the behaviour involves reading/storytelling)	•Adverse physical fitness including reduced standing long jump performance

(Note: most of the evidence is based on screen time which may have unfavourable impacts on health, but interactive non-screen based activities whilst sedentary, such as reading and storytelling, can have beneficial impacts e.g. *language development*)

2.3.3 Sleep and health and development in pre-school aged children

Historically, research on paediatric sleep focused on the medical model of sleep, including the presence or absence of sleep disorders and ‘problem’ sleep (Buysse, 2014), reflected in the vast number of questionnaires devised for this purpose (Spruyt and Gozal, 2011). However, there is an increased public health interest in healthy sleep (Chaput, 2019; Chaput and Shiau, 2019), including health promotion and prevention regarding sleep (Nelson and Bonuck, 2015), with the emergence of the term “sleep health” (Buysse, 2014; Chaput, 2019). In this sense, ‘sleep health’ or ‘healthy sleep’ can be determined by multiple dimensions, including sufficient duration, quality, and timing of sleep, rather than merely the presence or absence of sleep disorders (Buysse, 2014; Chaput and Shiau, 2019).

The emerging evidence suggests that healthy sleep habits for children in their early years are linked with health and developmental indicators; with shorter sleep duration in pre-school children associated with increased adiposity, poor emotion regulation and cognitive development, lower school readiness, and increased risk of injury (Chaput et al., 2017; Tso et al., 2016). A longitudinal analysis revealed that meeting sleep guidelines as a pre-school child was associated with better educational outcomes (reading, spelling, numeracy, language, writing) in later childhood (Hinkley et al., 2020). Evidence also suggests that sleep plays a role in early brain development of young children (Kocevska et al. 2017).

Figure 2.4. Overview of impacts of sleep duration on pre-school aged children

Adiposity	Psychosocial health	Cognitive development	Injury
• Higher body mass index (BMI)	• Improved self regulation • Improved relationships with others (e.g. peer acceptance, social skills, social engagement, understanding others emotions) • Adverse emotion responses • Risk factor for anxiety/depressive symptoms • Increased behavioural problems (e.g. anger, aggression, impulsivity, tantrums, externalising problems)	• Improved general conceptual ability (e.g. problem solving) • Improved vocabulary	• Increased medical attendance for injuries • Increased risk of accidental falls

(Note: evidence written in black text = associations with sufficient sleep duration, evidence written in brown text = associations with shorter/insufficient sleep duration).

Alongside sleep duration, sleep quality and timing/regularity of sleep, have been identified as important contributing factors to optimum health of pre-school children (Iwata et al., 2012). There are several indicators of sleep quality, with sleep continuity measures most frequently used to indicate quality from a public health perspective, these include: sleep latency, number of night-time wakings, wake after sleep onset and sleep efficiency, as well as naps. Table 2.2 displays classifications of 'good', 'uncertain' or 'poor' sleep quality for pre-school aged children for sleep continuity and nap indicators. These classifications were devised based on reviews of the literature and an expert panel to determine the National Sleep Foundation's (NSF) sleep quality recommendations (Ohayon et al., 2017). Sleep quality can also be determined through sleep architecture indicators which focus on the different stages of sleep rapid eye movement (REM) and non-REM sleep of different stages (Ohayon et al., 2017). There was no consensus among the expert panel as to what constituted differing levels of sleep quality for sleep architecture indicators for pre-school children. Sleep architecture requires examination of brain activity, which would include use of measures such as polysomnography, which would not be feasible for public health measurement. As such, sleep architecture will not be examined in this PhD.

Table 2.2. Indicators of sleep quality in pre-school children (Ohayon et al., 2017)

Sleep quality Indicator	Good	Uncertain	Poor
<i>Sleep continuity measures</i>			
Sleep latency	≤15 minutes, 16-30 minutes also classified as good.	31-45 minutes	46-60 minutes poor, more so when >60 minutes
Night wakings	≤1 per night	2-3 per night	≥4 per night
Wake after sleep onset	≤20 minutes	21-50 minutes	≥51 minutes
Sleep efficiency	≥85%	75-84%	≤74%
<i>Naps</i>			
Naps per 24 hours	No expert consensus on whether 0-3 naps was appropriate or not.		≥4
Nap duration	No expert consensus on nap duration for pre-school aged children		
Nap frequency	No expert consensus on nap frequency for pre-school aged children		

2.3.4 Standing Time

In addition to the key behaviours of PA, SB and sleep, there are additional behaviours that contribute towards the 24-hour cycle (Rosenberger et al., 2019). An example of such a behaviour is standing. Standing is difficult to classify, which has been displayed by uncertainty in the literature as to how standing should be defined for all age groups, ranging from active standing (may also be classified as light PA (LPA)), passive standing (which may encompass no movement), and frequently discussed in 'posture

allocation' literature (Davies, Reilly and Paton, 2012; Rosenberger et al., 2019; Tremblay et al., 2017a).

There is very limited evidence in any population on the association between standing and health outcomes, although, over the past two decades some key issues and areas for future exploration have arisen relating to posture allocation for the general population (not focusing on young children). These include: 1) Preliminary isotemporal analyses supporting the benefits of replacing sitting with standing and 2) the need to examine the effects of prolonged standing, as this may be negative for health in a similar way as prolonged sitting. The ideas around this include that there may be potential benefits from postural shifts, breaking up prolonged standing and sedentary periods may be beneficial for health (Biddle et al., 2019).

A limitation and challenge for research on standing in any age group is the measurement of this behaviour. There is much work to be done in the area of measurement of standing for all population groups, it is likely that the use of inclinometer-based accelerometers (e.g. ActivPAL) will help develop a better understanding of standing, how it contributes towards the 24 hour day, patterns and context of standing, and the associations between standing and health and developmental indicators (Biddle et al., 2019). The ActivPAL accelerometer has been shown to be able to identify posture allocation of pre-school children, but not always with optimal accuracy (Davies et al., 2012), whilst proxy reports of standing time by parents and carers of pre-school children are unlikely to correctly reflect the standing behaviours of their child (Davies, Reilly and Paton, 2012).

Although measurement is problematic in this area, research has identified that time spent in 'non-standard postures' including standing make up a very low proportion of pre-

schoolers daily activity, accounting for only 3% (Davies, Reilly and Paton, 2012).

Furthermore, there are no recommendations for 'standing' outlined in the new guidelines for pre-schoolers (Chief Medical Officers, 2019; Draper et al., 2020a; Ministry of Health New Zealand, 2017b; Okely et al., 2017; Tremblay et al., 2017b; WHO, 2019). Based on the available evidence thus far, and in light of the purpose of the current work to determine measures that could be used to monitor the recommendations, assessment of standing time of pre-schoolers will not be a focus of this research.

2.4 Health Inequalities

In this section I will discuss the broad context of health inequalities in the UK, followed by highlighting the significance and importance of the early years for improving health and reducing inequalities. I will then move on to outlining the available evidence on inequalities in movement behaviour opportunities, highlighting the need for representation of lower socioeconomic status (SES) groups in measurement, surveillance, and intervention studies, and ensuring that tools used to measure movement behaviours are equally appropriate for use by individuals across SES groups.

2.4.1 Health Inequalities in the UK

Health inequalities are a major public health issue in the UK (Public Health England, 2018). The relationship between social disadvantage and health is well documented (Marmot et al., 2010; Marmot et al., 2020a). Health is notoriously worse for individuals from low SES backgrounds in the UK (Marmot et al., 2010). In 2017-2019, individuals living in the most deprived areas in England, had a shorter life expectancy by an average of 7.6 years (females) and 9.4 years (males) than people living in the most affluent areas (Office for

National Statistics, 2021). Further, individuals living in the most deprived areas have much less years lived in a good health state, with healthy life expectancy being an average of 19.8 years (females) and 18.4 years (males) lower in those living in the most deprived areas in England, than those living in the most affluent areas (Office for National Statistics, 2021). Inequalities start from the early years, with children in more socioeconomically disadvantaged areas experiencing worse health than their advantaged peers, vast inequalities evident by the time children start school, and increased likelihood of poor mental and physical health outcomes in later life (Marmot et al., 2010; Marmot et al., 2020a; Pearce et al., 2019). The early years and giving children ‘the best start in life’ is highlighted as a key area of focus for improving health and reducing health inequalities (Marmot et al., 2010; Marmot et al., 2020a). Thus meaning that the early years are a crucial period in which inequalities can be addressed, as lifestyles and experiences established during this time impact upon health and well-being throughout the rest of a child’s life (Marmot et al., 2010; Marmot et al., 2020a). Evidence suggests that current policies continue to disadvantage young children from lower SES backgrounds, and do not have mechanisms in place to reduce inequalities for pre-school aged children (Griffin et al., 2021; Sutton Trust, 2021).

2.4.2 Inequalities of movement behaviour opportunities for pre-school aged children

Available evidence suggests that inequalities play a role in health across the 24-hour cycle for pre-school aged children. There is a ‘complex web’ of factors contributing towards activity and inactivity for children in their early years, that encapsulates community, economic, social, and personal factors; which may be more pronounced in deprived areas (Khanom et al., 2020). Children from lower SES backgrounds experience inequalities in

access and opportunity for PA (e.g. financial constraints, limited access to transport, limited access to safe outdoor spaces) (Hesketh, Lakshman and van Sluijs, 2017; Khanom et al., 2020). Further, a review of the literature identified that children from lower SES backgrounds and of non-white ethnicity spent more time in SB (Pate et al., 2011). Similar to this, children from a deprived community have been reported to have reduced sleep duration and longer sleep latency (Williamson et al., 2021; Wilson et al., 2014). Research has shown that pre-school children from more disadvantaged families and with parents with lower levels of education were less likely to have bedtime routines that facilitate healthy sleep than their more advantaged and educated counterparts (Hale et al., 2009). Alternative research has shown no differences in levels of PA, SB, or sleep duration between children from different SES groups (Jones and Ball, 2014; Kelly et al., 2006). Despite this, sleep hygiene behaviours (e.g. regular bedtimes) were still more prevalent in higher SES groups than lower SES groups (Jones and Ball, 2014). Additionally, the studies used different methods to categorise SES, which may explain and impact on the findings.

In terms of guideline adherence, recent research demonstrated that African American children and children who live at or below poverty level were less likely to meet the guidelines, in particular the sleep and sedentary screen time guidelines, than White or other race children or children not living in poverty (Kracht, Webster and Staiano, 2019). The analysis showed that no African American children met all three movement behaviour guidelines, and no children in the included sample at or below poverty level met all three guidelines, or the screen time guideline (Kracht, Webster and Staiano, 2019). Additionally, pre-school aged children from households with lower income and lower levels of education are less likely to meet movement behaviour guidelines, including engaging in higher screen time and less PA, than those from high income/education households (De Craemer et al.,

2020; Garriguet et al., 2016). Sleep deficiency in the form of not meeting recommendations and late and variable bedtimes has been demonstrated in children from families living in lower SES areas (Ordway et al., 2020).

However, these findings should be viewed with caution due to underrepresentation of lower SES groups in movement behaviour research (as described in section 2.10, p. 122 and section 2.11.4, p.133). Given the inequalities in movement behaviour opportunities, it is important to ensure that lower SES groups are appropriately represented in surveillance and intervention studies, and that tools used to measure movement behaviours in these studies are equally feasible and appropriate for use by individuals across SES groups.

2.5 Movement behaviour approach

In this section, I will explain the movement behaviour approach more broadly, including evidence and commentary on the movement behaviour approach that speaks to human movement generally, rather than focusing on specific age groups. I will explicitly state where the evidence is in relation to pre-school aged children.

2.5.1 Overview of movement behaviour approach

Recently, there has been a shift in thinking towards the ‘whole day matters’ with regards to movement behaviours for children in their early years, with the importance of examining PA, SB, and sleep in relation to each other, as opposed to in isolation as has previously been the case (Okely et al., 2018; Tremblay, 2020). Although it is clear that each of the movement behaviours has independent effects on health, the behaviours may also interact with one another in a way to impact health and development that cannot be determined by examining the behaviours in isolation (Rosenberger et al., 2019). Previously

research has focused on the impacts of PA on health; whilst PA provides important health benefits, it only accounts for a small part of the 24 hour day (estimated around 5%). By only focusing on PA, an estimated 95% of the day and the important behaviours that this encompasses (including sleep and SB, and an additional smaller proportion of LPA) are ignored (see figure 2.5, extracted from Chaput et al., 2014a). Ignoring the other movement engaged in during the day limits our understanding of how behaviours interact to impact children's health, and how to optimise health benefits of movement (Chaput et al., 2014a).

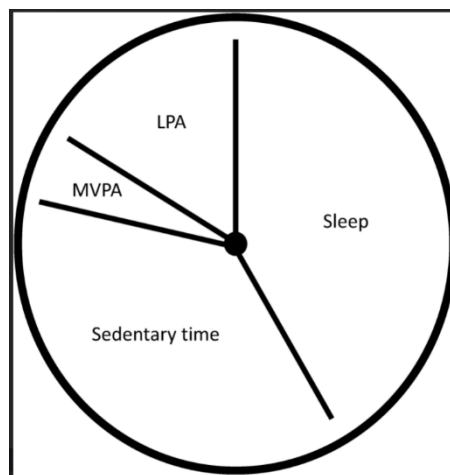


Figure 2.5. Estimated distribution of movement behaviours in a 24 hour day (Chaput et al., 2014a)

The rationale for viewing the whole 24 hour movement behaviours is nicely outlined by Rosenberger and colleagues (2019). The authors state that time spent in each behaviour affects the duration of time spent in the other behaviours (e.g. when engaging in PA, this has taken time away from engaging in other behaviours such as SB or sleep), which may modify the health-related influence of time spent in any other of the behaviours.

Consequently, by viewing the behaviours in isolation it is difficult to determine the extent to which a health benefit is as a result of change in one individual behaviour, or an alteration of the other behaviours that constitute the whole day (Rosenberger et al., 2019). With this,

there is a body of accumulating evidence across all age groups suggesting that combinations of the behaviours may impact health in a way that cannot be explained by exploring the behaviours independently (e.g. Buman et al., 2014; Chaput et al., 2014a; Chaput et al., 2014b; Chaput, Saunders and Carson, 2016; Chastin et al., 2015; Rao, Orpana and Krewski, 2016). For example, health benefits of engagement in moderate to vigorous PA (MVPA) may be lost if children have excessive SB or insufficient sleep (Chaput et al., 2014a; Chaput et al., 2014b).

Further to this, as research has moved on in this area and now examines relationships between meeting movement behaviour guidelines and health indicators there is preliminary evidence to suggest that the more guidelines that are met, for example, meeting all guidelines, as opposed to just one or two guidelines, are associated with most desirable health outcomes in pre-school children (Rollo, Antsygina and Tremblay, 2020).

This shift in the evidence base indicating that a broader, more integrated approach to understanding and promoting human movement, encapsulating the whole 24 hour day including PA, SB, and sleep is thought to be required to better address the current public health concerns, including physical inactivity and childhood obesity, and improve the health of children (Chaput et al., 2014a; WHO, 2017). The movement behaviour approach will help examine daily activity patterns, characterise the optimal 24-hour behaviour patterns for health, identify interrelationships between the behaviours and how combinations of the behaviours contribute to overall health and well-being (Rosenberger et al., 2019). A focus on the interrelationships between PA, SB, and sleep are important in public health messaging to emphasise that all of these behaviours matter for health of young children (Chaput, 2019). Preliminary evidence has demonstrated stakeholder and end user support for the

24-hour movement behaviour approach and associated guidelines for pre-school aged children (guidelines are explained in detail in section 2.6 of this thesis) (Okely et al., 2017; Reilly et al., 2020; Stanley et al., 2020; Tomaz et al., 2020).

At present, measurement issues are limiting the extent to which movement behaviours can be examined, interactions between the behaviours, and the influence on health and development. For example, a recent study examining associations between movement behaviours in pre-school children used accelerometer data to detect all behaviours (3 days, 24 hour wear protocol). Low movement throughout the day was regarded as SB, as such, it was unknown whether daytime napping occurred (and therefore subsequently how this may influence health) as no distinctions could be made between the behaviours during the period of measurement (Duraccio and Jensen, 2017).

2.5.2 Analysis of movement behaviour

Analysis methods predominantly used when exploring the 24-hour movement behaviour cycle across all age groups include isotemporal substitution analysis and compositional data analysis (Rosenberger et al., 2019). These methods acknowledge that the movement behaviours all take place during the same finite pool of time (the 24-hour cycle). Isotemporal substitution modelling involves exploring the associations of the impact of one behaviour on health, whilst holding the additional behaviours constant (e.g. Buman et al., 2014; Mekary et al., 2013). On the other hand, compositional data analysis is suited to understanding the best combinations of the behaviours on health outcomes. This is due to compositional data analysis dealing with the data as proportions of the 24-hour day, exploring how the behaviours make up the 24-hour day and how the behaviours interact with one another (Chastin, Dontje and Skelton, 2015). To date, much of the literature in the

area examines whether pre-school children have met the movement behaviour guidelines, and the relationships between this and health and developmental indicators; however, there is increasing evidence using compositional data analysis (Rollo, Antsygina and Tremblay, 2020).

2.5.3 Movement behaviour and health and development in pre-school aged children

A systematic review of studies identified that there was no available evidence examining the full spectrum of movement behaviours (PA, SB, and sleep) and health and development indicators in young children prior to 2016 (Kuzik et al., 2017). However, since the shift in thinking and policy move towards 24-hour movement behaviour assessment in young children, there is an increasing amount of evidence being produced on the topic. A systematic review by Rollo and colleagues, found that there was preliminary evidence suggesting that the composition of movement behaviours, in particular more sleep and less SB, was associated with better bone and skeletal health (Rollo, Antsygina and Tremblay, 2020). Additionally, pre-school children meeting the 24-hour movement guidelines demonstrated better health related quality of life, mental health indicators (lower depression scores), cognitive development, psychosocial skills (fewer behavioural and emotional problems), educational outcomes, and fitness (Carson et al., 2019; Chia, Tay and Chua, 2020; Cliff et al., 2017; Hinkley et al., 2020; Lemos et al., 2021; Rollo, Antsygina and Tremblay, 2020; Taylor et al., 2021). Whilst some evidence suggests there are no significant relationships observed between meeting the movement behaviour guidelines and motor skill development (Draper et al., 2020b; Kim et al., 2021), others have found improved gross motor development associated with meeting movement guidelines (Kracht, Webster and Staiano, 2020).

There has been some contention in the literature for the outcome of adiposity. The composition of movement behaviours has shown an association with adiposity, including children who spent 10% more time asleep relative to other behaviours was associated with a lower body mass index (BMI) score, whereas spending 10% more of the day sedentary or in LPA was associated with higher BMI scores. This was present only at the pre-school age group and no longitudinal effects were present (Taylor et al., 2018). There is inconsistent evidence regarding meeting the 24 hour movement guidelines and lower adiposity (Rollo, Antsygina and Tremblay, 2020). It is thought that the discrepancy was that comparing those meeting versus not meeting the guidelines may lack the sensitivity to detect differences in links between the 24 hour movement behaviours and health outcomes, but more evidence is needed before making firm conclusions (Rollo, Antsygina and Tremblay, 2020). Despite this, some research has demonstrated that pre-school children not meeting all three movement behaviour guidelines had a higher likelihood of being overweight or obese than those than meeting all three of the movement behaviours (Kim et al., 2020). An association has been found between pre-school aged children meeting all three guidelines and lower BMI in later childhood (Hinkley et al., 2020). Additionally, children who meet two or more of the guidelines have lower likelihood of being overweight or obese (as assessed by BMI), in particular the combination of meeting the PA and sleep duration guideline, than children who do not meet any of the guidelines (Lee et al., 2021; Leppänen et al., 2019). Meeting the sleep guidelines in particular were most pertinent for associations with lower BMI, however no associations between sedentary time and weight indicators were identified (Leppänen et al., 2019). Further research has demonstrated that there was no significant relationship between children from lower income countries meeting the guidelines and adiposity (Kim et al., 2021) emphasising the importance of cultural factors in relationships between

movement and health, as well as other important health behaviours such as having access to a nutritious and balanced diet (Brown et al., 2019).

2.6 Movement behaviour guidelines and recommendations for pre-school aged children

Government health guidelines, such as PA, sleep and movement behaviour guidelines, are developed through rigorous processes; primarily guided by systematic reviews of the literature, alongside input from topic experts from academia, and health professionals (Draper et al., 2020a; Reilly et al., 2020; Willumsen and Bull, 2020). The recommendations are a reflection of the evidence on the health and developmental benefits of engaging in appropriate amounts of each behaviour (Carson et al., 2017; Chaput et al., 2017b; Poitras et al., 2017; Hirshkowitz, 2015a). Guidelines are primarily aimed at policymakers, health care professionals, and practitioners (Bull et al., 2002; Chief Medical Officers, 2019; WHO, 2020). Guidelines can also be a valuable educational tool, to increase awareness on the importance of movement behaviours (Draper et al., 2020a). Public knowledge and education of the guidelines has been associated with higher adherence, particularly for individuals from lower SES backgrounds (Roth and Stamatakis, 2010). Despite this, literature has identified that parents of children in their early years are not always aware of the guidelines (Bentley, Jago and Turner, 2015; Riazi et al., 2017). Guidelines can also provide a standard to help monitor the prevalence of movement behaviours in young children (e.g. national surveillance, research) to further inform policy and practice.

2.6.1 Global movement behaviour guidelines for pre-school aged children

2.6.1.1 *Physical activity and sedentary behaviour guidelines for pre-school aged children*

PA and SB guidelines developed specifically for children in their early years (0-5 years old) do not have a long history. A review conducted in 2012 examining PA guidelines for pre-schoolers, reported that there were only three countries (Australia, USA, UK) that had PA guidelines for young children (Skouteris et al., 2012), with the Australian government publishing the first endorsed guidelines in 2009 (as cited in Skouteris et al., 2012, original document no longer available). Only the Australian and UK guidelines were government endorsed, whilst the USA guidelines were a product of the National Association for Sport and Physical Education (NASPE) (National Association for Sport and Physical Education, 2011), which were said to be the 2nd edition, with the first edition published in 2002 (although no records can be accessed for this). Guidelines for early years children shortly followed by Canada (Tremblay et al., 2012). At this time, the guidelines (UK, USA, Australia, Canada) were consistent in the recommendation for 180 minutes of PA of any intensity per day and to minimise SB. The main difference was that the Australian and Canadian guidelines also recommended limiting screen time to <1 hour per day. The USA guidelines differed most, suggesting that pre-school children should engage in at least 60 minutes of structured PA per day, engage in at least 60 minutes and up to several hours of unstructured PA, and should not be sedentary for more than 60 minutes at a time. The increase in the research and knowledge regarding the activity behaviours of children in their early years since this time has led to revised guidelines for children of this age (see section 2.6.2.1, p. 63), including the importance of MVPA for pre-school aged children.

2.6.1.2 *Sleep guidelines for pre-school aged children*

Sleep duration guidelines for pre-school aged children have been promoted by the National Sleep Foundation (NSF) in the USA since 2004. These guidelines initially suggested 11-13 hours of sleep for children aged 3-5 years old, which were updated in 2015 to suggest 10-13 hours (Hirshkowitz et al., 2015a; Hirshkowitz et al., 2015b). The 10-13 hour sleep duration guideline for pre-school aged children was also independently endorsed by the American Academy of Sleep Medicine (USA) in 2016 (Paruthi et al., 2016). Sleep guideline development was conducted by multidisciplinary expert panels (Hirshkowitz et al., 2015b; Paruthi et al., 2016), using the following procedures:

- 1) The available scientific evidence (systematic/detailed reviews examining sleep in relation to various health and developmental outcomes e.g. cognition, obesity),
- 2) Grading appropriateness of sleep time durations (stating the appropriateness of the sleep time duration in terms of whether the health benefit would exceed any negative outcomes)
- 3) Discussions to determine appropriate sleep time duration recommendations.

At present, sleep duration guidelines provide ranges of recommended sleeping hours, implying that there is a U-shaped association between sleep duration and health and developmental indicators; both short and long sleep duration associated with unfavourable health impacts (Chaput, Dutil and Sampasa-Kanyinga, 2018; Paruthi et al., 2016).

As outlined in a recent systematic review, there is a limited and weak evidence base that is currently being used to inform sleep duration recommendations for pre-school aged children (Chaput et al., 2017). Sleep duration guidelines are beneficial for public health to act as guidance at population level, however more high quality evidence examining dose-

response relationships between sleep duration and health outcomes are required to improve the evidence base on which the guidelines are developed (Chaput, Dutil and Sampasa-Kanyinga, 2018).

2.6.2 History of UK physical activity and movement behaviour guidelines for pre-school aged children

In 2011, the first UK PA guidelines for under 5's were published, which included advice on the amount of PA and SB young children should engage in (Department of Health and Social Care, 2011). Following this, in 2018, the UK drafted 24-hour movement guidelines for the early years (Reilly et al., 2018). An intensive process was conducted by an expert committee to update and develop the recommendations, with the committee recommending 24 hour movement behaviour guidelines. However, this was not accepted by the UK Chief Medical Officer's (Reilly et al., 2018; Reilly et al., 2020). The current UK government guidelines for pre-school aged children focus on PA only (Chief Medical Officers, 2019). Figure 2.6 displays a timeline of the PA guidelines for pre-school aged children in the UK.

2.6.2.1 *Movement behaviour guidelines for pre-school aged children*

Until recently, there was a main focus on PA, and more recently SB, and guidelines did not take into consideration other important behaviours across the movement continuum that occur throughout the day, that can impact on and moderate health (Chaput et al., 2014a). Since the initial introduction of PA and SB guidelines for young children, the shift in thinking pertaining to the 24 hour movement behaviour approach has occurred (as described in section 2.5, starting p. 55). This shift has led to changes in the recommendations for pre-school aged children, moving towards a 24-hour movement

approach, rather than viewing the behaviours in isolation. In 2019, the WHO published their first guidelines for children in their early years, which were also the first set of 24 hour movement behaviour guidelines (WHO, 2019). This was following countries across the globe that had implemented 24 hour movement guideline approach for pre-school age children: Canada (Tremblay et al., 2017b), Australia (Okely et al., 2017), New Zealand (Ministry of Health New Zealand, 2017b), South Africa (Draper et al., 2020a). In 2018, the USA also included pre-school specific guidelines in their government endorsed recommendations, however, like the UK these still only include PA and do not extend guidelines to other movement behaviours of SB and sleep (US Department of Health and Human Services, 2018). See table 2.3 for an overview of the guidelines in each of the countries.

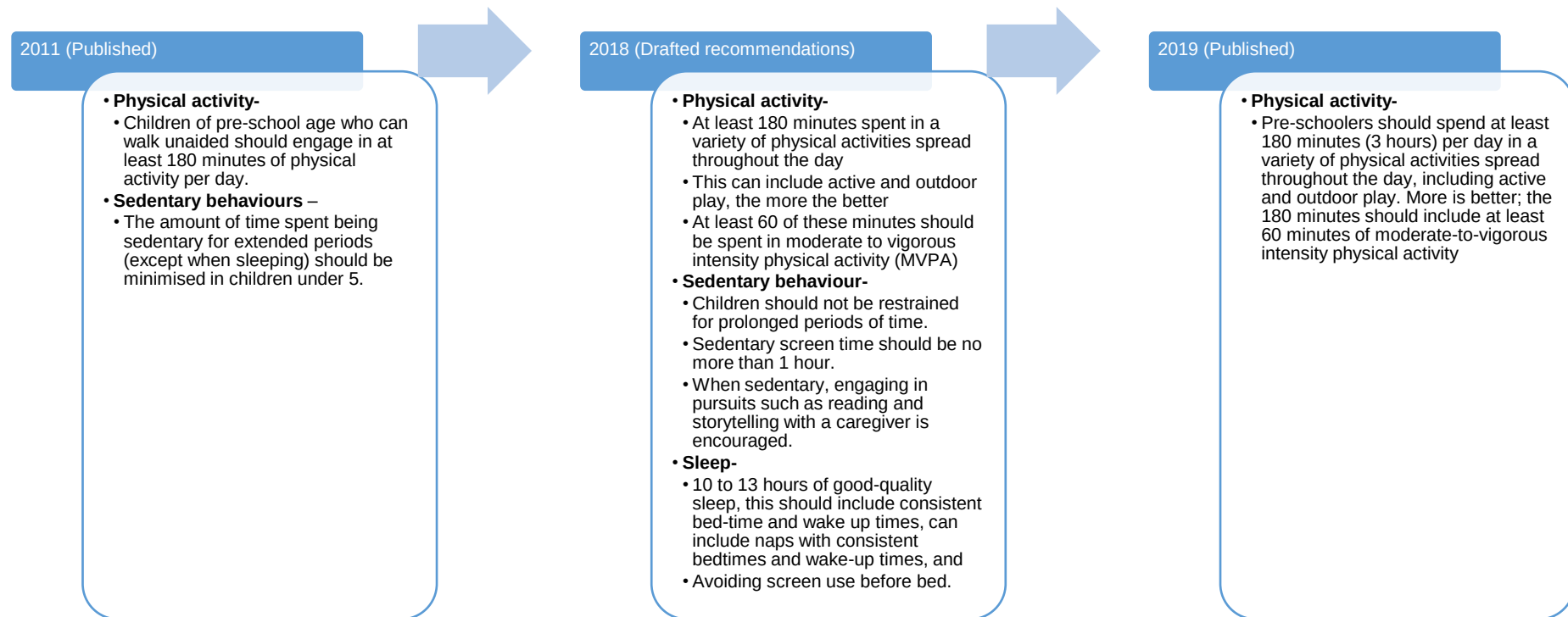


Figure 2.6. Timeline of UK physical activity guidelines for pre-school children

Table 2.3. Current global physical activity and 24 hour movement behaviour guidelines for pre-school children

Country	Year	Guidelines
24 hour movement behaviour guidelines		
Canada (Tremblay et al., 2017b)	2017	Physical Activity- - At least 180 minutes spent in a variety of physical activities spread throughout the day, of which at least 60 minutes is energetic play—more is better. Sedentary Behaviour- - Not being restrained for more than 1 hour at a time (e.g., in a stroller or car seat) or sitting for extended periods. - Sedentary screen time should be no more than 1 hour—less is better. - When sedentary, engaging in pursuits such as reading and storytelling with a caregiver is encouraged. Sleep- 10 to 13 hours of good-quality sleep, which may include a nap, with consistent bedtimes and wake-up times.
Australia (Okely et al., 2017)	2017	Physical activity- - Spend at least 180 minutes a day in a variety of physical activities, of which 60 minutes is energetic play such as running, jumping and kicking and throwing, spread throughout the day - noting more is better. Sedentary Behaviour- - Should not be restrained, for more than 1 hour at a time e.g. in a stroller or car seat) or sitting for extended periods. - Sedentary screen time should be no more than 1 hour in total throughout the 24-hour period -less is better. - When pre-schoolers are sedentary, caregivers are encouraged to engage with them through activities such as reading, singing, puzzles and storytelling. Sleep- 10 to 13 hours of good quality sleep, which may include a nap, with consistent sleep and wake-up times.
New Zealand (Ministry of Health New Zealand, 2017b)	2017	Physical activity- - Provide fun activities that support physical, social, emotional and spiritual growth (at least three hours every day for toddlers and preschoolers, spread throughout the day). Include plenty of opportunities for active play: - that develop movement competence and confidence

		• that provide sufficient challenges to build resilience and encourage creativity through exploration • where children are by themselves as well as interacting with others, such as parents, siblings, friends, whānau/family and other caregivers • that include a variety of indoor and outdoor activities, especially activities involving nature. Sedentary time- • Provide regular activity breaks to limit the amount of time a child spends sitting • Discourage screen time for under-two-year-olds and limit screen time to less than one hour every day for children aged two years or older – less is best! • Limit time in equipment that restricts free movement. Sleep- • Should have 10 to 13 hours of good-quality sleep every day, with consistent bedtimes and wake-up times.
South Africa (Draper et al., 2020a)	2019	Physical activity- • At least 180 minutes spent in a variety of physical activities, of which at least 60 minutes is very active play, spread throughout the day; more is better. Sedentary behaviours- • Engaging in activities such as reading, singing, puzzles, arts and crafts, and storytelling with a caregiver and other children. • NOT being strapped in and unable to move for more than 1 hour at a time (e.g., in a pram or car seat) or sitting for extended periods. • Screen time while sitting or lying down should be no more than 1 hour; less is better. Sleep- • 10 to 13 hours of good quality sleep, which may include a nap, with consistent sleep and wake-up times.
World Health Organization (WHO, 2019)	2019	Physical Activity- • At least 180 minutes spent in a variety of physical activities spread throughout the day, of which at least 60 minutes is energetic play—more is better. Sedentary Behaviour- • Not being restrained for more than 1 hour at a time (e.g., in a stroller or car seat) or sitting for extended periods. • Sedentary screen time should be no more than 1 hour—less is better.

		- When sedentary, engaging in pursuits such as reading and storytelling with a caregiver is encouraged. Sleep- 10 to 13 hours of good-quality sleep, which may include a nap, with regular sleep and wake-up times.
Physical activity guidelines		
USA (US Department of Health and Human Services, 2018)	2018	Physical activity- - Should be physically active throughout the day to enhance growth and development. - Adult caregivers of preschool-aged children should encourage active play that includes a variety of activity types.
UK (Chief Medical Officers, 2019)	2019	Physical activity- 180 minutes (3 hours) per day in a variety of physical activities spread throughout the day, including active and outdoor play. More is better. - This should include at least 60 minutes of moderate-to-vigorous intensity physical activity.

2.6.3 Guideline adherence among pre-school aged children

Research has demonstrated that very few pre-school children meet all components of the WHO movement guidelines (Carson et al., 2019; Chaput et al., 2017a; Chen et al., 2019; Chia, Tay and Chua, 2020; Cliff et al., 2017; De Craemer et al., 2018; de Lucena Martins et al., 2021; Draper et al., 2020b; Guan et al., 2020b; Hinkley et al., 2020; Hossain et al., 2021; Kim et al., 2021; Kim et al., 2020; Kracht, Webster and Staiano, 2020; Lee et al., 2021; Leppänen et al., 2019; Lucena Martins et al., 2020; McNeill et al., 2020; Rollo, Antsygina and Tremblay, 2020; Vale and Mota, 2020). Generally, children meet the PA and sleep duration guidelines, but not the sedentary screen time guidelines (Chaput et al., 2017a; Chia, Tay and Chua, 2020; Cliff et al., 2017; Draper et al., 2020b; Hinkley et al., 2020; Hossain et al., 2021; Kim et al., 2020; Kracht, Webster and Staiano, 2020; Lee et al., 2021; Leppänen et al., 2019; McNeill et al., 2020; Meredith-Jones et al., 2019).

Although the evidence is largely consistent, there is some variation, with few studies (with samples of children from Belgium and Canada) reporting that low proportions (11-19.3%) of children meet the PA guidelines (Carson et al., 2019; De Craemer et al., 2018). The authors of the studies suggested that different measurement tools and interpretation choices (e.g. use of 60s epoch, varying cut points) were likely reasons for the reported differences (Carson et al., 2019; De Craemer et al., 2018). This further highlights the importance of measurement in the outcomes of studies examining the behaviours and guidelines. Furthermore, two studies conducted in China demonstrated much higher compliance with the sedentary screen time guideline (70.2-88.2%), but lower compliance with the sleep duration guideline (13.7-29.5%) possibly suggesting cultural differences in compliance with specific guidelines (Chen et al., 2019; Guan et al., 2020b).

In the literature, guideline adherence primarily focuses on the recommendations of: 180 minutes of PA, 60 minutes of which is MVPA; sleep duration guideline, and <1 hour screen time (WHO, 2019). Very few studies to date have reported data on other aspects of the guidelines in pre-school children, including the quality of sleep, consistent sleep and wake-up times, period of time restrained / sitting for extended periods of time, or SB whilst engaging in pursuits such as reading and storytelling. Those that have, reported that only a quarter of children in their sample have consistent (calculated as within 30 minutes) sleep and wake times (Meredith-Jones et al., 2019).

Further to this, it was suggested by the UK early years expert committee that avoiding screen time one hour before bed time should be included in the sleep guidelines for pre-school age children (Reilly et al., 2018). Studies that have examined this alongside other movement behaviour guidelines have reported that high proportions of children

perform screen based activities close to sleep time (Draper et al., 2020b; Hossain et al., 2021).

At present, there are no published data available on the proportion of pre-school children in the UK meeting the WHO movement behaviour guidelines. The most recent national data for PA estimates of pre-school aged children in England was in 2015 from the Health Survey for England. Based on this proxy reported data, only 9% of children aged between 2 and 4 years met the previous (2011) guideline of at least 180 minutes of PA per day (NHS Digital, 2016). There are no data available for the 60 minute MVPA recommendation, as no national data has been collected since the publication of the new guidelines. Active Lives Children, a yearly national assessment of self reported PA levels in the UK, does not collect data for pre-school aged children (Sport England, 2020). However, academic literature, at a similar time to the most recent Health Survey for England data, suggested that high proportions of pre-school aged children meet the 180 minute PA guideline, but do not always meet the 60 minute MVPA guideline (Hesketh et al., 2014).

2.7 Brief overview of physical activity and sedentary behaviour measurement tools

In this section I briefly outline the measurement tools that can be used to assess PA and SB of pre-school children, including a description of how each tool is used and the primary purpose. I will describe each of these tools generally, drawing specifically on evidence of the application of the tools for use with pre-school aged children where available. The measurement tools used to assess PA and SB are grouped together, as similar methods are often used to assess these behaviours. The tools I will discuss include:

- Calorimetry (direct, indirect, and doubly labelled water)
- Direct observation

- Wearable cameras
- Heart rate monitors
- Accelerometers
- Combined heart rate and accelerometry
- Pedometers
- Consumer wearables
- Proxy reported measurement tools (questionnaires, diaries, and ecological momentary assessment)

Following discussion of each of these measurement tools, I present a summary table of the strengths and limitations of each of the tool (see Table 2.5). A detailed systematic review outlining the validity, reliability, feasibility of measurement tools used to assess PA and SB of pre-school aged children can be found in *Chapter 3* of this thesis.

2.7.1 Calorimetry

Calorimetry is considered a criterion method of measuring energy expenditure for all age groups, including pre-school children, classified as the most accurate method for identifying the amount of energy an individual expends during various activities. Calorimetry measures energy expenditure, an outcome of PA rather than the behaviour itself (Sirard and Pate, 2001). This method does not provide information on free living or habitual activity of an individual, and is not useful for determining frequency, duration, type, context or domain of PA. These methods are therefore most often, and most suitably used, for determining the energy expenditure of certain activities, or validity of other measurement tools used to assess energy expenditure aspects of PA and/or SB. There are two main methods of calorimetry: direct and indirect.

2.7.1.1 Direct Calorimetry

Direct calorimetry is a method used to determine energy expenditure of an individual. Whole room calorimetry is an example of direct calorimetry. This consists of an individual being in a calorimeter (a sealed chamber/controlled room), whereby the conditions are held constant, and the heat produced by the body is used to determine the energy cost of all activities carried out. This method is most often used in validation studies, when determining the validity of a new tool such as an accelerometer for use with pre-school aged children (e.g. Janssen et al., 2013a).

2.7.1.2 Indirect Calorimetry

Indirect calorimetry can be assessed through gas exchange, by continuously measuring the flow of oxygen consumed and carbon dioxide produced by the body to determine energy expenditure. This can be conducted using portable metabolic systems, making it more flexible to direct calorimetry, in that it is not necessary for individuals to be confined to a controlled room, but still not suitable for free living measurement due to the equipment used. An example of this used with pre-school children is the Cosmed portable metabolic system (Model K4b2; Cosmed, Rome, Italy) (e.g. Pfeiffer et al., 2006). This consists of a harness worn on the back, and a face mask fitted to ensure that all breath output is captured. All expired respiratory gases are collected, and the oxygen consumption is measured on a breath by breath basis to determine the amount of energy expenditure (e.g. Pate et al., 2006).

2.7.1.3 *Doubly labelled water*

Indirect calorimetry is most frequently conducted using doubly labelled water (DLW). This method involves an individual consuming water loaded with stable isotopes ^{18}O and deuterium ^2H . The deuterium ^2H is eliminated from the body as water only, whilst the ^{18}O isotopes are eliminated from the body as both carbon dioxide and water. The difference between the elimination rates of the isotopes and deuterium are a product of the rate of carbon dioxide production. The urine or saliva samples collected from the individual, usually daily for 7-14 days are used to calculate estimates of carbon dioxide production. The rate and amount of carbon dioxide production is used to determine energy expenditure using analytical tools such as an isotope ratio mass spectrometer (Westerterp, 2017). This method has also frequently been used as a comparison method in validation studies of accelerometers to determine energy expenditure of pre-school aged children (e.g. Delisle Nyström et al., 2017; Lopez-Alarcon et al., 2004).

2.7.2 *Direct Observation*

Direct observation consists of individuals engaging in their usual activity, whilst being directly observed by one or two trained observers. Observers document the PA and SB that is taking place; these are usually conducted using a time sampling method (predominant activity taking place each minute is recorded) or sometimes continuous (every change in activity is recorded, regardless of frequency of changes). Direct observation usually takes place in controlled settings, most often the child's nursery, in the playground of the nursery, or in some cases when the child is at home. Direct observation is classed as the best available reference method for measuring frequency, duration, type and domain of PA and SB (Sirard and Pate, 2001). However, direct observation usually takes place for a very

short period of time due to this being very intensive and burdensome for the researcher. It is therefore most often used in validation studies when determining the utility of other measurement tools, and is not feasible or practical for the assessment of these behaviours at scale.

There are a range of direct observation protocols that can be used to determine PA and SB of pre-school children. Those that have been developed with a sample that has included pre-school children include: Children's Activity Rating Scale (CARS) (Puhl et al., 1990), Children's Physical Activity Form (CPAF) (O'Hara et al., 1989), System for Observing Fitness Instruction Time for Preschoolers (SOFIT-P) (Sharma et al., 2011), Observation System for Recording Physical Activity in Children- Preschool (OSRAC-P) (Brown et al., 2006) and Fargo Activity Time Sampling (FATS) (Klesges, 1984). However, this is not always the case, and sometimes studies devise their own protocols dependent on the outcomes of interest (e.g. Sharp et al., 2017).

2.7.3 Wearable cameras

Wearable cameras, cameras worn on a lanyard around the individual's neck, can be used to measure PA and SB by taking images of the participant's front facing environment thousands of times per day. These can be used to determine contextual information about a participant's day and the activities that they engage in, including the social and environmental context in which activity occurs (Doherty et al., 2013b). Thus far, wearable cameras have primarily been assessed for feasibility in adults and school-aged children (Davies et al., 2020; Everson et al., 2019; Kelly et al., 2012) and used to understand the context in which PA and SB occurs (Doherty et al., 2013b). Example wearable cameras used

in PA research include Microsoft SenseCam and The Vicon Revue (Doherty et al., 2013a), however these are no longer commercially available (Downing, Janssen and Reilly, 2019).

To my knowledge, there are currently no studies examining the use of wearable cameras in pre-school children. However, some work is underway to assess the feasibility of wearable cameras (the PR6 wearable camera) for assessing SB, in particular screen time and time spent restrained, in a sample of pre-school aged children (Downing, Janssen and Reilly, 2019).

2.7.4 Heart Rate Monitors

Heart rate is a physiological method of measuring PA, recording heart beats to estimate how much energy is being exerted. Heart rate monitors measure the physiological parameters of PA and give a continuous recording that may reflect both intensity and duration. However, heart rate can only be used as an indirect indicator of PA. This is due to heart rate being affected by so many other factors, including: psychological stimuli (anxiety, stress), and environmental factors (temperature). Changes in heart rate cannot be truly attributable to engagement in PA, and an individual's heart rate could be affected even when sedentary. Heart rate monitoring, alone or combined with accelerometry, can be susceptible to 'noisy' heart rate data which can be as a result of poor electrode to skin contact which may impact on the estimates of movement given (Corder et al., 2008).

There are various types of heart rate monitors used to assess activity of pre-school aged children:

1. **Electrocardiography** – electrodes attached directly to the chest that measure the electrical potential of the skin to detect heart beats and are normally stored as

average heart beats per minute. Sometimes the device on the chest can store the data directly, whilst other devices have a wireless receiver often attached to the wrist that stores the data. Example: Polar Vantage XL (Bar-Or et al., 1996).

2. **Optical heart rate** – through the use of photoplethysmography (blood flow) sensors, able to detect heart rate based on fluctuations in blood flow occurring during a cardiac cycle. Example: Polar OH1 (Polar Electro Inc., Bethpage, NY, USA) (worn on the arm) (Sakzewski et al., 2021).

2.7.5 Accelerometers

Accelerometers are one of the most commonly used ways of measuring PA, including in pre-school aged children (Cain et al., 2013). Accelerometers are wearable devices, used to measure acceleration of the body resulting from human movement. Acceleration is the change in velocity overtime, which can be used to determine the volume and intensity of movement of the body. Accelerations from the accelerometer can be translated into meaningful information about the PA and SB of an individual. Although accelerometers have previously been referred to as an ‘objective’ way of measuring PA and SB, this method is fraught with methodological challenges and interpretation decisions that a researcher must make that can impact on the outcomes of studies (Banda et al., 2016; Ojiambo et al., 2011; Van Cauwenberghe et al., 2011b). Further, device based tools are subject to technical (e.g. inconsistencies in capturing signals) and human (e.g. accidental altering of the position of the device) sources of random errors (Pedišić and Bauman, 2015). There is currently no consensus on decision protocols for accelerometers in most age groups including pre-school children, making comparisons between the outcomes of studies

difficult. The decisions to be made when using accelerometry are outlined below, and will be discussed in turn in the subsequent sections:

1. Which device should be used?
2. Device placement - where on the body should the device be worn?
3. At what sampling frequency should the accelerometer collect the data?
4. What epoch length should be used? Or should raw acceleration data be used?
5. Which intensity cut points should be used?
6. How long should the device be worn for, and what is classified as 'wear time' and 'non wear time'?

2.7.5.1 *Types of Accelerometer*

There are a range of different types of accelerometer, with more continuously becoming available. One way in which accelerometers can be categorised is by the direction in which they can measure movement. The majority of accelerometers measure acceleration by piezo-electric sensors, these sensors can detect accelerations in the three axes of human movement: vertical, anteroposterior, and medio-lateral. However, there are four categories of accelerometers:

1. **Uniaxial** – measures activity in one plane: vertical

Examples of accelerometers: Actigraph GT1M, Actigraph 7164, Actiwatch Spectrum, Actiwatch-L, ActivPAL

2. **Biaxial**- measures movement in two planes: vertical and medio-lateral

Examples of accelerometers: Actiwatch AW16

3. **Triaxial**- measures movement in the three main axes of human movement: vertical; medio-lateral and antero-posterior.

Examples of accelerometers: Actigraph GT3X (+), Actigraph Wgt3x-BT, Stayhealthy RT3 (formerly Tritrac-R3D), Tracmor D Triaxial, GeneActiv

4. **Omnidirectional** – measures movement in all planes. This involves an accelerometer that is fixed at one end and free-moving at the other, which allows the sensor to rotate freely during movement and permits the measurement of PA in unspecified directions.

Examples of accelerometers: Actical

2.7.5.2 *Placement of Accelerometer*

The placement of the accelerometer is primarily dependent on the type of accelerometer. In some instances, accelerometers have to be placed on a very specific area of the body in order for them to capture the movement in the way they intend to. For example, the activPAL must be placed on the right thigh, midway between the hip and the knee (Janssen et al., 2013b); the Actiheart must be placed on the chest, as they are dependent on heart rate alongside acceleration (Hesketh et al., 2014). However, when measuring PA and SB of young children, the majority of accelerometers are worn on the hip, often the right hip (Miguelles et al., 2017), as a monitor as close to the body's center of mass is preferred to more accurately detect overall movement (Trost, McIver and Pate, 2005). In recent years, more accelerometers are being developed and tested using wrist placement, in part due to increased compliance with this wear location (Fairclough et al., 2016). Research suggests that both wrist and hip placement may be equally as accurate at classifying PA of pre-school children when using Actigraph accelerometers (Miguelles et al., 2017), with high correlation between placements for counts per minute and total count data (Hislop et al., 2016).

Some key considerations when deciding the placement of the accelerometer are:

1. The type of accelerometer and the location in which the device was validated (e.g. ActivPAL must be on the thigh; Actigraph can be wrist/hip worn).
2. Accuracy for the intended study design (e.g. if the research interest is to explore one movement behaviour or all behaviours then different locations may be used). For example, if assessing PA only, a hip placement may be most appropriate as most corresponding cut points for accelerometers were developed using this placement (Migueles et al., 2017). Whilst a wrist placement is most commonly used for assessing sleep only (as will be discussed in section 2.8.5.2, p.103). For assessing the full 24 hour movement behaviours, to my knowledge, the 'optimal' placement of an accelerometer for pre-school aged children has still not been determined. At present, studies with pre-school children examining the whole 24 hour movement behaviours using a single accelerometer have used a hip placement (Kuzik et al., 2020; Taylor et al., 2018).
3. Feasibility (e.g. which location is more likely to be complied with for study duration). Research shows wrist placement is associated with higher compliance than hip in school aged children (Fairclough et al., 2016). If a device is complied with to a greater extent there are improved chances of achieving a sufficient amount of data to obtain a representation of an individuals' activity.

2.7.5.3 Sampling Frequency of Accelerometer

The sampling frequency is another important consideration of accelerometry, to be determined prior to data collection. Sampling frequencies are the average number of samples recorded per second on the accelerometer. These can range from 1 to 100 Hertz

(hz), depending on the accelerometer. For example, older models of the Actigraph accelerometer (Actigraph GT1M and GT3X) had a default sample frequency of 30 hertz (hz), which could not be amended. However, the updated version (Actigraph GT3X+), allows the user to determine the sampling frequency from 30hz-100hz, in 10hz increments (Actigraph Manual, 2013). Higher sampling frequency rates are more accurate at capturing movement (Migueles et al., 2017). It has been suggested that the default 30Hz or its multiples provide more accurate estimates of PA when using the Actigraph (Migueles et al., 2017). Given this knowledge, 90z may be the most appropriate. Despite this, another aspect to consider when choosing sampling frequencies is the battery life and memory of the device (Hounslow et al., 2019). For most studies aiming to assess free-living behaviours of pre-school children, a measurement period of 7 continuous days is warranted. Therefore, there has to be a trade off between the highest sampling frequency possible with the device, and the battery life needed to capture the amount of data that is warranted for the purpose of the study .

Although it is important to be transparent in the sampling frequencies used in studies, as it has been recognised that the sampling frequency affects the generation of counts (e.g. >90 counts per minute detected for a slow walk when using the 40Hz compared to 30 Hz, a difference that increased with intensity of activity) (Clevenger et al., 2019; Brønd and Arvidsson, 2016). However, to date, research has suggested that the sampling frequency does not impact upon the overall classification of data into activity intensity levels in pre-school children (Clevenger et al., 2019).

2.7.5.4 *Filter*

Newer versions of accelerometer allow the possibility to apply different filters. The two filters available at this time are 1) normal or 2) low-frequency extension. The low-frequency extension increases the devices sensitivity to lower intensity activities, to more readily detect movement as measurement becomes more sensitive (Hjorth et al., 2012). To date, no data about the influence of the filter is available for the pre-school population (Migueles et al., 2017). More research is needed in this area having confidence to apply the different filters; as such, it is recommended to use the normal functions at this time (for pre-school children, children and adults) (Migueles et al., 2017).

2.7.5.5 *Epoch Length and Raw Acceleration*

Raw accelerations obtained from bodily movement are stored in the accelerometer as raw counts. The data collected are filtered by the monitor into user-specified time sampling intervals, called 'epochs'. At the end of each epoch (e.g. every 15 seconds), the summed value is stored on the monitor as 'activity counts'. Epoch length is another important interpretation choice that needs to be made when using an accelerometer. The length of the epochs vary, but are usually from 1 second to 60 seconds. Shorter epoch lengths (<15 seconds) are recommended for detecting PA and SB of pre-school aged children due to the sporadic and intermittent nature of their movement. In particular, 5 seconds is thought to be the most appropriate and feasible epoch length to accurately capture activity behaviours of young children, as with longer epoch lengths (e.g. 60 seconds) certain activity may be missed or incorrectly classified (Migueles et al., 2017).

Research has shown that using epoch lengths of varying lengths can provide different outcomes e.g. differing amounts of time engaged in different intensities of activity.

Data collected in 15 second epochs captured more MVPA and sedentary time than 60 second epochs when using the same data, as the longer epochs may not detect the shorter bursts of activity that are usually seen in young children (Colley et al., 2014; Kim et al., 2013). Similarly, Vale and colleagues found that using a 5 second epoch resulted in substantially more MVPA than the 60 second epoch (Vale et al., 2009).

In recent years, there have been technological advances in the use of accelerometry data, which has led to the ability to use raw acceleration data obtained from accelerometers (as opposed to summarising the acceleration data into activity counts/epochs). Raw acceleration may contain information such as direction and orientation of movement that may not be detected when the data is grouped into activity counts (Montoye et al., 2018). Using raw data may be beneficial as this means that data from different studies using different types of accelerometer may be compared more easily, necessary for data harmonisation. To my knowledge, GGIR software is the first non-commercial available software to help with processing and extraction of movement behaviour variables using raw data from three makes of accelerometer (GENEActiv, Actigraph, Axivity) (Migueles et al., 2019b).

2.7.5.6 *Cut Points*

Cuts points are thresholds used to determine the intensity of activity from accelerometers, usually from activity counts, to transform data into meaningful categories such as the amount of time spent in different intensities of activity. Calibration studies are used to develop and calibrate the cut points, using the accelerometer alongside a criterion measure (such as direct observation or calorimetry), to determine the thresholds at which different intensities of activity occur.

When analysing accelerometer derived data, the cut points used are a fundamental consideration. The cut points used can impact substantially on the results of a study including significantly influencing classification of intensity of activity and observed compliance with guidelines of young children (Ojiambo et al., 2011; Van Cauwenberghe et al., 2011b). Additionally, research has shown that cut points can have a higher influence on final estimations of activity behaviours than other considerations such as placement of the device (Migueles et al., 2019a) .

A host of cut points have been developed for use with pre-school children (Adolph et al., 2012; Evenson et al., 2008; Johansson et al., 2015; Pate et al., 2006; Pfeiffer et al., 2006; Puyau, Vohra and Butte, 2002; Reilly et al., 2003; Sirard et al., 2005; Van Cauwenberghe et al., 2011b). Table 2.4 shows a sample of some of the most frequently used cut-points in studies with pre-school children along with information on the studies where these were developed, to demonstrate how these can vary.

Table 2.4. Sample of calibration studies of accelerometer cut points for determining physical activity and sedentary behaviour of pre-school children

Authors	Sample	Details of accelerometer used to determine cut points (Type; placement; filter; epoch).	Criterion Method Used	Cut Points (15s epoch length)
Pate et al., 2006	N=29; 3-5 years Mean age 4.4 years 16 girls, 13 boys	Actigraph (MTI 7164); Right hip; Filter not reported; 15s	Indirect calorimetry (Cosmed portable metabolic system; Model K4b2; Cosmed, Rome, Italy)).	SB ≤37 LPA 38-419 MVPA ≥420
Evenson et al., 2008	N=33; 5-8 years; Mean age 7.3 years 21 girls, 12 boys	Actigraph (MTI 7164); Right hip; Filter not reported; 15s	Indirect calorimetry (Cosmed portable metabolic system; Model K4b2; Cosmed, Rome, Italy)).	SB ≤25 LPA 26-573 MVPA ≥574
Sirard et al., 2005	N=5; 4 years; N male and female unknown	Actigraph (MTI); Right hip; Filter not reported; 15s	Direct observation (modified CARS)	SB≤363 LPA 364-811 MVPA ≥812
Van Cauwenberghe et al., 2011b	N=18; Age = 4-6 years Mean age 5.8 years 10 girls, 8 boys	Actigraph (GT1M); Right hip; Filter not reported; 15s	Direct observation (modified CARS)	SB≤372 LPA 373-584 MVPA ≥585

2.7.5.7 *Duration of monitoring, wear time, and non wear-time*

For the measurement of PA of young children, a 7 day wearing protocol is recommended where possible to increase the chances of there being enough usable data (Addy et al., 2014; Penpraze et al., 2006), with a minimum of 3 days required for reasonable estimates (Bingham et al., 2016b; Hislop et al., 2014; Ricardo et al., 2020). For examining SB, research suggests that ≥ 4 days of monitoring is required for reasonable estimates (Byun, Beets and Pate, 2015). In addition to this, it is suggested that data should include at least one weekend day (Hinkley et al., 2012a). There are discrepancies in the recommended number of hours wear time of an accelerometer, ranging from 6 to 10 hours wear time per day required for the most reliable estimates of PA (Bingham et al., 2016b; Hislop et al., 2014; Lima et al., 2014; Penpraze et al. 2006). Despite this, there is an increase in 24 hour wear protocols, with the inclusion of sleep alongside PA, but also to increase compliance with accelerometer wearing protocols. Studies should ensure that they include the minimum amount of data, including both number of days and hours per day, to meet a reliability score of at least 0.7, and should report this in their study (Hinkley et al., 2012a).

Similarly, there is no consensus on defining non-wear time of accelerometry data (Cliff, Reilly and Okely, 2009; Janssen and Cliff, 2015; Migueles et al., 2017). Non-wear time is often determined by consecutive zero counts in the data sets, however, this varies between studies and becomes more challenging when also examining SB using accelerometry (Janssen and Cliff, 2015). A common and useful way in which non-wear time can be determined is by completion of a diary or log to state when the accelerometer was removed and reasons for this, which can then be cross-validated with numbers of consecutive zeros within the data sets.

2.7.6 Combined heart rate and accelerometry

Another method of PA measurement are devices that integrate different methods, including a combination of heart rate monitoring and accelerometry into one device. This allows the combination of detecting movement from the accelerometer and intensity of activity based on heart rate from this element of the device. The Actiheart (CamNtech, Cambridge, UK) is an example of such a device, that has been validated for use in pre-school children (Adolph et al., 2012). However, there are some feasibility concerns of this device with young children (aged 2-3 years old) due to the attachment directly on to the chest (Costa et al., 2013).

2.7.7 Pedometers

Pedometers are a wearable device that respond to oscillations in the acceleration of the body during gait cycles that provide output as 'step counts'. They often have a small screen that displays the number of steps recorded. Some devices are only able to store accumulated step count, meaning that participants may have to log steps recorded at certain time intervals (e.g. the end of every day). Pedometers are the cheapest, and a feasible, device for the measurement of PA. However, their inability to accurately detect movement behaviours beyond step count limit their use.

Pedometers can differ in their internal mechanisms (Crouter et al., 2003; Schneider, Crouter and Bassett, 2004):

- 1) **Spring-lever mechanisms**- spring suspended lever moves up and down in response to accelerations, which opens and closes an electrical circuit, the lever arm makes electrical contact which is then registered as a step.

Example: Yamax Digiwalker SW-200

- 2) **Magnetic reed proximity switch**- uses a spring-suspended horizontal lever, with a magnet attached to the lever, the magnetic field causes 2 overlapping pieces of metal to touch, which registers as a step.

Example: Omron walking style pedometers

- 3) **Horizontal beam and piezo-electric crystal**- piezoelectric systems with horizontal beam that has a small weight on the end, walking motion generates accelerations, the small weight compresses a piezoelectric crystal that generates voltage oscillations recorded as steps.

Example: New Lifestyles

Pedometers are used in PA research with pre-school children and are usually worn on the hip in this age group (Chow and Au, 2009; Louie and Chan, 2003; Oliver et al., 2007). There are various types of pedometers that have been tested and evaluated in pre-school children. These include: Yamax Digiwalker SW-200 (Hands and Larkin, 2006; Louie and Chan, 2003; Oliver et al., 2007) and Omron Walking Style pedometers (Fotini, Antonis and Dimitra, 2015). The validity, reliability and feasibility of pedometers used to examine PA and SB of pre-school aged children are discussed in detail in *Chapter 3*.

2.7.8 Consumer Wearables

In recent years, there has been an expansion in the interest in the use of consumer wearables (e.g. Fitbit) in the general population, and subsequently interest in the utility of these wearables for PA and SB research (Fuller et al., 2020). Despite this, commercial wearable devices are shown to be particularly poor at determining energy expenditure across age groups (Bunn et al., 2018; Evenson, Goto and Furberg, 2015; Feehan et al., 2018;

Fuller et al., 2020). The devices appear to be more accurate, yet still underestimate, heart rate and step count; however, this has been shown within laboratory based studies more so than for free living activity, a context in which these devices are most likely to be used (Fuller et al., 2020). Literature has raised some key issues regarding the use of commercial devices for research purposes. The systematic review by Fuller and colleagues examining the validity and reliability of consumer wearables reported that the majority of devices tested during the period of the review were outdated or discontinued by the time of publication of the review. The fast moving technology industry results in new devices continuously being updated and renewed, leaving little room for accurate validity and reliability testing in time for use for research. Additionally, there is variability in the way in which the data from the devices is sourced and extracted; some studies counted steps on the device in short time intervals instead of connecting the device to a platform after recording, whilst other studies exported and downloaded data from user accounts or from the brand application programme interface. Finally, algorithms used in consumer wearables are constantly changing based on sensor development and technological advances; thereby limiting the ability to determine true validity and reliability of such devices. Commercial devices may be used for estimates for PA rather than accurate quantities (Fuller et al., 2020).

To date, only three studies have been identified that examined the measurement properties of consumer wearables in pre-school aged children, including the Fitbit (Flex) and Fitbit (Zip) (Byun et al., 2018b; Byun, Kim and Brusseau, 2018; Sharp et al., 2017). As will be discussed in *Chapter 3*, these studies show promising validity for a range of outcome measures including SB, MVPA, total PA, and step counts; however, this is based on a limited amount of studies. The important considerations of using wearable technologies and continuous advancements of devices need to be addressed, as the use of such devices at

scale for research purposes and surveillance are limited due to the known lack of ability to standardise across devices and without knowing the algorithms that are being used behind the commercial based devices (see section 2.11.2 p.128 for more detailed information).

2.7.9 Proxy-reported measurement tools

Self-reported measurement tools involve participants reporting their own engagement in various PA and SB. Due to limited language, speech and cognitive abilities of pre-school aged children to recollect specific details on intensity, frequency and duration of their own activities, proxy reported methods are used instead. Proxy reports are where the responses are completed by someone else on behalf of the child, such as by their parent, carer, or teacher. Proxy report measures of PA and SB measurement of pre-school children can be broadly grouped into two main categories: questionnaires and diaries. Another method increasingly used is ecological momentary assessment, as such, I will discuss this briefly in this section. The information provided on self/proxy reported tools can be used to generate PA outcomes e.g. estimation of amount of activity someone engages in, whether they meet the guidelines. There can be some specific challenges with proxy reported tools, such as parents not being with their child for the whole day so many not know their every activity.

In previous years, there has been a false hierarchy of measurement tools used to assess PA and SB (Kelly, Fitzsimons and Baker, 2016). Proxy and self report tools have often been referred to as 'subjective' methods of PA measurement. However, it is becoming increasingly accepted that proxy reported tools provide a valuable insight into the behaviours of PA and SB and can capture information at scale that cannot be captured by other methods of measurement. For example, these tools can determine context of

behaviour – if a child is sedentary, a proxy tool can tell us what behaviour the child engaged in during that time, which cannot be detected from device based tools. There are a host of uses for self/proxy reported tools that posits them as a very valuable source of PA and SB measurement, including: monitoring and surveillance, large scale evaluation, screening and health promotion, assessing changes in targeted behaviours following interventions, providing information about domain, type and context of activities (Sattler et al., 2021).

2.7.9.1 *Questionnaires*

Questionnaires are a widely used method for PA and SB measurement with pre-school aged children, as they are convenient, cost effective and have the least amount of burden on participants and researchers alike. Questionnaires can differ in the time period that they cover (e.g. past week, past month, typical activity), length of questionnaire (e.g. number of questions), format (e.g. paper, electronic, interviewer led) and what they aim to measure (e.g. frequency of activity, context of activity). Questionnaires are often completed retrospectively. Questionnaires can be used for surveillance purposes, in research studies examining PA and/or SB, or can be used in studies whereby the outcome of PA/SB is a secondary outcome. An example of a questionnaire developed and used for assessing PA and SB of pre-school aged children is the 'Pre-school age physical activity questionnaire (PRE-PAQ)' (Dwyer et al., 2011).

2.7.9.2 *Diaries*

Diaries are another form of self/proxy report to assess PA and SB, that are slightly more intensive than questionnaires but can extract more detailed information. Similar to questionnaires, diaries can differ in time period, length, format and what they aim to measure. The advantage of diaries over questionnaires is that they can be used more 'in the

moment' reporting which eliminates the need for retrospective reporting and relies less on memory.

2.7.9.3 *Ecological Momentary Assessment (EMA)*

EMA is a 'real time' method of data collection for PA and SB. EMA is a method of self/proxy report that involves prompting a participant to complete a questionnaire/few short questions at intervals throughout the day. The increase in technology, and particular increase in the use of mobile phones has allowed advances in data collection techniques such as EMA (Kuntsche and Labhart, 2013). Information obtained can include what the participant is currently doing, any activity that is taking place (or if an individual is inactive), who they are with, where the activity is taking place, and thoughts/feelings. EMA is most often conducted through time sampling (prompts throughout the day), but sometimes involves event-based sampling (where participants record information when they are engaging in the behaviour of interest) (Degroote et al., 2020). It is suggested that EMA is a more ecologically valid way of examining PA and SB as it captures the activity at the exact time and place (Dunton, 2017). EMA has been used to examine the PA and SB (and context in which this happens, where, and with whom) in children aged 8 years and older (Dunton, et al., 2012; Liao et al., 2014; Romanzini et al., 2019), and is an acceptable and valid method of measurement in older children (Dunton et al., 2011). At present, there is limited use of this method within the pre-school population. One study assessing the validity of a TV diary in pre-school aged children suggested that EMA showed moderate convergence with the TV diary ($r=0.47-0.51$) (Mendoza et al., 2013). Much more research is needed to assess the EMA method for the collection of PA and SB data in the pre-school population.

2.7.10 Strengths and limitations of tools

Table 2.5. Strengths and limitations of measurement tools used to assess physical activity and sedentary behaviour

Measurement Tool	Strengths	Limitations
Whole room calorimetry	• Can accurately measure energy expenditure • Accurate method for validation studies when aim of comparison tool is to determine energy expenditure	• Expensive • High researcher and participant burden • Can often only be used with small sample sizes • Cannot assess habitual or free living PA • Trained researchers, specialist knowledge, equipment and facilities are needed to conduct and analyse
Indirect calorimetry	• Can measure free-living energy expenditure as it's portable	• Expensive • Can often only be used for small sample sizes • High participant burden in that it can be bulky for participants to wear • Does not provide any indication of frequency, duration, type and domain of PA • Specialist knowledge, equipment and facilities are needed to conduct and analyse

Doubly labelled water	• Can measure free-living energy expenditure • Relatively easy to administer	• Expensive • Can often only be used for small sample sizes • Time consuming for researchers and participants (lots of samples to collect) • Dose and analyses can vary depending on body weight of participant • Does not provide any indication of frequency, duration, type and domain of PA • Specialist knowledge, equipment and facilities are needed to conduct and analyse
Direct observation	• Can provide information on context of activity, including: duration, intensity, frequency, type and domain.	• Requires trained observers • Time intensive • High researcher burden • Likely to cause reactivity bias • Can only be conducted for limited periods of time, so does not give an indication of habitual activity

Wearable cameras	• Can provide detailed information on context and setting in which PA takes place	• Expensive • High participant burden to wear a camera all the time - low acceptability and feasibility with older children and adult samples, testing of feasibility with pre-school children warranted • Ethical issues –not only relating to the participant wearing the camera themselves, but with other people captured in the recordings (although ethical guidelines have been proposed (Kelly et al., 2013)). • High researcher burden with time consuming data handling and analysis • Images can be of varying quality • Can be linked to accelerometer data, but without this does not provide information on the intensity of activity
Heart rate monitor	• Can be used in free-living situations • Can collect data over a reasonable period of time (usually 1 week).	• Not directly measuring PA only, as heart rate is affected by multiple other factors, including psychological stimuli (anxiety, stress), and environmental factors (temperature).

		• Does not provide any contextual information about activity • Risk of reactivity bias • Wear period can increase participant burden • Placement on chest can cause discomfort • Time consuming data handling and analysis
Accelerometers	• Can measure free living activity (time stamped data) • Can provide information on intensity, frequency, duration and timing of activity	• Can be quite burdensome for participants (long wear time) • Does not provide information on type or context of activity • Risk of reactivity bias • Certain activities can exacerbate activity counts (e.g. cycling, children playing on swing) • Often removed during water based activities which may impact compliance (forgetting to put the device back on) • Interpretation and protocol decision (e.g. placement, cut points) can impact on what activities are recorded and outcomes • Time consuming data handling and analysis

		• Large amount of data storage required
Combined heart rate and accelerometry	• Can measure free living activity (time stamped data) • Can provide information on intensity, frequency, duration and timing of activity	• Relatively expensive • Measurement 'noise' from heart rate data • Time consuming data handling and analysis • Large amount of data storage required • Feasibility concerns for use with young children (Costa et al., 2013).
Pedometer	• Inexpensive • Suitable for medium-large scale studies • Low participant and researcher burden • Can provide a good measure of steps taken (dependent on validity/reliability of model)	• Open display (where individual can see how many steps they have done) can cause reactivity bias • If diary is used to report daily step count recall and social desirability bias may be present (would also require proxy report for young children) • Individuals may tamper with device • Can incorrectly measure movement (e.g. from car/bus travel) • Not possible to assess intensity, frequency, duration, type or domain of PA. • Participant sometimes has to log daily steps (for pre-school children, parents would need to record this)

Consumer wearables	• Comfort and attractive design • Low participant burden • Often linked to smartphone apps for easy tracking	• Lack accuracy for measurement of PA for research purposes • Changing algorithms and lack of access to way in which data is internally processed means they are not entirely accurate • Knowledge of validity/reliability of these devices is largely unknown due to differing algorithms used • New models continuously being developed – difficult for research purposes
Questionnaires	• Suitable for large scale studies • Low cost • Low participant burden • Can provide contextual information, including: duration, intensity, frequency, type and domain of PA/SB.	• More prone to subject bias (recall bias, social desirability bias) and error (reporting on child's behalf so may accidentally miss some behaviours) • Requires proxy report for young children • Increased likelihood of misinterpretation of PA due to language and cultural factors • Cognitively demanding – requires participant literacy when using self-administered method.
Diaries	• Suitable for large scale studies • Relatively inexpensive	• High participant burden -time consuming/cognitively demanding - requires participant literacy when using self-administered method.

	• Can provide contextual information, including: duration, intensity, frequency, type and domain of PA/SB.	• More prone to subject bias (recall bias, social desirability bias) and error (reporting on child's behalf so may accidentally miss some behaviours) • Requires proxy report for young children
Ecological momentary assessment	• Suitable for large scale studies • Low cost • In the moment activity recording, can provide indication of free living activity • Potential to reduce recall bias in comparison with other self/proxy reported tools • Can provide contextually rich information, including both quantitative and qualitative about the context of PA/SB • Can provide information further than that of solely the behaviour including the opportunity to report on psychological and social aspects	• Only provides a snapshot of activity when the participant responds – does not give us information on frequency, intensity or duration of activity taking place • Can be quite burdensome – if rate of prompting is too high may lead to drop out, but not enough prompting may not get an accurate account of PA/SB • Potential for social desirability bias • High likelihood of missing data for several reasons - participants not carrying smartphone, unable to respond to EMA prompts if busy doing something else, technological issues such as phone battery • Additional challenges when using this method with young children due to the need for proxy report including whether the child is at school or home, and who will report on the child's behalf.

2.8 Brief overview of sleep measurement tools

In this section, I provide an overview and discussion of the measurement tools that can be used to assess sleep. Similar to both PA and SB, the ways in which sleep is measured vary in line with different research aims and questions, the parameters and dimensions of sleep being measured (quantity, quality, architecture, schedule), the age of the child, the availability of the various resources (time, financial costs) and equipment, and the setting (laboratory based versus natural environment). The various methods used to assess sleep in the pre-school population will be discussed in turn, including how each tool is used and the primary purpose. I will describe each of these tools generally, drawing specifically on evidence of the application of the tools for use with pre-school aged children where available. The tools I will discuss include:

- Polysomnography
- Electroencephalography
- Videosomnography
- Direct observation
- Accelerometers
- Consumer wearables
- Proxy reported tools (questionnaires and diaries)

Following discussion of each of these measurement tools, I present a summary table of the strengths and limitations of each of the tool (see Table 2.6). A detailed review of the quality of the validity, reliability, feasibility of the measurement tools used to assess sleep of pre-school aged children can be found in *Chapter 4* of this thesis.

2.8.1 Polysomnography

Polysomnography (PSG) is a method of measuring sleep that involves electrodes and sensors being attached to the individual prior to their night time sleep, for one to two nights, in a laboratory setting (Sadeh, 2015). PSG includes the monitoring of electrical brain activity (electroencephalography – EEG), eye movements (electrooculography), electrical activity produced by the skeletal muscles (electromyography), respiration activity (e.g. breathing flow, oxygen saturation) and cardiac activity (electrocardiogram), alongside video recording of the sleep (Iber et al., 2007; Mouthon and Huber, 2015). This method is considered the ‘gold standard’ of sleep assessment, due to the detailed information and data that this method can detect. This method is more likely to be used in studies exploring the underlying processes and mechanisms of sleep, as well as for clinical purposes to diagnose sleep disorders, rather than to detect everyday behaviour patterns, as such an extensive method is not usually necessary (or feasible) for habitual assessment of sleep (Mouthon and Huber, 2015). This method has been used as the criterion method of validation studies of accelerometers in pre-school aged children (e.g. Meltzer et al., 2015).

2.8.2 Electroencephalography

Electroencephalography uses electrodes to assess brain activity, eye movements and submental (space between face and neck) activity, whilst individuals are asleep. This can be useful for identifying when individuals are in the different stages of sleep (non-rapid eye movement sleep versus rapid eye movement sleep). This method can also identify sleep waves to determine when an individual is in deep sleep. Although electroencephalography is used during PSG, electroencephalography can sometimes be used as a lone, less intensive (due to less electrodes and sensors attached to the body) method (Mouthon and Huber,

2015). Similarly high density electroencephalography can be used which involves the use of more electrodes than traditional electroencephalography so can detect whole brain activity (Mouthon and Huber, 2015). Electroencephalography has been used in studies with pre-school aged children (e.g. Hatzinger et al., 2013), but is not used for large scale public health purposes due to the cost.

2.8.3 Videosomnography

Videosomnography is a method of measuring sleep based on video recordings. This often involves several cameras being set up in a child's room, or in a sleep laboratory setting, so that patterns of sleep, quantity, behaviour during sleep and night wakings can be detected. This method enables a researcher or health care professional to view the whole night-time sleep of an individual in their natural sleeping environment, which can be conducted over several nights. This method can be used when researchers are interested in measuring certain sleep behaviours, including bedtime interaction in infants and young children or night time disturbances such as night terrors or sleepwalking, and can be used as a screening tool for some clinical sleep disturbances such as sleep apnoea (Sadeh, 2015).

2.8.4 Direct Observation

Direct observation involves in the moment assessment of sleep by trained observers. This involves the researcher observing and scoring the sleep of an individual including states of sleep and wakefulness over a specified period of time, which can be conducted in the individuals natural sleeping environment. This method is almost only ever conducted with infants or very young children whose sleep is less likely to be impacted by the presence of other people (Sadeh, 2015).

2.8.5 Accelerometers

As discussed in section 2.7.5 (p.76), accelerometers can also be used to measure movement behaviour of pre-school aged children. Although the accelerometers work in the same way as described for PA and SB measurement, the main difference is that different makes of accelerometer are commonly used for sleep than those used to assess PA and SB, which I will discuss in more detail. Accelerometers are small devices, usually worn on the wrist for measuring sleep of young children (Meltzer et al., 2012b), that monitor acceleration to detect body movements and intensity of movement (Mouthon and Huber, 2015). For the assessment of sleep, accelerometers work on the basis that when movement is detected this reflects a state of wakefulness, whilst inactivity is reflective of sleep (Galland et al., 2014). The devices can be worn 24 hours a day, and for extended periods of time, often around 7 days. Accelerometers can be used to assess both sleep duration and some elements of sleep quality (e.g. sleep latency), however; it is suggested that accelerometers are particularly poor at detecting night time awakenings in young children (Bélanger et al., 2013; Meltzer et al., 2012a; Sitnick, Goodlin-Jones and Anders, 2008). Accelerometers are primarily recommended for documenting the sleep of an individual in their natural environment for prolonged periods of time and to determine 24 hour sleep time that does not rely solely on self or parental report.

2.8.5.1 Types of Accelerometer

There are many different brands of accelerometer that are used to assess sleep of pre-school children including the Micromini Motionlogger (Ambulatory-Monitoring Inc.) and Actiwatch 2 (Philips Respironics) (Schoch, Kurth and Werner, 2021). Devices most frequently validated to assess PA and SB of pre-school children are also starting to be used to monitor

sleep for 24 hour monitoring, these include the Actical (e.g. Galland et al., 2016; Taylor et al., 2021) and Actigraph GT3X+ (e.g. (Duraccio et al., 2018; Tracy et al., 2021) accelerometers.

2.8.5.2 Placement of Accelerometers

As written in section 2.7.5.2 (p.78), placement of accelerometers often depends on the device. For the measurement of sleep, accelerometers are usually placed on the wrist of young children (Meltzer et al., 2012b; Schoch, Kurth and Werner, 2021). Studies have also used an ankle or waist/hip placement to assess sleep of young children (Duraccio et al., 2018; Sekine et al., 2002; Tracy et al., 2021), and devices are usually placed on the ankle or calf in infants (Chaput et al., 2017b). Studies in older children have shown that hip placement provides comparable results with wrist-worn accelerometry data for detecting sleep duration (Kinder et al., 2012; Paavonen et al., 2002). Others have suggested that whilst hip placement could be used as a ‘proxy’ measure of sleep in children (mean age 10 years), hip and wrist data should not be used interchangeably (Hjorth et al., 2012). In adults, wrist placement has shown to have higher accuracy for detecting sleep when compared with accelerometry than that of hip placement (Slater et al., 2015; Zinkhan et al., 2014). The accuracy of different placements for the measurement of sleep of pre-school aged children is an area for future research (Migueles et al., 2017).

2.8.5.3 Epochs

As outlined in section 2.7.5.5 (p.81), accelerometers collect data based on movement and summarise overall intensity across pre-defined epochs (a period of time in which data is collected and summed) that provides the output of activity counts. The epoch length used in studies using accelerometry to assess sleep of young children is often not

reported (Schoch, Kurth and Werner, 2021). When reported, one minute epochs are most frequently used, with 15-30 second epochs also used in some instances (Meltzer et al., 2012b; Schoch, Kurth and Werner, 2021). Once the data from the accelerometer has been downloaded, the activity counts are processed to determine different intensities of activity (e.g. high versus low activity counts) that are used to determine wake and sleep status.

2.8.5.4 Algorithms

Scoring algorithms for sleep measurement are the equivalent to the term cut points for PA and SB measurement, in that they are indicators used to determine sleep and wake status. A recent review revealed that nearly half of studies (47.5% studies) using accelerometry to determine sleep of infants and young children did not report the specific algorithms used, sometimes reporting the software used but this did not include the algorithms of the software (Schoch, Kurth and Werner, 2021). Reporting the software only is problematic as different algorithms may be applied between different software, different versions of a similar software, and updates within the software may mean that results are not comparable and the validity of the algorithms applied are unknown (Schoch, Kurth and Werner, 2021). Until recently, there were no available algorithms that had been developed to determine sleep in the pre-school population (Migueles et al., 2017). The most commonly used algorithms to determine sleep of pre-school aged population samples included those developed using adolescent and adult samples (Cole et al., 1992; Sadeh, Sharkey and Carskadon, 1994). Research in older children (aged 9-10 years) reported that the most commonly used algorithm, the Sadeh algorithm (Sadeh, Sharkey and Carskadon, 1994), significantly overestimated sleep period by nearly 45 minutes; emphasising that the algorithms developed with adolescents and adult population samples are not transferable

to the sleep of children (Tudor-Locke et al., 2014). Belanger and colleagues assessed three algorithms using the Actiwatch-L to detect night time sleep and wake periods in children aged 2-5 years. Results suggested that the '*AlgoSmooth*' algorithm developed previously but not validated was the best of the three algorithms for detecting sleep of the pre-school children (Sitnick et al., 2008). This algorithm showed lower concordance with PSG for detecting wakefulness and was poor at determining night awakenings, but was still the most accurate of the three (Bélanger et al., 2013). In pre-school specific populations, a recently developed decision tree algorithm (R software package- '*PhysActBedRest*') for use with the Actigraph GT3X+ was able to detect bedrest, daytime sleep, and wake time similarly to visual identification of the data and was better at identifying bedrest than the most frequently used algorithms developed in other age groups usually applied to young children (Tracy et al., 2021). Algorithm development for the accurate detection of sleep and wake periods, including day time sleep, specifically for pre-school aged children is in its infancy. Further development of algorithms to detect sleep, and to distinguish sleep from other low movement behaviours including SB (and non-wear time) are needed. This is particularly important if using just one device to detect 24 hour movement.

2.8.5.5 Duration of monitoring, wear time, non-wear time

Research examining the wear time of devices for reliable estimates of sleep has shown that at least five nights of accelerometer data are needed for children and adolescents (Acebo et al., 1999). For three-year old children, six nights was needed for a reliable estimate of sleep duration. However, this research demonstrated that there was low reliability for even seven days of monitoring of children aged three-years old for a range of sleep parameters including sleep minutes, wake minutes and sleep efficiency, with

individual children varying widely from night to night for these parameters (Acebo et al., 1999). Suggested reasons are that this may be a representation of developmental instability at this age. Further research with children aged 3 to 7 years suggested that four to seven days of accelerometry data was needed to obtain reliable estimates of sleep duration and quality. However, at least eleven days were needed to reliably estimate sleep latency. Authors suggested that sleep is incredibly variable for young children even within a one-week cycle and shows limited stability across these early years (Taylor et al., 2015).

Wear time required may vary depending on which elements of sleep are being examined. For example, if only assessing night-time sleep then children may be asked to wear the accelerometer when going to bed and remove when they wake up (Meltzer et al., 2015). Most frequently participants are asked to wear the accelerometer continuously for a period of time, usually seven days (e.g. Lam et al., 2011b; Vaughn et al., 2015). It is recommended that a 24 hour wear protocol should be used for both the measurement and analysis of accelerometer data for the assessment of sleep, particularly in infants and young children where naps and daytime sleep are still prevalent (Schoch, Kurth and Werner, 2021).

Being able to identify sleep episodes from SB and non-wear time when using accelerometers is difficult due to accelerometry providing data based on detecting movement, and all three of these conditions being of similarly low/no movement (Quante et al., 2015). The standard method for sleep detection when using accelerometry is through use of a concurrent sleep diary/log, which is often the case when assessing sleep in young children (Schoch, Kurth and Werner, 2021). Using a diary has been shown to significantly improve the accuracy of the output data and provides useful relevant information for the scoring of sleep, compared to an automated approach that did not include an accompanying

diary being completed (Tétreault et al., 2018). Diaries/logs help ensure more accurate scoring of the data for the following reasons:

- 1) Score sleep onset/offset (through reporting bed/wake time)
- 2) Identify important artifacts in the data such as non-wear time (through reporting time/reason for removal)
- 3) May include reporting of prolonged periods of quiet activity (e.g. watching TV)
- 4) Reporting of activity that should be scored as sleep but may show movement (e.g. if child naps during a car journey).

(Meltzer et al., 2012b; Sadeh and Acebo, 2002; Quante et al., 2015)

2.8.6 Consumer wearables

As discussed in section 2.7.8 (p.87), in recent years there has been an increase in the development and availability of consumer wearables used to measure health behaviours, including sleep. These devices are installed with the technical hardwiring of accelerometers, thereby being able to detect movement and motion. Fitbits are a commonly used wrist worn commercially available devices that were initially designed to track steps, and movement. However, with individuals increasingly monitoring their health behaviours, Fitbit models have become advanced, and now have 'sleep-tracking' capabilities. A systematic review of the literature revealed that earlier generation 'non-sleep staging' devices were poor at detecting sleep and wakefulness, whilst newer generation 'sleep-staging' Fitbit models have improved ability to differentiate sleep from wake (Haghighat et al., 2019). However, this conclusion was based on the results of three studies only, containing only adult participants. Others have urged caution with the use of commercial wearables to be used confidently in sleep research, due to very limited studies on the validity of the devices,

and limitations such as the inability to detect the data processing decisions and the continuously changing algorithms (Evenson, Goto and Furberg, 2015).

At present, only one study has examined the utility of the Fitbit (Ultra) in a pre-school population (discussed in *Chapter 4*), which has shown that the device significantly underestimated sleep time and sleep efficiency, and overestimated wake after sleep onset in comparison with PSG when using the sensitive mode. When using the 'normal mode' the Fitbit (Ultra) overestimated sleep time and sleep efficiency, but there was no significant difference for wake after sleep onset (Meltzer et al., 2015).

Research regarding the ability of wearable devices to accurately measure health behaviours including PA, SB and sleep, is in its infancy and focuses more on adult population samples. Further research is needed before such methods can be used confidently in research with young children, including intra-device reliability, confidence in data processing decisions and algorithms used within the devices.

2.8.6.1.1.1 Smartphones

Smartphones are gaining increasing popularity as contactless hardware devices that can be used to measure sleep (Ibáñez, Silva and Cauli, 2018). Smartphones contain multiple hardware devices: a microphone, camera, and accelerometer, making them able to measure and monitor sleep. These features within smartphones, alongside the increase in individuals' desire to survey their own health behaviours (Van den Bulck, 2015), has led to the development of multiple smartphone applications that can be used for the purpose of monitoring sleep (e.g Sleep Cycle Alarm Clock; Sleep Better; Smart Sleep Manager). Research in adult population samples has suggested that sleep applications are poor at detecting sleep parameters in comparison with PSG (Fino et al., 2020) and that most sleep

apps have not been validated in scientific research (Ong and Gillespie, 2016). There are very few studies examining smartphone application recorded data for the assessment of sleep in children. However, Patel and colleagues (2017) reported that a phone app assessing sleep (Sleep Cycle App) was not able to determine sleep duration, sleep latency, and did not have the ability to determine different stages of sleep, when compared with PSG in a sample of children (Patel, Kim and Brooks, 2017). It was concluded that phone applications monitoring sleep are not yet accurate enough to be used for research and scientific purposes (Patel, Kim and Brooks, 2017). Additionally, it is unlikely that pre-school children will have their own smartphone so this would not be an appropriate or applicable measure for the age group of interest in this thesis.

2.8.7 Proxy reported measurement tools

Similar to the behaviours of PA and SB, proxy reported measurement tools can be used to determine sleep of pre-school aged children. Proxy reported tools are the most commonly used tool to examine sleep of young children. Here I will briefly discuss questionnaires and diaries. *Chapter 4* presents a systematically conducted rapid review of the validity, reliability and feasibility of tools used to examine sleep of pre-school aged children, where I will discuss the evidence in relation to pre-school aged children in depth.

2.8.7.1 Questionnaires

Sleep based questionnaires are a time and cost efficient way to achieve data based on sleep patterns, problems and sleep behaviours (Mouthon and Huber, 2015). Questionnaires can be used when exploring patterns, schedules and behaviours of sleep. Similar to questionnaires used in PA and SB research, sleep questionnaires are often completed using retrospective reporting, can cover different time periods, be of varying

length, and differ in what they aim to measure (e.g. sleep duration, behavioural sleep problems). Questionnaires are advantageous over other methods as they are much less costly, intrusive and intensive, and can be used on a much larger scale than any other measurement tool used to assess sleep. Additionally, they do not require additional technical equipment. There are many different questionnaires that have been developed and are used to assess the sleep behaviours of young children (e.g. Children's Sleep Habits Questionnaire); however, most questionnaires have been devised with a focus on problematic sleeping behaviours (e.g. Pediatric Sleep Questionnaire, Epworth Sleepiness Scale). The measurement properties of sleep questionnaires used in children are often unknown (Spruyt and Gozal, 2011), which I will discuss more in depth in *Chapter 4*.

2.8.7.2 Diaries

Sleep diaries are a common method used to explore sleep patterns of individuals of all ages, with proxy-reported diaries being completed by parents and carers on behalf of young children. The diaries can provide rich information on the sleep schedule, patterns, behaviours, quantity and to some extent quality of the sleep. Sleep diaries have been suggested as a suitable measure to assess sleep duration, schedule or sleep related behaviours and interactions, that are readily available and usually do not require any specialist equipment or knowledge (Ibáñez, Silva and Cauli, 2018). Sleep diaries provide a detailed assessment of sleep patterns and daily contextual information on the child's sleep (Quante et al., 2015), and can help to determine important sleep parameters such as sleep onset and offset (Tétreault et al., 2018). Sleep diaries can determine duration but also sleep quality in terms of sleep latency, number of night awakenings, wake after sleep onset and sleep efficiency (Ohayon et al., 2017). Diaries usually consist of daily reporting on

parameters such as sleep start and end time, sleep period, night wakings, and sometimes day time sleep (Mouthon and Huber, 2015). However, other variations of sleep diary include intervals of time (e.g. 15 minutes) where participants self or proxy report sleep time (e.g. Werner et al., 2008).

2.8.8 Strengths and limitations of tools

Table 2.6. Strengths and limitations of measurement tools used to assess sleep

Measurement Tool	Strengths	Limitations
Polysomnography	• Very detailed view of all aspects of an individuals' sleep, including sleep stages based on brain activity • Ability to accurately detect and diagnose clinical sleep disturbances • Controlled environment – full control over the environment in which testing occurs and close observation of the tested participant	• Expensive • Time intensive • Burdensome for both researcher and participant • Can only be conducted on small scale (due to intensive nature and expense) • Only captures one or two nights of sleep (usually only conducted for short period of time due to expense) • Specialist equipment, knowledge and training required to both conduct research and analyse data • Not representative of usual sleep; does not take place in natural environment (laboratory based), as such the derived measures may not be representative of the individuals' usual sleep.

		• Physically intrusive- involves multiple sensors and electrodes being attached to the body which may be quite challenging with young children. • Cannot capture all natural sleep episodes, as PSG does not capture daytime sleep that young children may engage in.
Electroencephalography	• Can provide a detailed view on brain activity during an individuals' sleep • Slightly less intensive than polysomnography, can sometimes be conducted at home rather than in the laboratory due to simplified equipment (however, high density electroencephalography is laboratory based)	• Expensive • Time intensive • Burdensome for both researcher and participant • Physically intrusive • Can only be conducted with small sample sizes • Specialist equipment, knowledge and training required to both conduct research and analyse data • Often takes place in laboratory setting, so may not be reflective of usual sleep

		• Challenging with young children to attach the electrode cap
Videosomnography	• Can be used to assess sleep in natural sleep environment • Not physically intrusive • Can identify behaviour during sleep (e.g sleep walking; getting out of bed; parental intervention) • Can be used as an initial screening tool for clinical sleep problems e.g. sleep apnoea	• Specialist equipment, knowledge and training required to both conduct research, visually inspect and analyse data – makes this a time intensive method for researcher • Home installation of the equipment - safety concerns with installing equipment in child's bedroom and potentially invasive on privacy • Potential for data interferences (dependent on where the camera is positioned, child's movements or positions in bed may not always be captured) • Time intensive for researcher • Can only be conducted on small scale
Direct Observation	• Can be used to assess sleep in natural sleeping environment	• Only used with infants or very young children (unlikely with pre-schoolers), as sleep very likely to be impacted by a person being present. • Very invasive method

		• Specialist knowledge and training required to both conduct research and analyse data making this a very time intensive method for researchers • Can only be conducted with small sample sizes • Often only conducted for a couple of hours, due to this being wholly dependent on researcher directly observing the participant • Does not provide information on habitual sleep or daytime sleep – provides a very small snapshot
Accelerometers	• Relatively inexpensive (in comparison with above measures) • Enables 24 hour monitoring (able to identify both day and night sleep) • Can be used to assess sleep in natural setting, over multiple days • Not intrusive or invasive	• Is only based on body movement, so may not accurately distinguish when someone is asleep/awake. • Relies on different data interpretation choices (placement, algorithms) that can influence accuracy of the method • Often needs a diary/log used alongside the device to determine sleep artifacts such as

		sleep onset and offset – which may increase burden on participant • Requires trained researchers to analyse data
Consumer wearables and smartphone applications	• Potential for large scale sleep measurement	• Data processing decisions are unknown • Algorithms can be changed at any time making these not feasible for research • At present, there is very limited research on the validity and reliability of these tools, particularly in young children
Questionnaires	• Cost efficient • Can be used on larger scale • Can be used to assess sleep in natural setting/habitual sleep • Can be used to assess sleep over extended periods of time • Not intrusive or invasive	• Higher risk of misreporting or inaccuracies in reporting, may include: recall or social desirability bias (heightened with proxy reporting, as this relies on the extent to which parents/carers are aware on their child's sleep behaviours).
Diaries	• Relatively cost efficient • Can be used on larger scale • Can be used to assess sleep in natural setting, over multiple days	• Can be quite burdensome on participants • Higher risk of misreporting or inaccuracies in reporting, may include: recall or social desirability bias (heightened with proxy

	• Can report on 24 hour sleep patterns to identify day and night time sleep • Not intrusive or invasive • Less likely to incur recall bias over questionnaire based methods	reporting, as this relies on the extent to which parents/carers are aware on their child's sleep behaviours).

2.9 Measurement of all movement behaviours simultaneously in pre-school aged children

As discussed in the sections 2.7 and 2.8, multiple tools and methods can be used to measure the isolated movement behaviours. Measuring the behaviours independently in young children comes with key challenges, however attempting to measure the whole 24-hour day in a practical, yet accurate way comes with additional challenges. The measurement of the movement behaviours of children in their early years has been stated to be the main potential reason as to why there is a lack of evidence examining the whole spectrum of movement behaviours (Kuzik et al., 2017). The 'gold-standard' measures that exist to assess the behaviours are laboratory-based methods; with PSG used to assess sleep, and calorimetry used to determine energy expenditure for PA/SB research (Quante et al., 2015). Both PSG and calorimetry are costly, which means they are not suitable for measurement at scale. Additionally, these methods are conducted in laboratories, environments that are unfamiliar to children and may not necessarily be reflective, and not applicable for measuring free-living activity. Based on these limitations, these methods are not useful for assessing habitual movement, but can be used as criterion methods when determining the validity of tools that will subsequently be used in free-living contexts and for public health purposes.

At present, studies examining the full spectrum of movement in pre-school aged children generally use a combination of accelerometers and surveys, through the use of accelerometers to assess PA, and surveys to examine sleep and sedentary screen time (Chaput et al., 2017a; Cliff et al., 2017; De Craemer et al., 2018; de Lucena Martins et al., 2021; Hinkley et al., 2020; Kim et al., 2020; Lee et al., 2021; Leppänen et al., 2019; McNeill et al., 2020; Vale and Mota, 2020). Sedentary screen time is most frequently assessed via

questions asking for average screen time, with no further assessment for SB or varying types of SB. In many instances, just one question is proposed to determine sleep, such as *'how many hours of sleep does your child usually have during the night?'*. To my knowledge, there is only one study, the SUNRISE study, that is currently collecting data on children's movement behaviours in countries across the globe (currently only in Scotland in the UK) (Okely et al., 2021b). The protocol includes the use of accelerometry for PA, and a parent reported survey for questions on sedentary time and screen time (with questions extending beyond sedentary screen time only) and sleep (Okely et al., 2021b).

For the majority of research projects and for measurement in public health research, one measurement tool capable of measuring the whole spectrum of movement behaviours in free-living conditions is warranted to reduce participant burden (e.g. not requiring the use of multiple tools) and researcher cost (e.g. not needing to purchase multiple tools). Of the currently available tools, direct observation, accelerometers, and proxy reported tools are able to measure PA, SB and sleep of pre-school children. I will briefly discuss each of these methods in turn.

Direct observation may be classified as the highest quality reference tool that can be used when aiming to determine the validity of tools measuring movement behaviours in free-living contexts (Sadeh, 2015; Sirard and Pate, 2001). However, direct observation is a highly intensive method that usually takes place for a limited period of time in a specific setting, meaning that it is not a feasible or applicable tool to determine free-living or habitual behaviours, and would not be a feasible method to detect the whole 24 hours of the day.

Accelerometers also have the ability to determine the whole spectrum of movement behaviours, frequently used in research with samples of pre-school children when examining PA and SB (e.g. Barber et al., 2015) and sleep (Meltzer et al., 2012b). Different accelerometers are used in different ways to measure the behaviours, creating some specific challenges associated with the use of accelerometry for movement behaviour measurement. The interpretation choices made when using accelerometers, including sampling frequency, epoch length, placement, wear time, what constitutes non-wear time, number of days of monitoring, or dealing with missing data can all impact on the outcomes of studies. The different algorithms and cut points used to interpret accelerometer data into meaningful outcomes are generally calibrated and validated for either PA and SB, or sleep; placement varies with hip placement usually used for assessing PA, and wrist placement for sleep (Quante et al., 2015). Additionally, when using 24 hour accelerometry, the different methods of detecting sleep (e.g., parental diary, different time filter algorithms, non-wear time of 60 mins of consecutive zeros) have been shown to provide significantly different outcomes (Meredith-Jones et al., 2016). It is still unknown, due to a lack of research, how well a single accelerometer in one location can accurately detect the full spectrum of 24-hour movement (Quante et al., 2015). This is an area for future exploration to develop appropriate protocols (e.g. placement, epoch, algorithms) for detecting all movement behaviours (PA, SB, sleep), that can accurately identify and separate SB, sleep and non-wear time, ideally with a single device. It is thought that to be able to accurately distinguish between the different states (non-wear, sleep, SB), data from additional tools may be needed such as an inclinometer, heart rate monitor or diary, but these features have not yet been examined or validated for accuracy (Quante et al., 2015). I will expand further on issues that may limit use of accelerometers for public health research in section 2.11.2

(p.126), including logistical issues with administration and processing of such large amounts of accelerometer data.

The final group of tools able to determine the whole spectrum of movement for pre-school aged children are proxy-reported tools (e.g. questionnaires and diaries), with tools available to measure PA and SB (e.g. Dwyer et al., 2011), and sleep (e.g. Duraccio et al., 2018). As with any tool, there are limitations of proxy reported methods with the presence of certain biases, including recall and social desirability bias (discussed below in section 2.10, p.122). However, these tools offer a great advantage in many ways, including being less burdensome for participating families and are inexpensive. Additionally, these tools have the ability to capture detailed information when measuring the full spectrum of movement that cannot be distinguished through accelerometer data alone, but would be important in order to determine compliance with guidelines. These include location of PA (e.g. whether children are engaging in outdoor play), type of SB (e.g. whether children are sitting and reading/playing with toys, or sat watching television), and can more easily distinguish between seated time and sleep time.

2.10 Issues and error in movement behaviour measurement

In this section I will describe some of the main biases present when measuring movement behaviours in any population, including in pre-school aged children. All methods of measurement of movement behaviour are affected by error and bias, but this includes different biases and to different degrees. Although it was envisaged that there would be a focus on proxy reported tools in this thesis, due to this being a more feasible method for large scale data collection and surveillance (Bauman et al., 2006); it is still important to understand the range of methods that can be used to assess movement behaviour of pre-schoolers. This is particularly the case for tools that may be used as criterion/reference methods for evaluations of proxy-reported measurement tools. Errors associated with each of the methods used to assess PA and SB can be found in section 2.7.10 (p. 92) and sleep in 2.8.8 (p.112).

Firstly, as with any research, self-selection or volunteer bias may be present, by only people from a certain group or with certain characteristics choosing to take part in the research. It may be the case that more active families or parents with higher education may respond to a study on movement behaviour; which would impact on the internal validity of the study and conclusions. For example, there is evidence demonstrating lower proportions of pre-school children (and their families) from lower SES groups participating in research examining movement behaviour than those from high SES groups (62% from high SES group versus only 7% from low SES group) (Hinkley et al., 2020). Additionally, loss to follow up may occur where there is non-random drop out of a group of individuals from a study, which can similarly confound findings. For example, research assessing adherence of pre-school children in relation to 24-hour movement behaviour guidelines showed that those who did

not provide complete data had a lower income and were predominantly African American compared with the overall included sample (Kracht, Webster and Staiano, 2019), and were from lower SES areas (Cliff et al., 2017). Examples of how low SES and minority groups can be under-represented in movement behaviour research.

Secondly, the process of being observed can alter behaviour, known as the 'Hawthorne effect' (Wickström & Bendix, 2000). When measuring movement behaviour, individuals are generally aware that their behaviour or the behaviour of their pre-school child is being monitored which may lead to alterations in behaviour. In physical activity literature, responding to the presence of being monitored is usually referred to as 'reactivity bias' or the 'reactivity effect' (e.g. Dössegger et al., 2014). Reactivity bias may occur if participants obtain direct feedback (e.g. are able to see step count on a pedometer), as a result of increased motivation to be active in line with social desirability bias (e.g. parents may take their children to play at the park more/reduce screen time if they know these behaviours are being monitored), or for young children through curiosity of the device.

As proxy reported tools are most frequently used when assessing behaviours at scale, here I will present some of the biases that may be more likely when measuring movement behaviours of young children using proxy reported measurement tools. Firstly, recalling activity behaviours is a cognitively demanding task. As such, pre-school aged children are not able to perform the detailed recall required to report their own activity, and so their activity behaviours must be proxy reported usually by a parent or carer (Oliver, Schofield and Kolt, 2007). As the parent or carer may not be with their child 24 hours of the day, it is plausible that activities will be missed or incorrectly reported, as estimations of what children do during the nursery day will be made. Proxy reported tools are more prone

to recall bias (Troost, 2007), whereby parents or carers incorrectly recall activity because they do not correctly remember. The degree to which recall bias may be present will depend on the type of tool being used and duration of recall, for example, past month would be more prone to bias than the past 24 hours. Additionally, questionnaires asking for 'habitual' behaviour risk participants recalling more active / more socially desirable weeks (Mindell, Coombs and Stamatakis, 2014).

Secondly, with proxy (and self) reported tools, there is increased chances of social desirability bias, which refers to participants giving responses that are perceived to be desirable to present themselves in a favourable way, instead of providing responses that are reflective of their actual behaviour. Research has shown that social desirability is associated with over-reporting of PA, leading to overestimation of PA duration and PA energy expenditure (Adams et al., 2005). It is plausible that when using these methods, parents may report their child's movement behaviours in line with what is perceived to be desirable.

Further, with self/proxy reported tools, there is the concern that questions may be misunderstood or incorrectly interpreted. One way in which this may be overcome is through piloting the tool prior to use, to identify if there are any questions in particular where this may be the case.

It is possible to minimise biases and potential sources of error in research through careful study design e.g. considering appropriate recall time of proxy reported tools, piloting tools prior to use. However, bias is ultimately an unavoidable issue in any research project using any measurement tool, and so it is important to be aware that some bias will be present and these must be considered in the interpretation of results.

2.11 Population level measurement of health behaviours

In this section I will outline the importance, challenges, and feasibility concerns of large scale surveillance of health behaviours generally, across adults and children, including pre-school aged children. I will then outline the current activity surveillance systems specific to pre-school aged children. I will finish this section by providing a discussion on the available evidence on the social disparities and inequality in surveillance and measurement of health behaviours across the general population, demonstrating the need for ensuring tools are equally as acceptable across all SES groups.

2.11.1 Importance of large scale surveillance of health behaviours

Population surveillance of movement behaviours is an important part of non-communicable disease surveillance, by assessing and monitoring risk of chronic disease through regular surveillance of key behavioural and physiological contributing factors (Pedišić and Bauman, 2015). The *Global Action Plan on Physical Activity 2018–2030*, published by WHO, identified regular population surveillance as a key strategic action (WHO, 2018). In adults, the development of the International Physical Activity Questionnaire (IPAQ) and the Global Physical Activity Questionnaire (GPAQ), led to improved standardised PA population data between countries (Hallal et al., 2012), which helped to provide clearer indication on global PA and inactivity risks, and to guide policy and practice including updated guidelines (Hallal et al., 2012). International surveillance systems for measuring PA of children are primarily targeted at adolescents, with no data presented for early years children at the time Aubert and colleagues conducted their review in 2020 (Aubert et al., 2021). Additionally, only 2 of 8 initiatives (in 5 to 17 year olds) had the ability to provide estimates of meeting the 2019 WHO guidelines (Aubert et al., 2021). Monitoring

and surveillance of movement behaviours of young children is said to be the first important step to promote healthy movement (Okely et al., 2018). A standardised approach to surveillance of movement behaviours is important for policy making and communication to the public (Brownson et al., 2000). Population surveillance of the movement behaviours are essential to:

- Determine prevalence rates and identify if recommendations and guidelines are being met (Biddle et al., 2011; Tremblay et al., 2016)
- Provide the ability to detect changes in these behaviours over time (Shephard, 2003) and to identify associations between the behaviours and other health and developmental outcomes, and map changes against health concerns such as childhood obesity (Okely et al., 2018)
- Monitor changes in population activity and help develop understanding of determinants and inequalities of movement behaviours (Okely et al., 2018)
- Help the development, monitoring, and evaluation of public health policy and practice, including the development of more specific recommendations which will in turn help guide the development and evaluation of interventions, policies and programmes to improve levels of movement behaviours (Chaput et al., 2014a; Chaput, 2019; WHO, 2007).

2.11.2 Challenges and feasibility concerns for large scale measurement

Movement behaviours are particularly difficult to measure as a result of their complexity and the multifaceted nature of these behaviours. As discussed in sections 2.7 and 2.8, all methods of measuring movement behaviours contain their own strengths and limitations, and the most appropriate methods depend on the aspect of the behaviour that

is being measured, the population of interest, the different aims of the research and available resources (Ainsworth et al., 2015).

Although there is growing interest in the use of devices for surveillance of PA and related behaviours, with some evidence of success in adult populations such as the UK biobank (Doherty et al., 2017), there are some key concerns that may limit their use. These include the short life span of devices, due to continuous technological advances and development of new and improved activity monitoring devices, with devices typically being prominent in the market for only 1 or 2 years before newer 'better' devices are available (Troiano, Stamatakis and Bull, 2020), making standardisation within and between studies and countries difficult (Bauman, Pedišić and Bragg, 2016). There have been vast technical advances in accelerometry over the past 20 years, including: larger memory and battery, smaller size, and reduction in cost (Troiano et al., 2014). Despite this, a standard accelerometer suitable for PA research purposes is still relatively expensive and there is ongoing debate around the affordability of devices (Pedišić and Bauman, 2015), which may especially be the case where budgets are restricted or movement behaviours are not the primary research outcome. Administration of devices, device malfunctions, loss or failure to return devices, and processing of the data can result in a large financial cost when the accelerometers are being distributed on a very large scale (Lee and Shiroma, 2014). Further, there is not only the cost associated with purchasing the devices, but experts would be needed for the data cleaning, analysis and interpretation of accelerometer data which would be time intensive and may not be feasible at population level with such large samples. Non-compliance with accelerometer wearing protocols has been outlined as a major issue when accelerometry was used at scale in the US National Health and Nutrition Examination Survey (Troiano et al., 2014). Similarly, there are logistical and administrative

concerns due to the need to distribute and collect accelerometers, initialise and calibrate devices and appropriately instruct participants (Pedišić and Bauman, 2015). This can extend to difficulties with keeping track of the stage of each participant, with days when accelerometers were posted to participants being incorrectly included in analyses as wear days (Keadle et al., 2014). For large scale measurement, remote data collection may be most viable, but there is still the need to ensure that participants are appropriately briefed and instructed on the accelerometer wearing protocol, and researcher assistance is still essential (Troost, McIver and Pate, 2005). The use of device based measurement tools for international surveillance may be problematic by increasing the gap between low and high income countries, with concerns of non-compliance (Aubert et al., 2021).

Further, device based tools are not the optimal tool for compliance with the guidelines that were developed based on evidence from self-report data. Although accelerometers are being used much more frequently in research and, therefore, there is more knowledge around accelerometer derived activity, this would not be coherent with the data in which the guidelines were primarily developed. As such, it would not be appropriate to fully transition from measuring the guidelines based on self-reported data using device based tools (Troiano, Stamatakis and Bull, 2020). As discussed in sections 2.7.5 (p.76) and 2.8.5 (p.102), there are many key interpretation choices and decisions that need to be considered when using accelerometers in any study, including: type, placement, cut points, wear time, non-wear time. This is a large area of concern for the use of accelerometers for national and global surveillance for ensuring consistency. A lack of standardisation and use of differing methods of data collection can result in discontinuity in time-trend data, meaning that using accelerometers for large scale surveillance does not always provide meaningful data for comparisons between studies, or at national and

international levels (Pedišić and Bauman, 2015). Thus, researchers have argued that the widespread use of accelerometers for population surveillance may be premature (Pedišić and Bauman, 2015). This highlights the continued need for survey-based methods that can be used to collect prevalence data using a standardised approach, important for surveillance of activity and inactivity at national and global levels (Hallal et al., 2012). Although it may be that accelerometers can begin to be introduced into surveillance to complement self/proxy reported data to provide additional insights into movement behaviours at population level.

Accelerometers measure acceleration and movement, providing information on duration of movement, but they do not provide any contextual information or describe the type of behaviour being performed. From accelerometer data, it is not possible to determine what activities children are engaging in. For example, we may see that there is limited movement that would be in line with the intensity of SB (Reilly et al., 2015), but it would not be possible to distinguish whether this time is engaged in screen time or reading and storytelling (Carson et al., 2015). Additionally, when examining sleep alongside these behaviours, it is not necessarily possible to determine sleep from SB /non-wear time from the accelerometer data. In line with this, sleep diaries are often used alongside accelerometers to increase the accuracy of the output data (Tétreault et al., 2018). In some circumstances, such as large population studies, the use of sleep diaries alongside accelerometry may be too onerous and not feasible due to participant burden (Quante et al., 2015). Although accelerometers provide useful information on movement, they should not be regarded as a better or interchangeable tool with self-report that includes reports on behaviours (Troiano et al., 2014).

In recent years, there has been an increase in research to determine whether smartphones and consumer wearables could be used for population level surveillance (Althoff et al., 2017). However, there is large concern regarding the low quality and inaccurate data that is produced by such methods (Brodie et al., 2018). The fast moving technology industry results in new devices continuously being updated and renewed, variability in the way in which the data from the devices is sourced, and the constantly changing algorithms used in consumer wearables based on sensor development and technological advances, without user or researcher knowledge limit the feasibility of such devices for population surveillance (Fuller et al., 2020). Further, representative sampling, ways in which devices are used (e.g. wear time placement), measurement properties of the tools and compatibility between different consumer devices limits the use of consumer wearables for population surveillance at this time (Strain et al., 2022). Other concerns with the use of smartphone and commercial devices for PA population level surveillance include that just because individuals have these devices it does not necessarily mean that they would be happy for their data to be downloaded and used for national surveillance purposes which could lead to a further bias in the data. Most importantly, for measurement research with pre-school children, it is unlikely that children of this age would have access to their own smartphone or consumer wearable, as such, the degree to which these methods could be used to survey movement of young children is questionable.

This taken together highlights the continued importance for the use of self/proxy reports as a primary measure for population surveillance at this time (Sattler et al., 2021; Troiano, Stamatakis and Bull, 2020). Further to the reasons just outlined, self and proxy reports are a cost-effective way of collecting information from large cohorts of individuals

on the types and domains of PA (Sattler et al., 2021), and at population level are often administered alongside a host of other questions assessing various different behaviours e.g. Health Survey for England (NHS Digital, 2020).

2.11.3 Current activity surveillance and monitoring systems

In the UK, there is a lack of standardisation of surveillance measurement of PA and SB across the home nations (Strain et al., 2020). There are also no measures specific to assessing the PA and SB recommendations of pre-school children (NHS Digital, 2016; NHS Digital, 2020) and no measures assessing sleep. Given the importance of sleep for health (evidenced by international guidelines) it is important that sleep is also considered in national surveys as a public health issue (Chaput, 2019).

In the UK, the devolved nations have separate health surveys; England, Wales and Scotland all have health surveys that collect data on PA of young children. However, the age range targeted varies across the surveys, the questions are not consistent across nation surveys and are not specific to the movement behaviours or recommendations for pre-school children. The National Survey for Wales collects data on PA of children aged 3-7 years, asking parents to report the PA their child engaged in each day for the past 7 days using prescribed categories from '*none*' to '*more than an hour*' (Welsh Government, 2018). The Scottish Health Survey collects data for children aged 2-15 years, but groups this age category together and does not report on guideline adherence in children aged under 5 (Scottish Government, 2020; Strain et al., 2020). The Health Survey for England is the only UK survey that analyses data on national PA prevalence for pre-school aged children separately from older children (NHS Digital, 2016; Strain et al., 2020). The Health Survey for England includes a standardised set of questions to assess PA prevalence across the ages of

2-15 years, with the only difference being that children aged 13-15 years are asked to self-report, whilst parents report for those aged 2-12 years (NHS Digital, 2016). The standardised measure is used despite the difference in recommendations between these age ranges (Chief Medical Officers, 2019; Department of Health and Social Care, 2011), along with the vast difference in development; growth and maturation, throughout childhood and adolescence (Malina, Bouchard and Bar-Or, 2004), and differences in the activities in which children engage in (Sims et al., 2021). Further to this, the questionnaire used as part of the Health Survey for England showed an overestimation of activity levels of children in comparison with the results obtained from accelerometry in 130 6-7 year old children (Basterfield et al., 2008).

Across the globe, the Australian Health Survey sporadically asks questions on PA, including outdoor and indoor PA/active play, and sedentary screen time of pre-school aged children. However, the most recent survey including this age group was in 2011/12 (Australian Bureau of Statistics, 2013), with children aged 15 years and over only included in the most recent two survey rounds: 2014/15 (Australian Bureau of Statistics, 2016) and 2017/2018 (Australian Bureau of Statistics, 2018). The New Zealand Health Survey asks questions for pre-school aged children in relation to screen time, but PA questions are restricted to children aged 5 years and older only (Ministry of Health New Zealand, 2017). The Canadian Health Measures Survey is the first national data collection scheme to collect data on the movement behaviours of pre-school aged children using accelerometers for PA estimates, two questions on average hours of screen time for weekday and weekend days separately, and one question on the usual number of hours the child sleeps within a 24-hour period (Chaput et al., 2017a; Tremblay and Gorber, 2007).

The initiative 'International Study of Movement Behaviors in the Early Years' (SUNRISE) has been presented as the first of its kind to attempt to measure the WHO guidelines across countries of different gross domestic product, as well as to make evidence informed recommendations based on the associations of these behaviours with health and developmental outcomes; at scale, with children from different countries (SUNRISE, 2020; Okely et al., 2021b). A further initiative is the Sleep and Activity Database for Early Years (SADEY), a recently funded project that aims to create a harmonisation of device based surveillance of young children's movement behaviours (Cliff et al., 2021).

However, it has been recommended that further efforts are needed to develop surveillance methods to capture the new WHO guidelines (Troiano, Stamatakis and Bull, 2020). In particular, a specific focus is essential on key priority groups who are currently under-represented in surveillance systems, one of which is pre-school aged children (Aubert et al., 2021; Chen et al., 2021; Troiano, Stamatakis and Bull, 2020). This is particularly the case in the UK where we currently do not have any monitoring systems in place (although Scotland are involved in the SUNRISE study, England, Wales and Northern Ireland are currently not involved) (Okely et al., 2021b). There is a need to advance the evidence base on movement behaviours in the UK. The methods used for large scale studies and surveillance need to undergo some element of validation to ensure the questions/method captures the desired outcome (Chaput, 2019), and to ensure that a tool is available that is low burden for participants and feasible for use.

2.11.4 Social disparities of surveillance and measurement of health behaviours

As discussed in section 2.11.1 (p.125), population surveillance through health surveys provide valuable information about the health of the general population that can be

used for policy making and health promotion activities. The representativeness of the data obtained through the surveys is highly dependent on the sample, as outcomes will only be known for those who participate in the surveys. There is evidence to suggest that response rates to national health surveys is decreasing (De Heer and De Leeuw, 2002; Mölenberg et al., 2021). There is also a body of evidence to suggest that there are disparities in the sociodemographic characteristics in response rate to health surveys used for measurement at population level (Galea and Tracy, 2007; Korkeila et al., 2001; Tolonen et al., 2005). Individuals from lower SES backgrounds are less likely to participate in health surveys (Demarest et al., 2013; Harald et al., 2007; Tolonen et al., 2006; Van Loon et al., 2003), whilst individuals from high SES backgrounds are more likely to participate (Boshuizen et al., 2006; Galea and Tracy, 2007). Reduced participation in health surveys have been shown by parents, teachers and children from more deprived areas (as assessed by the Index of Multiple Deprivation), as opposed to their counterparts in less deprived areas (Goodman and Gatward, 2008). The reduced levels of participation seen in health surveys can lead to bias and inaccurate estimates of health behaviours and poor health, with the possibility of the survey outcomes not being representative of the whole population which impacts on the generalisability and usefulness of the findings (Van Loon et al., 2003). Additionally, this may lead to inappropriate provision and services. It is important to accurately measure health behaviours across different socioeconomic groups, and to ensure that lower SES groups are represented in measurement property assessment research. As such, it is fundamental that efforts are made to ensure that population level surveys are accessible to lower SES groups (Demarest et al., 2013), to ensure that interventions in response to prevalence data do not widen inequalities.

Additionally, as a result of worse health displayed in lower SES communities, a large number of interventions are targeted specifically at this group, in an attempt to improve health and reduce inequalities. A few examples of such initiatives aimed at pre-school children in the UK include:

- 1) '*A Better Start*', an initiative taking place in five of the most deprived communities in England. This includes the implementation of science and evidence based preventative programmes, policies and services, to improve outcomes and health of children from disadvantaged backgrounds (A Better Start, 2021).
- 2) HENRY, an intervention targeted at parents and their young children, which aims to improve nutrition, emotional wellbeing, parenting skills and breastfeeding, along with increasing activity behaviours of parents and their young children (HENRY, 2020).
- 3) The Startwell project, a training programme to improve PA and nutrition of pre-school children, as well as improving environmental factors including opportunities for active play (The Startwell Project, 2020).

In order to understand health behaviours across groups, including those in low SES communities, and to recognise the effectiveness of targeted and universal initiatives and interventions in place, it is important to ensure that measurement tools are appropriate and accessible across SES groups. As such, the review work in this thesis will include extracting data specific to the socioeconomic contexts of participants, and recruitment of the empirical work focused on areas of high deprivation (as determined by the Index of Multiple Deprivation).

2.12 Impact of the COVID 19 pandemic on children

The recent COVID-19 pandemic has highlighted and exacerbated pre-existing inequalities (Marmot et al., 2020b). The most devastating impacts of the pandemic are likely to fall on children as they grow and develop (Sinha, Bennett and Taylor Robinson, 2020; Whitehead, Taylor-Robinson and Barr, 2021). The COVID-19 pandemic has impacted on health and development of children that extend far beyond PA, SB, and sleep. There was large disruption to early years provision, with pre-school aged children not attending settings for a period of time and rather spending most of their time in their home environments due to lockdowns (Hefferon et al., 2021). Children from disadvantaged backgrounds were less likely to have the resources, receive support to replicate learning done in schools, and more likely to have parents working outside the home who were unable to provide their children with support in the same way that wealthier parents who were home were able to (Marmot et al., 2020b). The loss of learning and socialisation as a result of the pandemic is likely to impact children living in poverty and from disadvantaged backgrounds disproportionately, which may subsequently widen the gap in attainment between poorest and wealthy families (Sinha, Bennett and Taylor Robinson, 2020; Whitehead, Taylor-Robinson and Barr, 2021). Preliminary evidence suggests that there are large educational and developmental impacts on pre-school aged children, including a predicted 66% decrease in reading ability due to school closures than what would usually be expected of children of this age (Bao et al., 2020).

Alongside this, deteriorating family finances, poverty and overcrowded households had an effect on parents' capacity to support their young children during lockdowns; lower income parents were shown to have experienced greater stress when young children were

at home, with worsening mental health within families (Hefferon et al., 2021). A review of the literature has suggested that the limited movement and interaction young children have faced during the COVID-19 pandemic will have long term impact on social skills and increased risk of physical difficulties through insufficient activity as a result of the COVID-19 pandemic (López-Bueno et al., 2021).

The aftermath of the COVID-19 pandemic (including a predicted global recession) is likely to exacerbate child poverty and have immeasurable impacts on child health, learning, and development (Hefferon et al., 2021; Marmot et al., 2020b). To negate some of the devastating impacts of the pandemic on child health and health inequalities, direct and collaborative policy focus on, and with, children is required in the recovery (Hefferon et al., 2021; Larcher and Brierley, 2020). In particular, early years have been suggested as an area that must be prioritised, including education, health and welfare services being appropriately funded and supported (Kyeremateng, Oguda and Asemota, 2021; Marmot et al., 2020b).

2.12.1 COVID 19 and movement behaviours of pre-school aged children

The COVID-19 pandemic has had a substantial impact on the movement behaviours of young children (Clarke et al., 2021; Owen and Bould, 2021). There has been a global effect of the COVID-19 pandemic on the movement behaviours of pre-school aged children, with children engaging in more sedentary screen time (Aguilar-Farias et al., 2020; Hossain et al., 2021; López-Bueno et al., 2020; Okely et al., 2021a), more overall sedentary time (Alonso-Martínez et al., 2021), less PA (Aguilar-Farias et al., 2020; Alonso-Martínez et al., 2021; Hossain et al., 2021; López-Bueno et al., 2020; Tulchin-Francis et al., 2021), less time outdoors (Hossain et al., 2021; Okely et al., 2021a), later bed and wake up times and

napping less (Liu et al., 2021; Okely et al., 2021a) than before COVID-19. There is mixed evidence for sleep quality, with proxy reports indicating reduced sleep quality than before COVID-19 (Aguilar-Farias et al., 2020), but alternative research suggesting lower levels of overall sleep disturbance through less bedtime resistance, sleep onset delay, sleep anxiety, night wakings and parasomnias (Liu et al., 2021). Sociodemographic factors including living situation (e.g. location, space, living in house versus apartment) and parental education level were associated with differences in movement behaviours recorded (Aguilar-Farias et al., 2020; Kracht, Katzmarzyk and Staiano, 2021). PA and sedentary screen time levels for children from lower to middle income countries were less impacted by the pandemic than high income countries, explained by children in lower to middle income countries being in rural areas with more access to outdoor space (Okely et al., 2021a). In Sweden, increases in the PA and outdoor time of a small sample of pre-school aged children were seen, but there were also reports of increased screen time (Nystrom et al., 2020). However, Sweden had less restrictions in place during the pandemic, including pre-schools and childcare centres remaining open throughout.

Changes in emotional and behavioural states were observed in pre-schoolers during the COVID 19 pandemic (Aguilar-Farias et al., 2021; Alonso-Martínez et al., 2021). Research with toddlers and pre-schoolers in Latin America demonstrated that changes in emotional and behavioural states during the COVID-19 pandemic were associated with movement behaviours of the children. Maintaining pre-pandemic levels of PA and increases in sleep duration were associated with children being more affectionate and displaying less negative emotions (e.g. less worried, angry, sad, afraid), whilst increases in screen time were associated with increased negative emotions and behaviours (e.g. aggression, irritability, frustration, tantrums) (Aguilar-Farias et al., 2021).

The importance of healthy movement behaviours in pre-school aged children is amplified by the recent pandemic (Guan et al., 2020a). It has been suggested that the promotion of movement behaviours in children should be a key part in public health messaging, in the response and recovery from COVID-19 (Guan et al., 2020a). Appropriate methods of measurement for assessing movement behaviours of young children need to be in place to advance the evidence base for health promotion in this area.

3 [Chapter Three: A systematic review of the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children.](#)

The study in this chapter has been published in the International Journal of Behavioural Nutrition and Physical Activity:

Phillips, SM., Summerbell, C., Hobbs, M., Hesketh KR., Saxena S., Muir, C., and Hillier-Brown, FC. (2021). A systematic review of the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children. *International Journal of Behavioural Nutrition and Physical Activity*, 18(1), 141.

3.1 Prelude

In the endeavour to assess the best way of measuring movement behaviours of pre-school children at scale, knowing about the measurement properties of tools used to assess physical activity and sedentary behaviour was an important first step. I scoped the literature and it was apparent that there were currently no available overviews on the state of the evidence base on the validity, reliability, and feasibility of tools used to examine physical activity and sedentary behaviour of pre-school aged children. On this basis, the first step of the PhD was to conduct a study to fill this knowledge gap, to review the evidence base and determine the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children. In this review, I deliberately extracted information relating to the socioeconomic status and ethnicity of participating families, to determine representation of different groups in the pre-existing measurement tool literature.

3.2 Introduction

This chapter presents a systematic review of the validity, reliability and feasibility of measurement tools used to examine physical activity (PA) and sedentary behaviour (SB) of pre-school aged children. On starting this PhD, which aimed to assess options for measuring movement of young children, it was important to search the literature to identify what tools were already available, and the measurement properties of the tools, to determine which tools were valid, reliable, and feasible for measuring the PA and SB of pre-school children, (Bauman et al., 2006; Kohl, Fulton and Caspersen, 2000). On examining the literature, it was clear that there were no up-to-date recommendations, guidance, or available summaries on

the evidence base of measurement properties (validity, reliability, feasibility) of pre-existing tools to measure PA and SB in pre-school aged children.

Previous reviews had examined measurement properties of tools but had focused on selected measurement tools (Adamo et al., 2009; Kohl, Fulton and Caspersen, 2000; Loprinzi and Cardinal, 2011; Lubans et al., 2011; Pate, O'Neill and Mitchell, 2010; Sirard and Pate, 2001), including those which focused specifically on questionnaire based measures (Chinapaw et al., 2010; Helmerhorst et al., 2012; Hidding et al., 2018; Hidding et al., 2017). These reviews did not examine the full range of measurement tools used to assess PA and SB, and did not focus specifically on pre-school aged children. One review was identified by Oliver and colleagues, conducted in 2007, where authors had summarised studies that had examined the validity and reliability of a range of measurement tools used to assess PA, as well as examining prevalence rates of PA, of pre-school aged children (Oliver, Schofield and Kolt, 2007). However, a rapid scoping search of relevant studies identified a large number of potential studies for inclusion that were published after Oliver et al's review. Additionally, this review did not examine tools used to assess sedentary behaviour: the most likely reason for this is that the important associations of SB with health and developmental outcomes have only started to emerge in the last 10-15 years (LeBlanc et al., 2012; Poitras et al., 2017; Vale et al., 2010). However, many tools developed and used in this time have focused on measuring both behaviours simultaneously.

Prior to developing any new tools, or using any tools in this research, it was important to examine and understand the full evidence base on all tools that already exist, and the validity, reliability and feasibility of these tools, to measure both PA and SB specifically in the pre-school population.

3.3 Aims and objectives

The aim of this study was therefore to:

- Systematically review studies examining the validity, reliability, and feasibility, of measurement tools used to assess PA and/or SB in pre-school aged children.

The objectives of this study were to:

- Identify, collate, and synthesise studies on all pre-existing tools (that had been evaluated) used to examine PA and/or SB of pre-school aged children
- Determine which tools were valid, reliable, and feasible for measuring PA and/or SB of pre-school aged children

3.4 Methods

To achieve this aim, a systematic review was the method of choice to identify the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children. A systematic review is the most robust, comprehensive, and intensive way of examining the literature within a given field. The purpose of a systematic review is to identify, collate and synthesise all research on a given topic, to provide an understanding and up-to-date recommendations on the evidence base that can help guide decisions of researchers and practitioners (Lasserson, Thomas and Higgins, 2021). As there were no currently available reviews of this kind, it was an important first step in the PhD in order to thoroughly examine and determine the current state of the evidence.

This systematic review was conducted and is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria

(Moher et al., 2009) (see appendix 3.1). The protocol for this systematic review was registered at the International Prospective Register for Systematic Reviews (PROSPERO), registration number CRD42019133613 (see appendix 3.2, online file can be accessed [here](#)). No ethical approval was needed for conducting this study.

3.4.1 Search Strategy

Systematic searches of seven major electronic databases (Scopus, Web of Science, PsycARTICLES, PsycINFO, MEDLINE, CINAHL and SPORTDiscus); topic specific journals (*Journal for the Measurement of Physical Behaviour* and *Pediatric Exercise Science*); and the grey literature (opengrey.eu and Research Gate), were conducted to identify relevant studies. Searches were conducted in March 2019, and updated in March 2020. No restrictions were placed on language, year of study, or publication status.

The search strategy included combinations of key words relating to the: construct (physical activity, sedentary, sitting), population (pre-school, early years, early childhood, young children and kindergarten) and measurement properties (assessment, measurement, method, valid, reliable, feasible). Searches were adapted to each database, alongside the use of appropriate boolean operators and database specific filters (see Table 3.1). I worked with colleagues at the Durham University information science team, to refine the search strategies. References and citation searches of included studies, as well as checking the reference lists of selected existing reviews (Adamo et al., 2009; Chinapaw et al., 2010; Helmerhorst et al., 2012; Hidding et al., 2017; Hidding et al., 2018; Kohl, Fulton and Caspersen, 2000; Loprinzi and Cardinal, 2011; Lubans et al., 2011; Oliver, Schofield and Kolt, 2007; Pate, O'Neill and Mitchell, 2010; Sirard and Pate, 2001) were conducted for completeness.

Table 3.1. Search strategy and outcomes

Source of Search Conducted	Date of Search	Searches Conducted	Search Outcome
Scopus- <i>Science Direct</i>	05/03/20	(TITLE-ABS-KEY ("physical activit*") AND TITLE-ABS-KEY (sedentary OR sitting) AND TITLE-ABS-KEY (valid* OR reliab* OR feasib*) AND TITLE-ABS-KEY (pre*school OR "early years" OR "early childhood" OR "young children" OR kindergarten*) AND TITLE-ABS-KEY (assess* OR measure* OR method))	203
		(TITLE-ABS-KEY ("physical activ*") AND TITLE-ABS-KEY (assess* OR measure* OR method) AND TITLE-ABS-KEY (valid* OR reliab* OR feasib*) AND TITLE-ABS-KEY (pre*school OR "early years" OR "early childhood" OR "young children" OR kindergarten*))	878
		(TITLE-ABS-KEY (sedentary OR sitting) AND TITLE-ABS-KEY (assess* OR measure* OR method) AND TITLE-ABS-KEY (valid* OR reliab* OR feasib*) AND TITLE-ABS-KEY (pre*school OR "early years" OR "early childhood" OR "young children" OR kindergarten*))	346
Web of Science- <i>Web of Science</i>	05/03/20	ALL FIELDS: (physical activit*) AND ALL FIELDS: (sedentary OR sitting) AND ALL FIELDS: (assess* OR measure* OR method) AND ALL FIELDS: (valid* OR reliab* OR feasib*) AND ALL FIELDS: (pre\$school OR "early years" OR "early childhood" OR "young children" OR "kindergarten")	282
		ALL FIELDS: (sedentary OR sitting) AND ALL FIELDS: (assess* OR measure* OR method) AND ALL FIELDS: (valid* OR reliab* OR feasib*) AND ALL FIELDS: (pre\$school OR "early years" OR "early childhood" OR "young children" OR "kindergarten")	329
		ALL FIELDS: (physical activit*) AND ALL FIELDS: (valid* OR reliab* OR feasib*) AND ALL FIELDS: (pre\$school OR "early years" OR "early	993

		childhood" OR "young children" OR "kindergarten") AND ALL FIELDS: (assess* OR measure* OR method)	
CINAHL, PsycARTICLES, PsycINFO, Medline SPORTdiscus- EBSCOhost	05/03/20	TX "physical activit*" AND TX (sedentary OR sitting) AND TX (assess* OR measure* OR method) AND TX (valid* OR reliab* OR feasib*) AND TX ("pre#school" OR "early childhood" OR "early years" OR "young children" OR "kindergarten")	487
		TX (sedentary OR sitting) AND TX (assess* OR measure* OR method) AND TX (valid* OR reliab* OR feasib*) AND TX ("pre#school" OR "early childhood" OR "early years" OR "young children" OR "kindergarten")	1,789
		TX "physical activit*" AND TX (assess* OR measure* OR method) AND TX (valid* OR reliab* OR feasib*) AND TX ("pre#school" OR "early childhood" OR "early years" OR "young children" OR "kindergarten")	2,293
Journals- Pediatric Exercise Science <i>Human Kinetics Journals</i>	05/03/20	"physical activity" OR "sedentary" OR "sitting" AND "pre?school" OR "early years" OR "early childhood" OR "young children" OR "kindergarten"	63
Journals- Journal for the Measurement of Physical Behaviour- <i>Human Kinetics Journals</i>	05/03/20	"physical activity" OR "sedentary" OR "sitting"	69
Grey Literature- Opengrey.eu	05/03/20	"physical activity" OR "sedentary" AND "preschool" OR "early years" OR "young children"	1408
Grey Literature- Research Gate	05/03/20	"physical activity" OR "sedentary" AND "preschool" OR "early years" OR "young children"	310

3.4.2 Eligibility Criteria

3.4.2.1 Inclusion Criteria

Studies were eligible for inclusion if their aim was to examine the measurement properties (validity and/or reliability and/or feasibility) of a tool used to measure PA and/or SB of a general population sample of pre-school aged children between the ages of 3 and 7 years old. This broad age range for what constitutes a 'pre-school aged child' was selected to ensure international relevance, using the definition of '*any child who has not yet reached the age of formal schooling*'. Although this is usually aged between 3 and 5 years old, this varies internationally as indicated by the inclusion criteria (World Bank Data, 2020).

Additionally, in line with others, the current review included studies where the mean age of children was 6 years or below to ensure children who had not entered primary school were captured (Carson et al., 2015). Where studies had sampled children of different age categories together, we observed whether these studies performed stratified analyses by age groups, isolating the preschool children from the other age groups; if this was the case, the study was included and information concerning pre-school children was reported. If this was not possible, the study was excluded (Ortega et al., 2015).

3.4.2.2 Exclusion Criteria

Studies were excluded from the review if:

- 1 The aim of the study *was not* to examine the measurement properties of the tool.

For example, studies were excluded if the aim was to examine the reproducibility or tracking of behaviours over time, inter-observer reliability, epoch lengths, wear time, the calibration of cut points or prediction equations (with no separate validation study);

- 2 The aim of the study was *not* to examine a tool measuring PA and/or SB. For example, studies were excluded if their aim was to examine physical fitness, motor skills, PA environment, correlates of PA or SB, or the impact of interventions;
- 3 The study included children under 3 years or over 7 years of age, or if the population sample included children with chronic conditions;
- 4 The study was a study protocol or a review. Higher degree theses and conference abstracts were included, however, where the relevant information from the theses were also provided in published peer-reviewed journal articles, the journal article was included and the thesis excluded.

3.4.3 Screening for relevant studies

Following the searches, all identified articles were imported into a referencing manager software (Endnote X9.1) and duplicates were removed. I screened titles and abstracts of included articles for inclusion, and a further 10% were screened by a second reviewer (Dr Matt Hobbs, co-author on the publication of this review). There was very high agreement between the two reviewers for title and abstract screening, with a Cohen's kappa statistic of $k=0.94$ (McHugh, 2012). We discussed any discrepancies on the eligibility of particular studies, without the need for escalation to a third reviewer. I screened full texts of potentially relevant studies, which were then double screened by one of two other reviewers (Dr Matt Hobbs and Dr Frances Hillier-Brown (supervisor and co-author) for inclusion. An inclusion checklist was used to determine whether each study met the eligibility criteria (see appendix 3.3).

3.4.4 Data extraction

I extracted data from all relevant studies, and the studies were extracted independently by one of two other reviewers (Dr Frances Hillier-Brown or Cassey Muir (co-author of on the publication of this review)) using a pre-piloted data extraction form (see appendix 3.4). Any discrepancies were resolved by discussion. Extracted information included: study characteristics, participant characteristics, the measurement tool explored (e.g. accelerometry), the measurement tool(s) used as a comparison, data interpretation choices for device based measurement tools (e.g. cut points, epoch, placement), statistical method used to compare measurement tools, behaviour category (PA and/or SB), and the details of the units of measure (e.g. moderate-to-vigorous PA), measurement properties assessed (e.g. criterion validity), the results of the study, and the sources of funding for the study.

3.4.5 Risk of bias

Assessing risk of bias of studies included in systematic reviews is important to prevent inappropriate recommendations being made, and to be able to acknowledge the quality of the evidence on which the conclusions are made (Guyatt et al., 2008). I extensively searched various risk of bias assessments to see which would be most suitable for studies of this kind. There are multiple different risk of bias assessment available, however, these are mostly developed for assessing risk of bias of intervention based studies, which aim to identify flaws in the design between the intervention and outcome of interest (e.g. Effective Public Health Practice Project, 2009; Sterne et al., 2019). Previous reviews examining measurement properties of tools used to assess PA or SB focusing on questionnaires assessed the risk of bias of included studies using the Consensus-based

Standards for the selection of health Measurement Instruments (COSMIN) (Hidding et al., 2017; Hidding et al., 2018; Mokkink et al., 2010a; Mokkink et al., 2018a; Mokkink et al., 2018b) and the Qualitative attributes and measurement properties of physical activity questionnaires (QAPAQ) or similar modified versions (Chinapaw et al., 2010; Lubans et al., 2011; Terwee et al., 2010). Despite this, the COSMIN checklist has been devised specifically for the assessment of patient reported outcomes (Mokkink et al., 2010a; Mokkink et al., 2018a; Mokkink et al., 2018b). Whilst the QAPAQ is also specific to physical activity questionnaire-based outcomes (Terwee et al., 2010). These two checklists have a strong emphasis on questionnaire design and development which were not applicable for all studies as the aim of this review was to include all types of measurement tool, including device based tools. A further risk of bias tool considered was the QUADAS-2 which assesses risk of bias of studies examining diagnostic accuracy (Whiting et al., 2011), but I decided not to use this tool as it would not be an applicable tool to assess studies on all measurement properties including in this review, in particular studies examining feasibility. In discussion with the review team, we wanted to use one risk of bias assessment that we could apply consistently to all of the studies, irrespective of the measurement property being examined. A review of risk of bias assessments identified the Downs and Black instrument (Downs and Black, 1998) and the Newcastle-Ottawa Scale (Wells and Shay, 2000) as the most useful tools for assessing risk of bias of non-randomised studies (Deeks et al., 2003). Additionally, previous systematic reviews examining PA assessment measures had successfully used a modified version of the Downs and Black checklist (Adamo et al., 2009; Prince et al., 2008). As such, I decided that this was the most appropriate assessment checklist to use on all studies included in the review.

Dr Frances Hillier-Brown and I independently conducted the risk of bias assessment on each of the included studies, with any discrepancies resolved through discussion. The risk of bias of all individual studies included in this review was assessed using the modified version of the Downs and Black Checklist (Downs and Black, 1998). The tool includes seventeen questions: eight focus on the quality of reported criteria, three on the external validity and five the internal validity. The maximum quality score a study could receive was 17, with study quality rated as good (13-17), fair (9-<13) or poor (<9), based on the protocols used in previous reviews (Jutai, Strong, and Russell-Minda, 2009; Korakakis et al., 2018; Maidment et al., 2018).

We conducted an additional risk of bias assessment on studies examining proxy reported measurement tools, using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) risk of bias checklist (Mokkink et al., 2018b; Prinsen et al., 2018). Although this could not be conducted on all studies, due to this tool being available, we felt it was useful to complete this additional risk of bias assessment. This checklist was devised specifically for assessing the risk of bias of patient reported measurement property studies. Based on the studies included in our review we conducted the assessment using the sub-sections relating to reliability and construct (convergent) validity. It was not possible to assess the quality of the content validity of the studies due to such minimal information available. Each item was scored using a 4 point scale (very good, adequate, doubtful, inadequate). The overall methodological rating of a study was determined using the COSMIN protocol of 'the worst score counts' principle (Terwee et al., 2012) (e.g. if the lowest rating of all items was 'doubtful', the overall methodological quality of the measurement property in that study would be rated as 'doubtful').

3.4.6 Definitions of validity, reliability and feasibility

There remains some debate over the terminology used for validity in the field of PA and SB measurement, particularly in relation to criterion validity (Kelly, Fitzsimons and Baker, 2016). For clarity, an overview is provided in Table 3.2 provides of the definitions of measurement properties that were examined in this review.

Table 3.2. Definition of each of the measurement properties included in this review

Measurement property	Definition
Validity	<i>Ability for a measure to accurately reflect the construct it is designed to measure.</i>
Criterion validity	Output of a measure produces similar results to a 'criterion' measure. This includes studies that have examined a tool for determining energy expenditure with calorimetry (including doubly labelled water) used as the criterion measure.
Convergent validity	Output of a measure produces similar results to a reference measure not considered a criterion.
Content validity	Extent to which a measure covers all aspects of the intended domains or dimensions that it claims to measure, includes face validity which is the appearance of a measure, in that it appears to measure what it claims to measure
Reliability	<i>Extent to which a tool gives measurements that are consistent, stable and repeatable.</i>
Test-retest reliability	The extent to which a measure can obtain similar results in repeated trials, keeping as many conditions stable as possible
Inter/Intra Instrument Reliability	The extent to which scores are consistent when measurements are taken by different versions of the same instrument (inter-instrument) or by the same version of an instrument repeatedly (intra-instrument)
Feasibility	The extent to which a measurement tool: is suitable for the target population; can be successfully delivered in the target population/context; shows promise of being successful within the intended population. This can include: participant acceptability, researcher acceptability and cost, which can be assessed for all measurement tools through qualitative feedback of participants and through missing or lost data occurred from the measurement tool (with the exception of proxy or self-reported tools that can only be determined through qualitative means including the comprehensiveness and relevance of items).

Definitions guided by: Bowen et al., 2009; Evenson, Goto and Furberg, 2015; Forouhi, Brage and Wareham, 2017; Kelly, Fitzsimons and Baker, 2016; Terwee et al., 2018a

Additionally, for the purpose of this review, we developed a level of evidence scheme to distinguish between validity studies that used different comparison measurement tools. Level 1 includes criterion validity studies; the only methods considered 'criterion' were calorimetry based methods (e.g. whole room calorimetry and doubly labelled water) when compared against a measurement tool aiming to measure energy expenditure. Levels 2-4 include convergent validity, separated by the quality of the comparison measurement tool used. Level 2 includes studies where a tool has been compared against a measure which is considered to have relatively high validity (but not criterion), such as direct observation. Level 3 includes studies where a tool has been compared against a measure which is considered to have relatively low validity, such as device based measurement tools. Level 4 includes the comparison of two (or more) of the same type of measurement tool where neither tool is considered to have known higher validity, such as the outcomes of two makes of accelerometer being compared. Table 3.3 provides a full explanation on what constitutes each level of evidence.

Table 3.3. Level of evidence for validity studies included in this review

Level of evidence	Explanation	Criteria of comparisons (examples)	
		Measurement tool under study	Comparison tool/measure
1	Criterion validity	Any measurement tool to determine energy expenditure	Calorimetry, including doubly labelled water
2	Convergent validity: measurement tool compared with a measure which is considered to have <i>relatively high validity</i> (but not considered criterion)	Direct observation protocols (newly devised)	Direct observation protocol (pre-existing)
		Device based measurement tools: combined heart rate and accelerometer, heart rate monitor, accelerometer, pedometer	Direct observation Electrodiagram (for heart rate monitors)
		Proxy reported measurement tool: questionnaire, diaries	Direct observation
3	Convergent validity: measurement tool compared with a measure which is considered to have <i>relatively low validity</i>	Direct observation protocols (newly devised)	Device based measurement tools: combined heart rate and accelerometer, heart rate monitor, accelerometer or pedometer
		Device based measurement tools: combined heart rate and accelerometer, heart rate monitor, accelerometer, pedometer	Device based measurement tools with known higher validity than tool under study: combined heart rate and accelerometer, heart rate monitor, accelerometer
		Proxy reported measurement tool: questionnaire, diaries	Device based measurement tools: combined heart rate and accelerometer, heart rate monitor, accelerometer or pedometer
4	Convergent validity: two (or more) of the same type of measurement tools being compared, where <i>neither tool is considered to have known higher validity</i>	Device based measurement tools: combined heart rate and accelerometer, heart rate monitor, accelerometer	Device based measurement tools: combined heart rate and accelerometer, heart rate monitor, accelerometer
		Proxy reported measurement tool: questionnaire, diaries	Proxy reported measurement tool: questionnaire, diaries

3.4.7 Interpretation and synthesis of results

3.4.7.1 *Validity and Reliability*

Studies commonly use a number of different statistical analyses to define validity and reliability (Forouhi, Brage and Wareham 2017; Kelly, Fitzsimons and Baker, 2016). These frequently include correlations (Pearson's; Spearman's; Kendall's; Intraclass), linear regressions, root mean square error (RMSE), Bland Altman, kappa statistics and receiver operating curve (ROC-AUC) (De Vet et al., 2011; Evenson, Goto and Furberg, 2015; Koo and Li, 2016; Lohr, 2002; Schmidt and Steindorf, 2006; Terwee et al., 2007; Twomey and Kroll, 2008). In order to demonstrate consistency in the interpretation of the results across studies, and also to compare and rank these results, I scoped the relevant literature to search for guidance. Through this searching and through discussions with my supervisors, it was apparent that currently there is no consensus on the values that constitute reasonable validity or reliability for any of the statistical analyses. Despite this, based on available evidence and in line with previous systematic reviews, as a review team we established what constitutes a 'poor', 'reasonable, or 'good/excellent' statistical result (Chinapaw et al., 2010; Evenson, Goto and Furberg, 2015; Hidding et al., 2017; Hidding et al., 2018; Lubans et al., 2011; van Poppel et al., 2010), and have provided an overview of this information in Table 3.4.

3.4.7.2 *Feasibility*

There is no standardised way of determining feasibility of most measurement tools. It is argued that the feasibility of certain measurement tools (self/proxy reported) can only be truly identified by qualitative methods to determine if the tool is comprehensive and understandable (Terwee et al., 2018a). However, studies use different approaches and

methods of analysing and reporting the feasibility of measurement tools, including amount of missing and usable data. For this purpose of this review, we took the following approach: for studies evaluating proxy reported measurement, feasibility was only recorded if determined using qualitative methods. For all other measurement tools, feasibility was based on qualitative acceptability or feasibility of the measurement tools where data was available. Additionally, an indication of feasibility was also based on compliance and usable data. We used the values from the Effective Public Health Practice Project (EPHPP) quality assessment tool for withdrawal and drop out (EPHPP, 2009). The EPHPP is primarily used for assessing the quality of quantitative intervention based studies in systematic reviews and rates studies as 'strong', 'moderate', or 'weak' based on the percentage of participants completing the study. We applied the criteria to provide an indication on the feasibility of the measurement tools in our review, to indicate a 'weak', 'moderate', or 'good' level of feasibility based on the percentage of usable data from the measurement tools (as a result of missing data and/or drop out) (Effective Public Health Practice Project, 2009). This was used to provide an indication on feasibility only. When interpreting summary scores for feasibility, more weight was given to qualitative findings; the scores of these studies were based on the qualitative data provided in the original study. Information on the interpretation of the studies can be found in Table 3.4.

In addition to measurement property analysis, I also extracted information relating to socioeconomic status and ethnicity of participants, so that it would be possible to analyse results from these studies to identify how well tools worked across or within different sociodemographic groups.

Table 3.4. Main statistical analyses and interpretation of statistics

Statistical analysis		Relative or Absolute Validity/reliability?	Weak	Moderate	Good
Correlations (<i>r</i>)	Pearson's	Relative	<0.60	0.60-0.79	≥0.80
	Spearman's	Relative	<0.60	0.60-0.79	≥0.80
	Kendall's	Relative	<0.60	0.60-0.79	≥0.80
	Intraclass correlation coefficient	Absolute	<0.60	0.60-0.69	≥0.70
Linear Regression (% variance explained by the measurement tool)		Relative	<60%	60-79%	≥80%
Root mean squared error		Absolute	*	*	*
Bland Altman (mean difference, limits of agreement, bias)		Absolute	*	*	*
Kappa (<i>r</i>)		Absolute	<0.60	0.60-0.69	≥0.70
Area under the receiver operating curve (AUC-ROC)		Relative	<0.70	0.70-0.79	≥0.80
Feasibility (% of usable data)			<60%	60-79%	80-100%

(References: Chinapaw et al., 2010; De Vet et al., 2011; EPHPP, 2009; Evenson, Goto and Furberg, 2015; Forouhi, Brage and Wareham, 2017; Hidding et al., 2017; Hidding et al., 2018; Koo and Li, 2016; Lohr, 2002; Lubans et al., 2011; Schmidt and Steindorf, 2006; Terwee et al., 2007; Twomey and Kroll, 2008; van Poppel et al., 2010)

*Depends on the units of measurement

3.4.8 Combining the results of individual studies for overall interpretation

The systematic review has taken the form of narrative synthesis, including a structured and substantial focus on displaying the evidence by type of measurement property explored: validity (construct, criterion, convergent, face/content), reliability (test-retest, intra-instrument and internal consistency) and feasibility. The study designs of included studies were too heterogeneous to employ meta-analyses; studies included in the review examined different measurement tools, using different comparison measurement tools, using different protocols and designs. Studies that assessed the same measurement tool (e.g. Actigraph GT3X), in comparison with similar reference methods (such as direct observation, but often using different observation protocols), was often limited to 3 studies maximum and used different interpretation choices (e.g. cut points of the accelerometer), and so this was not sufficient for meta-analyses. Table 3.5 provides an overview of the measurement properties that were examined for each tool, to help determine which type of evaluation was conducted on each of the measurement tools. This table does not state anything about the quality of the tools.

In this review, we summarised the results of studies where they aimed to compare a particular measurement property (separated by level of evidence for validity) of a particular measurement tool (e.g. Actigraph GT3X accelerometers). The summaries of this information, including the quality of the tools based on the available evidence, can be found in: Table 3.7 (level 1 validity evidence), Table 3.9 (level 2 validity evidence), Table 3.11 (level 3 validity evidence), Table 3.13 (level 4 validity evidence), and Table 3.17 (reliability and feasibility evidence). The evidence outlined in the tables was interpreted as follows:

- Where the specific measurement property for a specific measurement tool was deemed 'good' in the majority of studies, the summary assessment was deemed to be 'good'.
- Where the specific measurement property for a specific measurement tool was deemed 'moderate' in the majority of studies, the overall assessment was deemed to be 'moderate'.
- Where the specific measurement property for a specific measurement tool was deemed 'weak' in the majority of studies, the overall assessment was deemed to be 'weak'.
- In instances where the specific measurement property of a measurement tool had mixed evidence in the studies, such as studies with outcomes of 'weak' and 'moderate', or 'moderate' and 'good', the overall assessment was deemed to be the most positive of the two outcomes.

For completeness, I have included information about all tools of reasonable quality in the summary tables (Tables 3.7, 3.9, 3.11, 3.13 and 3.17) where there was any available evidence, including where there was only one or two studies that reported results. The information is greyed out to highlight where the evidence is based on ≥ 3 studies to demonstrate where there is a more established evidence base.

3.5 Results

3.5.1 Study selection

A total of 6,088 records were screened for inclusion, of which 146 were included for full text screening and 69 articles were included in the review, describing 75 individual studies (See Figure 3.1). Sixty seven articles were retrieved following the initial searches, and a further two from the updated searches (Dobell et al., 2019; Steenbock et al., 2019). All included articles were in English and, in most cases, were peer reviewed journal articles (n=66). One abstract (Lee et al., 2014) and two Masters theses (Murray, 2009; Shin, 2015) were included due to the articles meeting the inclusion criteria and providing sufficient information. A list of excluded studies with reasons can be found in appendix 3.5.

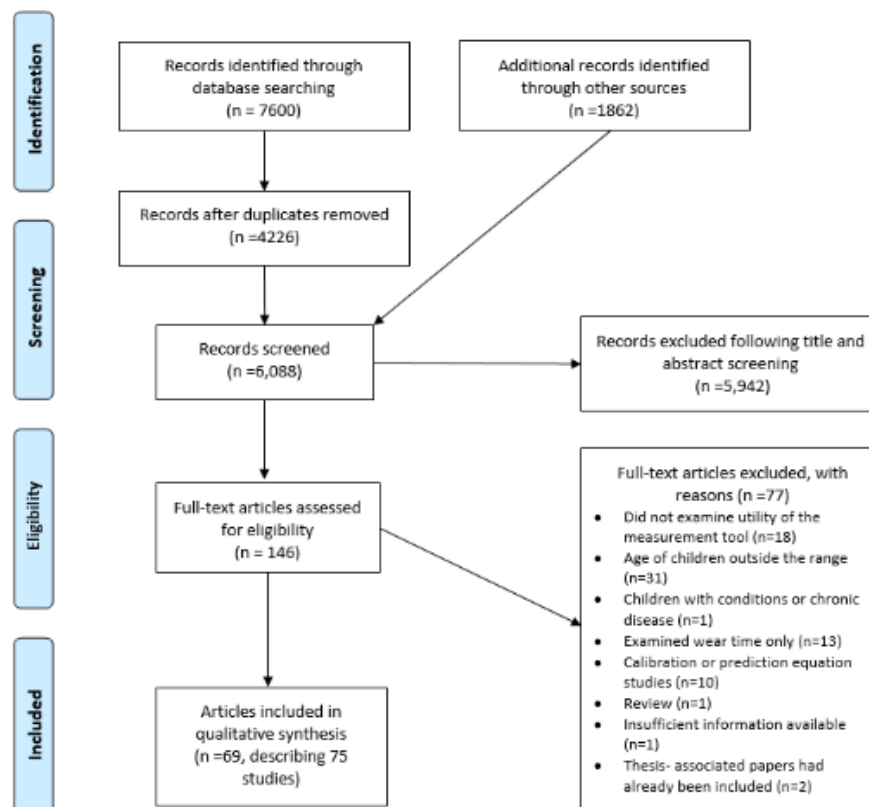


Figure 3.1. PRISMA flow chart of included studies

3.5.2 Description of studies

The majority of the studies were conducted in high income countries: USA (n=24), UK (n=17), Australia (n=10), Belgium (n=3), Canada (n=3), Hong Kong (n=3), New Zealand (n=3), Greece (n=2), Japan (n=2), Netherlands (n=2), Germany (n=1), Korea (n=1), Sweden (n=1), and both the USA and Sweden (n=1). One study was conducted in Mexico, an upper to middle income country (Bacardi-Gascón et al., 2012) and one study was a collaboration between high income and an upper to middle income country: Belgium, Bulgaria, Germany, Greece, Poland and Spain (González-Gil et al., 2014). No studies were conducted in low income countries. Study sample sizes ranged from 4 (Larson, Normand and Hustyi, 2011) to 269 (Sirard et al., 2005) (median n= 34). Based on criteria outlined in the COSMIN measurement property checklist (Mokkink et al., 2018b) only 7 of the included studies had excellent (n= >100), 15 had good (n=50-99), 26 had moderate (n=30-49), and 27 had small (n= <30) sample sizes. All studies reporting child sex (n=69) included both male and female participants. The median age of children included in the studies was 4.5 years. The studies commonly explored the measurement properties of the measurement tools in free-living conditions (n=54), used structured protocols that were reflective of habitual behaviour of pre-school aged children (n=11) or were conducted in laboratory settings (n=10). A large proportion of the studies did not report on any funding received for the research (n=31). Many of the studies that reported funding were supported by more than one funding source.

Studies assessed the measurement tools for PA only (n=27), SB only (n=5), and both PA and SB (n=43). Units of measure varied across studies, and included: activity preferences, activity levels, activity counts, activity energy expenditure, energy expenditure, frequency of

activity, heart rate, metabolic equivalent of task (MET) values, posture allocation, step count, time spent in different intensities of activity and VO_2 .

The majority of the studies examined the measurement properties of one measurement tool only (n=65). Several studies examined the measurement properties of two (n=8) or infrequently three (n=2) measurement tools simultaneously, in comparison with the reference methods. Articles examined the measurement properties of calorimetry (n=2), direct observation (n=4), combined heart rate and accelerometry (n=1), heart rate monitors (n=1), accelerometers (n=44), pedometers (n=13), and proxy report measures (questionnaires or diaries) reported by parent, carer or teacher (n=13).

Validity of the measurement tools was the most frequently examined measurement property; level 1 validity (n=12), level 2 validity (n=36), level 3 validity (n=23) and level 4 validity (n=9). Only two studies examined the face and content validity of the measurement tools. Ten studies described the test-retest reliability of the measurement tools and one study the intra-instrument reliability. Feasibility data was abstracted from forty one studies; thirteen of these studies had a primary aim of determining the feasibility of the measurement tool, the remaining twenty eight studies commented on reasons for drop out or exclusion of data, which was also classified as assessing feasibility. Table 3.5 provides an overview of all measurement tools examined and measurement tools explored for each.

Table 3.5. Overview of all measurement tools examined and measurement properties explored

Measurement tools examined	Validity*					Reliability		Feasibility
	Criterion: Level 1	Convergent: Level 2	Convergent: Level 3	Convergent: Level 4	Face/Content	Test- retest-	Intra- instrument	
Calorimetry								●
Direct observation		●	●					●
Combined heart rate and accelerometer								
Actiheart	●	●		●				
Heart rate monitor								
Polar Vantage XL Monitor		●						●
Accelerometers								
Actical	●	●		●				●
Actigraph (CSA/MTI)	●	●		●				●
Actigraph (GT1M)		●		●				●
Actigraph (GT3X)	●	●		●				●
Actigraph (GT3X+)		●		●				●
Actigraph (wGT3X-BT)	●							●
Actometer		●						
ActivPAL	●	●	●	●				●
Actiwatch (AW16)	●	●		●				●
Actiwatch (MiniMitter)		●						
Actiwatch (Spectrum)		●		●				●
Actiwatch-L		●						●
Best Fit Friend			●					●
Caloriecounter								●
Caltrac								●
Fitbit (Flex)		●	●					

Fitbit (Zip)		●					●	●
GENEActiv	●							
New Lifestyles NL-1000		●						
Tracmor _D	●							●
Triaxial Research Tracker 3 (RT3)	●	●		●		●		
Pedometers								
MVP 4 Walk4Life Digital		●				●		●
Omron Walking Style Pro (HJ-720IT-E2)		●	●					●
Yamax Digi-Walker (SW-200)		●	●					●
Yamax Digi-Walker (SW-700)			●			●		●
Yamasa AM-5 Pedometer		●						
Pedometer (type not specified)		●	●					
Proxy reported measurement tools								
Children's Leisure Activities Study Survey (CLASS)			●			●		
Children's physical activity questionnaire (CPAQ)	●		●			●		
Habitual Activity Estimation Scale (HAES)			●					
Leisure time report			●			●		
Netherland's Physical Activity Questionnaire (NPAQ)			●			●		
Nursery teacher's report			●					
Pre School Physical Activity Questionnaire (PRE-PAQ)			●		●	●		
Questionnaire developed for parents of pre-schoolers in Mexico			●		●	●		
Teacher activity rating			●			●		

Teacher/mother reported habitual PA		●	●	●				
'Toybox' Primary Caregivers Questionnaire						●		
TV Diary			●			●		
7 day activity diary (adapted from CLASS)			●					

*Validity separated per level of evidence depending on the quality of measurement tool used (see Table 3.3 for full explanation).

3.5.3 Risk of bias assessment

Risk of bias was assessed for all included studies using the modified Downs and Black checklist (Downs and Black, 1998). Studies consistently described the main aims of the research, the main outcomes to be measured, the exposures of interest and the main findings. In most cases, the staff, places and facilities were representative of what would usually be the case for the children under study, as testing often took place in nursery settings, at home, or was assessing habitual activity behaviours consisting of children's usual behaviours and routines. However, a proportion of the studies were not reflective of usual activity for children due to them being laboratory based or involving structured protocols.

The main concern regarding potential bias of the studies was related to the lack of reporting of key information in the studies. This included lack of clarity on the representativeness of the sample population compared with the population from which they were recruited (n=63). Also, many studies failed to consistently report reasons for drop out (e.g. non-completion, or missing or incomplete data) (n=34). Many of the studies only reported the number of children included in analysis, but did not include the number who started the study sample, and so it was unclear as to whether drop out was an issue in these studies.

The majority of the studies demonstrated fair to good methodological quality ratings. Only two studies received a low methodological quality score (Fotini, Antonis and Dimitra, 2015; Saris and Binkhorst, 1977) and were removed from the overall summary analysis, however this did not affect the overall results of the review. The full Downs and Black risk of bias assessment for all included studies can be found in appendix 3.6.

An additional risk of bias assessment was conducted on studies reporting proxy reported measurement tools, using the COSMIN risk of bias (Mokkink et al., 2018b; Prinsen et al., 2018). The checklist demonstrated that the majority of tools examining convergent validity were of low quality; five studies were rated doubtful (Dwyer et al., 2011; Mendoza et al., 2013; Noland et al., 1990; Telford et al., 2004; Wen et al., 2010) and seven studies inadequate (Bacardi-Gascón et al., 2012; Chen et al., 2002; Chow and Au, 2009; Corder et al., 2009; Janz, Broffitt and Levy, 2005; Manios, Kafatos and Markakis, 1998; Nishikido, Kashiwazaki and Suzuki, 1982). In the majority of the studies it was clear what the comparator tool measured and the statistical methods used were generally appropriate. Main concerns with the methodological quality of the studies related to the measurement properties of the comparator tools, with inadequate information provided. Proxy reported tools examined for test-retest reliability showed a range of quality, with some studies rated as adequate (Dwyer et al., 2011; Mendoza et al., 2013; Telford et al., 2004), two as doubtful (Corder et al., 2009; Manios, Kafatos and Markakis, 1998) and some inadequate (Bacardi-Gascón et al., 2012; González-Gil et al., 2014; Janz, Broffitt and Levy, 2005). The full COSMIN risk of bias assessment can be found in appendix 3.7. This risk of bias assessment highlighted that the studies on proxy reported measurement tools were generally of low quality. However, no studies were removed from the overall summary analysis based on this additional assessment, due to this evidence being the best available evidence for proxy reported measurement tools. The quality of the evidence is regarded in the interpretation of the studies throughout the review.

3.5.4 Summary of measurement properties of tools

Results presented here are in line with the level of evidence scheme displayed in Table 3.3. The results will therefore be discussed as follows: Level 1 validity evidence where the tool under study is compared with calorimetry; level 2 validity, where the tool under study is compared with a measurement tool with relatively high validity e.g. direct observation; level 3 validity, where the tool under study is compared with a measurement tool with relatively low validity e.g. device based method such as accelerometry; level 4 validity, where two of the same type of measurement tool are compared where neither tool is considered to have known higher validity e.g. two makes of accelerometer; reliability; followed by feasibility.

3.5.4.1 Level 1 Validity

The criterion method of calorimetry (including doubly labelled water) for the outcome of energy expenditure was used in 12 studies. Multiple accelerometers were shown to have reasonable ability to determine energy expenditure, but often based on very limited evidence (one study only). There was however stronger evidence to suggest that the Actigraph (in particular the MTI and GT3X) and Actical were both able to determine energy expenditure and VO₂ max (Janssen et al., 2013a; Pate et al., 2006; Pfeiffer et al., 2006; Reilly et al., 2006; Steenbock et al., 2019), and the ActivPAL to determine energy expenditure (Janssen et al., 2014; Steenbock et al., 2019).

Table 3.6 provides detailed information and table 3.7 provides a summary table of the quality of all included studies that examined level 1 validity evidence of measurement tools.

Table 3.6. Detailed information of studies examining level 1 validity evidence

Study details	Methods			Units of measure	Validity results
	Measurement tool(s) under study <i>Placement, epoch, cut points, wear time, non-wear time and number of valid days</i>	Criterion Method	Study protocol <i>(lab/free living)</i>		
Combined heart rate and accelerometry and accelerometers (n=1)					
Physical activity (PA) and Sedentary Behaviour (SB) (n=1)					
Adolph et al. (2012); USA; n=64; 3-5 years, mean age=4.5 years; 37 male, 27 female 50% White, 27% Black, 22% Hispanic, 2% Asian	Actiheart (MiniMitter) <i>Chest</i> <i>15 sec epoch</i> Actical (Respironics, MiniMitter) <i>Right hip</i> <i>15 sec epoch</i> Triaxial Research Tracker 3 (RT3-Stayhealthy) <i>Right hip</i> <i>1 sec epoch</i>	Whole room Calorimetry	3-hour whole room calorimeter protocol <i>(laboratory based)</i>	Accelerometers: activity counts per minute Room respiration calorimetry: heart rate (HR), energy expenditure (EE), activity energy expenditure (AEE), physical activity rating	Pearson’s Correlation: In increasing order - Actical, Actiheart, RT3 total counts were positively correlated with AEE and physical activity ratings r=0.69-0.82 (p=0.001). Actiheart significantly correlated with

	<i>Cut points: Piecewise linear regressions determined thresholds into sedentary behaviour (SB), light physical activity (LPA), moderate physical activity (MPA), vigorous physical activity (VPA)</i>				AEE $r=0.72$ ($p<0.001$). Overall classification accuracy of accelerometers compared to calorimetry: Actical -71% Actiheart – 72% RT3-73%
Accelerometers (n=10)					
PA and SB (n=3)					
Janssen et al. (2013a); Australia N=40; 4-6 years, mean age 5.3 years; 22 male, 18 female.	Actigraph (GT3X) <i>Right hip 15 sec epoch Cut points: AEE and VO₂: Puyau et al., 2002, Pate et al., 2006; and activity intensity: Puyau et al., 2002, Pate et al., 2006,</i>	Whole room calorimetry	150-minute structured activity laboratory protocol (<i>lab based</i>)	Actigraph- Activity counts converted to AEE, VO ₂ and activity intensity (SB, LPA, moderate to vigorous (MVPA)). Whole room calorimetry- AEE	Dependent t- tests, with Bonferroni correction: VO₂- Actigraph using Pate cut point significantly overestimated VO ₂ during SB and LPA and for total VO ₂ ($p<0.001$) but did

	<i>Evenson et al., 2008, Sirard et al., 2005, Van Cauwenberghe et al., 2011b and Reilly et al., 2003</i>				not show a significant difference between measured and predicted VO₂ during MVPA (p=0.072). Energy expenditure- Actigraph using Puyau cut point significantly underestimated AEE during MVPA and LPA and for total AEE (p<0.0125) but did not show a significant difference for AEE during SB (p=0.5481). Classification accuracy determined using area under the receiver operating curve (ROC-AUC),

					sensitivity, specificity: Both PA intensity and EE: Classification accuracy for Evenson cut point for SB (ROC-AUC = 0.90, sensitivity = 90.7%, specificity= 89.7%), LPA (ROC-AUC = 0.76, sensitivity = 69.9%, specificity= 82.2%) and MVPA (ROC-AUC = 0.76, sensitivity = 60.5%, specificity= 91%); higher accuracy compared to all others expect Pate cut point for MVPA (ROC-AUC = 0.78, sensitivity = 69.6%, specificity= 86.4%).

Janssen et al. (2014); Australia; N=18; 4-6 years, mean age 5.2 years; 9 male, 9 female	ActivPAL <i>Thigh</i> <i>15 sec epoch</i> <i>Cut points:</i> <i>produced</i> <i>during ROC-</i> <i>AUC analysis</i>	Whole room calorimetry	150 minute room calorimeter protocol (<i>laboratory based</i>)	ActivPAL- activity counts to determine MVPA and metabolic equivalent of task (MET) equation Whole room calorimetry – EE	Classification accuracy determined by ROC-AUC, sensitivity, specificity: ActivPAL METs equation overestimated METs during SB (+6.0%) and underestimated METs for LPA (- 15.3%), MVPA (- 32.8%) ad total METs (-13.6%) (all p<0.001). Classification accuracy for activPAL determined MVPA when using both EE and direct observation as criterion: Sensitivity = 94.8%, Specificity = 84.8%, ROC- AUC = 0.90.

Janssen et al. (2015); Australia; N=40, 4-6 years, mean age 5.3 years; 22 male, 18 female	Actical (Respironics, Philips) <i>Right hip 15 and 60 sec epoch Cut points: Pfeiffer et al., 2006, Adolph et al., 2012 and Evenson et al., 2008</i>	Whole room calorimetry	150 minute room calorimeter protocol (<i>laboratory based</i>)	Actical- AEE; activity intensity - SB, LPA, MVPA Whole room calorimetry – EE	Actical for predicting AEE: Dependent t- test: Pfeiffer equation overestimated AEE during SB and underestimated AEE during LPA (0.017 vs 0.005 kcal·kg ⁻¹ ·min ⁻¹ , 0.023 vs 0.030 kcal·kg ⁻¹ ·min ⁻¹ , respectively; both p<0.01). No difference for MVPA (p=0.76) or total AEE (p=0.80). Classification accuracy determined by ROC-AUC, sensitivity and specificity: Broken down by intensity: SB: Adolph (15s): ROC-AUC=0.80, sensitivity = 89.1%,

					specificity= 71.4% Adolph (60s): ROC-AUC=0.82, sensitivity = 82.5%, specificity=81.1% Evenson (15s): ROC-AUC=0.79, sensitivity = 91%, specificity= 68% Classification accuracy was significantly higher for Adolph compared to Evenson (p< 0.05). LPA: Adolph (15s): ROC-AUC=0.68, sensitivity = 51.9%, specificity= 84.3% Adolph (60s): ROC-AUC=0.73, sensitivity = 66.2%, specificity=80.2% Evenson (15s): ROC-AUC=0.65,

					sensitivity = 51%, specificity= 79.1% Adolph showed a significantly higher sensitivity (66.2%), which resulted in significantly higher classification accuracy compared to Evenson ($p < 0.05$). MVPA: Adolph (15s): ROC-AUC=0.82, sensitivity = 71%, specificity= 93.6% Adolph (60s): ROC-AUC=0.85 sensitivity = 74.3%, specificity=95.3% Evenson (15s): ROC-AUC=0.75, sensitivity = 54.9%, specificity= 96%

					Pfeiffer (15s): ROC-AUC=0.70, sensitivity = 43.9%, specificity= 96.5% Classification accuracy was significantly higher for Adolph compared to all others ($p < 0.05$). Actical compared with direct observation and EE: Kappa: κ values were slightly increased compared to using direct observation only. Adolph ($\kappa = 0.72$, 95% CI = 0.70– 0.74), Adolph(15s) ($\kappa = 0.65$, 95% CI = 0.64–0.65) and Evenson ($\kappa = 0.61$ 95% CI = 0.60– 0.62) exhibited substantial

					agreement predicting PA and SB determined by direct observation and calorimetry. Classification accuracy determined by ROC-AUC, sensitivity and specificity : ROC-AUC values when using direct observation combined with EE were slightly higher but similar to using direct observation only. Adolph was significantly better than all others ($p < 0.05$) when classifying LPA and MVPA. For SB, classification accuracy was good for the

					Adolph and Evenson (ROC-AUC = 0.82–0.85) with Adolph performing significantly better than Evenson. However, when comparing Evenson and Adolph ^(15s) this difference disappeared.
PA (n=7)					
Delisle Nyström et al. (2017); Sweden; N=40; 5.2- 5.7 years, mean age 5.5 years; 22 male, 18 female	Actigraph (wGT3X-BT) <i>Non-dominant wrist</i> <i>1 sec epoch</i> <i>No cut points reported for PA</i> <i>Wear time: When the awake wearing time (non-sleeping data) was ≥600 minutes</i>	Doubly labelled water (DLW)	7 days activity (<i>free living</i>)	Actigraph- Counts per minute to determine AEE DLW- total energy expenditure (TEE) and AEE	Bland Altman: Mean difference between the methods was 2.4kj/24h with limits of agreement of 621.6kj/24h (mean vector magnitude (VM) total) and mean difference - 2.4kj/24h with limits of agreement of 584.8kj/24h (mean VM waking).

	<i>Non-wear time: Determined using raw acceleration data. Standard deviation of each axis calculated over a 30-min period. If standard deviation of acceleration of any two axes less than 0.002 g for the same window= non-wear time. Non-wear time confirmed by diary completed by parents.</i> <i>Number of days: Minimum of 3 days of valid data.</i>				
Pate et al. (2006); USA;	Actigraph (MTI 7164) Right hip	Indirect calorimetry- Cosmed	20 minutes of unstructured activities in pre-school setting (<i>usual activity- children were wearing metabolic</i>	Actigraph- activity counts	Spearman's correlation:

N=29; 3-5 years, mean age 4.4 years; 13 male, 16 female 55.2% African American, 44.8% White	<i>15 sec epoch</i> <i>Cut points:</i> <i>Pate et al.,</i> <i>2006</i>	portable metabolic system (Model K4b2)	<i>system at time of activity and were asked to select new activity every 4-6 minutes)</i>	Cosmed metabolic system- VO₂ estimation	Correlation between the measurement tools was $r=0.66$ ($p<0.001$). Pearson's correlation: Correlation between VO₂ and accelerometer counts across the activities was $r=0.82$ Intraclass correlation coefficient: Measured and predicted VO₂ was $r=0.57$. % agreement, kappa, modified kappa, sensitivity, specificity: MVPA: % agreement was 0.69, kappa 0.36, modified kappa 0.38. Sensitivity

					96.6% and specificity 86.2%. VPA: % agreement was 0.81, kappa 0.13 and modified kappa 0.62. Sensitivity 65.5% and specificity 95.4%.
Reilly et al. (2006) ; Scotland, UK; N=85; 3-6 years, mean age 4.6 years; 51 male, 34 female	Actigraph (CSA/MTI) <i>Right hip</i> <i>1 min epoch</i> <i>Cut points:</i> <i>Ekelund et al., 2001 and Puyau et al., 2002</i> <i>Wear time:</i> <i>>360 minutes</i> <i>Non wear time:</i> <i>Parents to record when accelerometer removed</i> <i>Valid number of days: 3 days (2 weekdays)</i>	DLW	3 days of habitual activity (<i>free living</i>)	Actigraph-activity counts to determine TEE DLW- TEE	Bland Altman: Ekelund equation- Mean paired difference between predicted TEE (Actigraph data) and DLW was +0.3 MJ/d. Difference was not statistically significant (p=0.10), limits of agreement were +4.3 and -3.7 MJ/d. Errors were significantly negatively correlated with TEE (r=-0.65, p<0.001). Errors in estimation

	<i>and 1 weekend day)</i>				larger in boys compared with girls. Puyau equation-Mean paired difference between TEE (Actigraph data) and DLW was - 0.3 MJ/d. Difference was not statistically significant ($p < 0.10$), limits of agreement were +3.2 to -3.8 MJ/d. Errors were significantly negatively correlated with TEE $r = -0.89$ ($p < 0.001$). No gender differences in accuracy.
Steenbock et al. (2019); Germany; n=41; 3yrs – 6.3 years, mean age 4.8 years; 22 male, 19 female 100% Caucasian	GENEActiv <i>Right hip, left wrist, right wrist</i> Actigraph (GT3X)	Indirect calorimetry: MetaMax3b	75 minute protocol of both pre-determined and free play activities (<i>usual activity</i>)	Accelerometers – EE and METs Indirect calorimetry - EE	Root mean squared error (RMSE): For predicting EE: GENEActiv devices obtained mean root mean

	<i>Left and right hip</i> ActivPAL <i>Thigh</i> <i>Cut points: Regressions used to predict EE and METs</i>				squared error (RMSE) of 2.56 kJ/min (left and right wrist) and 2.73kJ/min (right hip); GT3X on the left and right hip obtained mean RMSE of 2.60 and 2.74 kJ/min, respectively; thigh worn activPAL provided a mean RMSE of 2.76kJ/min.
Pfeiffer et al. (2006); USA; n=18; 3.4-5.7 years, mean age 4.4 years ; 7 male, 11 female 89% African American	Actical (MiniMitter) <i>Right hip</i> 15 sec epoch <i>Cut points: Pfeiffer et al., 2006</i>	Indirect calorimetry- Cosmed portable metabolic system (Model K4b2)	20 minutes of unstructured activities at the pre-school, but child instructed to choose a different activity every 6 minutes (<i>usual activity</i>)	Actical: Activity counts to determine VPA and MVPA, and estimated VO₂ Cosmed: VO₂	Intraclass correlation coefficient (ICC): ICC between VO₂ estimated from Actical counts and by Cosmed = 0.59 using log transformed data. Spearman's correlation:

					Spearman correlation coefficient between the VO₂ estimated from Actical and Cosmed $r=0.80$ ($p<0.001$). % agreement and Cohen's kappa: Percent agreement, kappa, modified kappa for VPA were: 0.85, 0.26 and 0.71 respectively. Percent agreement, kappa, modified kappa for MVPA were: 0.73, 0.40, 0.46 respectively.
Lopez-Alarcon et al. (2004); USA; N=29; 4-6 years, mean age 4.9 years; 17 male, 12 female 58% White	Actiwatch (AW16) <i>Right ankle</i> <i>1 min epoch</i>	DLW	8 days of habitual activity- first day not included in the analyses (<i>free living</i>)	Actiwatch: Activity counts DLW: TEE	Pearson's correlation: Activity counts and TEE were not significantly associated $r=0.27$ ($p=0.15$).

42% African American 6 of the children recruited from Head Start Center					Multiple linear regression analyses: Activity counts were not significantly related to TEE in any regression models ($p=0.31$).
Sijtsma et al. (2013); Netherlands; N=30; 3.1-4.1 years, mean age 3.5 years; 12 male, 18 female	Direct life triaxial accelerometer (TracmorD)- <i>Middle of lower back</i> <i>1 min epoch</i> <i>Wear time:</i> <i>Parents reported sleep and wake time, and times of wearing accelerometer.</i> <i>Day invalid when time during which accelerometers worn and time spend sleeping did not add up</i>	DLW	3-5 days of habitual activity, during waking hours apart from water based activities (<i>free living</i>)	TracmorD Triaxial Accelerometer- Activity counts DLW- TEE, AEE, physical activity level (PAL)	Pearson's correlations: Activity counts per minute and per day were significantly correlated with PAL $r = 0.61$ ($p=0.004$) and $r = 0.46$ ($p=0.042$) respectively. Activity counts per minute correlated with activity EE $r = 0.56$ ($p=0.011$), but no correlation between activity counts per day and activity EE

	<i>to at least 19 hours.</i> <i>Valid number of days: At least 3 days of wear time</i>				was found $r = 0.38$ ($p=0.098$). Activity counts per minute and per day were not correlated with TEE $r = 0.34$ ($p=0.094$) and $r = 0.21$ ($p=0.326$) respectively. Linear regression: Activity counts per minute explained 31% of the variance in AEE ($p=0.011$).
Proxy reported measurement tools (n=1)					
PA (n=1)					
Corder et al. (2009); UK; N=27; 4-5 years, mean age 4.9 years; 17 male, 10 female	Children's physical activity questionnaire (CPAQ)	DLW	Questionnaire for past 7 days, remaining protocol (accelerometer and DLW) for 11 days assessing habitual activity (<i>free living</i>)	CPAQ- Physical activity energy expenditure (PAEE) and time spent in MVPA Actigraph- Time spent in MVPA DLW- physical activity energy expenditure (PAEE)	Spearman's correlations: PAEE assessed by CPAQ was not significantly correlated with PAEE from DLW $r=0.22$ ($p=0.28$) Bland Altman: Mean bias (PAEE) = -14.4 ± 52.4

					(95% CI -35.2, 6.3)

Abbreviations in this table: SB= sedentary behaviour; LPA= light physical activity; MVPA =moderate to vigorous physical activity; HR= heart rate; EE= energy expenditure; AEE =activity energy expenditure; TEE= total energy expenditure; VM= vector magnitude; PAL =physical activity level; MET = metabolic equivalent of task; ROC-AUC= area under the receiver operating curve; RMSE= Root mean squared error; Tracmor_D= Direct life triaxial accelerometer; RT3= Triaxial Research Tracker 3; CPAQ= Children's physical activity questionnaire; CI= Confidence interval

Table 3.7. Summary table of level 1 validity evidence of the measurement tools

Measurement tools under study	Outcome measures		References
	Energy expenditure	VO ₂	
Combined heart rate and accelerometer			
Actiheart			Adolph et al., 2012
Accelerometers			
Actigraph (MTI)	 1-2	 3	Pate et al., 2006; Reilly et al., 2006
Actigraph (GT3X)	 2-3	 3	Janssen et al., 2013a; Steenbock et al., 2019
Actical	 4-6	 5	Adolph et al., 2012; Janssen et al., 2015; Pfeiffer et al., 2006
Actigraph (wGT3X-BT)			Delisle Nyström et al., 2017
ActivPAL			Janssen et al., 2014; Steenbock et al., 2019
GENEActiv			Steenbock et al., 2019
Triaxial Research Tracker 3 (RT3)			Adolph et al., 2012
Actiwatch (AW16)			Lopez-Alarcon et al., 2004
Tracmor _D			Sijtsma et al., 2013
Proxy reported measurement tools			

Children's physical activity questionnaire (CPAQ)			Corder et al., 2009

Cut points: ¹Ekelund et al., 2001; ²Puyau, Vohra and Butte, 2002; ³Pate et al., 2006; ⁴Evenson et al., 2008; ⁵Pfeiffer et al., 2006; ⁶Adolph et al., 2012

***Methodology used to assess the ability of the tool is detailed in the methods above and is indicated in the summary table as:**

Good = ● Moderate=● Weak=●

Key for colour of boxes:  = evidence from ≥3 studies

 = evidence from <3 studies

This table shows a summary of the results of studies where they aimed to compare a particular measurement tool (e.g. Actigraph GT3X accelerometer) against calorimetry (including doubly labelled water). The summary ratings were based on the quality of the tools for this specific measurement property. Where the measurement tool was deemed 'good' in the majority of the studies, the summary assessment was deemed 'good'. Where the measurement tool was deemed 'moderate' in the majority of the studies, the summary assessment was deemed 'moderate'. Where the measurement tool was deemed 'weak' in the majority of the studies, the summary assessment was deemed 'weak'. In instances where the measurement tool had mixed evidence in the studies, such as studies with outcomes of 'weak' and 'moderate', or 'moderate' and 'good', the overall assessment was deemed to be the most positive of the two outcomes. All tools of reasonable quality where any evidence was available are included in this table, including where only one or two studies reported that result.

3.5.4.2 *Level 2 Validity*

The most commonly used comparison methods for convergent validity of device based measurement tools was direct observation. Studies demonstrated that a number of device-based measurement tools were reasonably accurate at determining different PA and SB outcomes (Ettienne et al., 2016; Fairweather et al., 1999; Finn and Specker, 2000; Hands and Larkin, 2006; Kelly et al., 2004; McKee et al., 2005; Reilly et al., 2003). The Actigraph accelerometer was the most frequently evaluated tool. Overall, these studies showed the Actigraph devices (in particular GT3X versions) had a good ability to determine SB, vigorous PA (VPA) and MVPA (Alhassan et al., 2017; Dobell et al., 2019; Hislop et al., 2016; Hislop et al., 2012a; Hislop et al., 2012b; Janssen et al., 2013a; Kahan, Nicaise and Reuben, 2013). Fewer studies evaluated the Actical accelerometer but these showed similar results (Adolph et al., 2012; Janssen et al., 2015). The ActivPAL accelerometer was shown to be suitable at assessing SB, MVPA and posture allocation (Alghaeed et al., 2013; Davies et al., 2012; Janssen et al., 2013b; Janssen et al., 2014); a unique quality that is not identified by other measurement tools. However, there is space for this to be developed further, as the accuracy of this measurement tool for identifying posture allocation is lower than in other population samples due to the amount of time that pre-school aged children will spend in 'other' postures, such as kneeling and crawling (Janssen and Cliff, 2015). Fitbits also show some promising results for the measurement of SB, MVPA, total PA (TPA), and step counts (Byun et al., 2018b; Sharp et al., 2017); however, these conclusions are based on a very limited number of studies.

For pedometers, when compared against direct observation, study results were mixed but there is limited evidence to suggest that the Yamax Digiwalker SW-200 is able to determine step counts with reasonable accuracy (Louie and Chan, 2003; McKee et al., 2005).

Table 3.8 provides detailed information and table 3.9 provides a summary table of the quality of included studies that examined level 2 validity evidence of measurement tools.

Table 3.8. Detailed information of studies examining level 2 validity studies

Study details	Methods			Units of measure	Validity results
	Measurement tool(s) under study <i>Placement, epoch, cut points, wear time, non-wear time and number of valid days</i>	Comparison Tool	Study protocol <i>(lab/free living)</i>		
Direct observation (newly devised) (n=1)					
Physical activity (PA) and sedentary behaviour (SB) (n=1)					
Klesges et al. (1985); USA; n=30; 3.5-6.5 years, mean age 4.8 years; 10 male, 20 female	Fargo Activity Timesampling Survey (FATs)– Continuous observation.	Fargo Activity Timesampling Survey (FATs)– Time sampling	Child observed for 1 hour during free play in nursery setting <i>(free living)</i>	Behaviour (sleeping/lying down, crawling, climbing, standing, walking, running) Intensity (minimal, moderate, extreme)	Pearson’s correlations: Correlations between the 2 observation sampling methods (p<0.001). Total behaviour r=0.79, total intensity r=0.77 and total behaviour by intensity = r=0.90. Correlations between the 2 sampling methods for all behaviours and intensities (p<0.001). T-tests: Continuous method yielded higher levels of crawling (t (28) = 2.87, p<0.001), walking (t (28) = 7.86, p<0.001), running (t (28) = 4.23, p<.001), moderate

					levels of activity ($t(28) = 4.16, p < 0.001$), and extreme levels of activity ($t(28) = 3.26, p < 0.01$) than the time sampling method. Continuous method yielded lower levels of sitting ($t(28) = 4.35, p < 0.001$) and minimal levels of activity ($t(28) = 4.64, p < 0.001$) compared with the time sampling
Combined heart rate and accelerometry and accelerometers (n=1)					
PA and SB (n=1)					
Adolph et al. (2012); USA; n=64; 3-5 years, mean age=4.5 years; 37 male, 27 female 50% White, 27% Black, 22% Hispanic, 2% Asian	Actiheart (MiniMitter) <i>Chest</i> <i>15 sec epoch</i> Actical (Respironics, MiniMitter) <i>Right hip</i> <i>15 sec epoch</i> Triaxial Research Tracker 3 (RT3-Stayhealthy) <i>Right hip</i> <i>1 sec epoch</i> <i>Cut points: Piecewise linear regressions determined</i>	Direct observation (Children's activity rating scale (CARS)).	3 hour whole room calorimeter protocol (<i>laboratory based</i>)	Accelerometers: activity counts per minute Direct observation: Activity levels	Pearson's Correlation: In increasing order - Actical, Actiheart, RT3 total counts were positively correlated with activity energy expenditure (AEE) and physical activity ratings $r=0.69-0.82$ ($p=0.001$). Similar correlations observed between accelerometer counts and CARS and heart rate (HR) $r=0.64-0.80$ ($p=0.001$). True positive predictive rates for moderate to

	<i>thresholds into Sedentary behaviour (SB), light physical activity (LPA), moderate physical activity (MPA), vigorous physical activity (VPA)</i>				vigorous physical activity (MVPA): Actical- 69% Actiheart – 82% RT3- 79%
Heart Rate Monitors (n=1)					
PA and SB (n=1)					
Bar-Or et al. (1996); Canada; n=27; 3-5 years, mean age=4.3 years; 18 male, 9 female	Polar Vantage XL Monitor <i>Chest and wrist Records every 5s</i>	Electrodiagram (ECG)	19 minute structured activity protocol (<i>laboratory setting</i>)	HRM- Heart beats per minute Electrodiagram- ECG recordings	Pearson's correlation: Correlations between ECG and heart rate monitor values ranged from r=0.92 to r=0.99 Values very similar for resting (ECG = 97.3 ± 7.5 , XL = 97.9 ± 7.2 for supine and ECG = 111.1 ± 16.5 , XL = 110.4 ± 16.3 for upright, both r=0.99). However, the XL monitor tended to underestimate heart rate during cycling (ECG = 142.7 ± 11.0 , XL = 140.2 ± 11.5 , r=0.93), and overestimate heart rate during recovery (ECG = 112.4 ± 12.8 , XL = 118.0 ± 12.3 , r=0.92).
Accelerometers (n=26)					
PA and SB (n=19)					

Sirard et al. (2005); USA; N=269; 3-5 years, mean age not reported; 125 male, 144 female 3yr; 37.7% white, 4yr; 25.6% white, 5yr; 38.7% white	Actigraph (MTI) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: Sirard et al., 2005</i>	Direct observation (CARS)	Children wore Actigraph for the entire time they were at school for up to 10 consecutive weekdays, direct observation for 1-3 hours of activity in preschool (<i>free living</i>)	Actigraph – Activity counts to determine SB, total physical activity (TPA), LPA, MPA, VPA and MVPA Direct observation – Activity levels (SB, LPA, MPA, VPA, MVPA.)	Pearson's correlation: Significant ($p<0.001$) correlations between Actigraph and direct observation: TPA $r=0.58$, SB $r=0.70$, LPA $r=0.59$, MPA $r=0.50$, MVPA $r=0.46$, VPA $r=0.61$.
Hands et al. (2006); Australia; N=23; 5-6 years, mean age 5.5 years; 12 male, 11 female	Actigraph (MTI AM7164) <i>Waist</i> <i>10 sec epoch</i>	Direct observation (CARS)	Children monitored for 30 minutes over 5 days during free playing in their school setting (<i>free living</i>)	Actigraph- Activity counts Direct observation- Activity levels	Pearson's correlations: Accelerometer and direct observation $r=0.77$ ($p<0.01$). Accelerometer shared 59% of variance with direct observation. One way ANOVA: When the children were grouped into low, moderate, and high activity levels using observation, there was statistically significant differences between groups for the accelerometer. However, data from a pedometer data was better able to separate the groups than from the accelerometer.

Dobell et al. (2019); UK; N=62; 3-4 years, mean age 3.5 years; 36 males, 30 females (sex of participants prior to exclusion of 4 data points)	Actigraph (GT3X) <i>Right hip and non-dominant wrist 5, 10, 15 and 30 sec epochs</i> <i>Cut points: produced during ROC-AUC analysis.</i>	Direct observation (Observational system recording physical activity in children – preschool version (OSRAC-P))	60 minutes of activity in pre-school setting (<i>free living</i>)	Actigraph- Raw acceleration and activity counts to determine SB, LPA, MVPA Direct observation – Activity levels	Spearman's correlation: Correlations between accelerometer raw counts and observations were weak to moderate (ranging from $r=0.34$ to $r=0.55$, all $p<0.05$). Area under the receiver operating characteristic: Agreement for accelerometer counts and observation levels for both SB and MVPA (ROC-AUC ≥ 0.7). Agreement between observation and accelerometer for LPA (ROC-AUC = 0.53 to 0.56). Hip and wrist placement resulted in similar levels of agreement with the wrist performing marginally better for SB and LPA, and the hip performing better for MVPA.
Hislop et al. (2016); Scotland, UK; N=32; 3-5 years, mean age 4.2 years; 21 male, 11 female	Actigraph (GT3X+) <i>Non-dominant wrist and hip 1 sec epoch</i>	Direct observation (CARS)	Children wore accelerometers whilst being video recorded for 60 minute activity in	Actigraph- activity counts to determine minutes in SB, LPA, MVPA and TPA	Bland Altman: Agreement between CARS and actigraph (mean minutes (and standard deviation)):

	<i>Cut points: Johansson et al., 2015 for wrist data and Evenson et al., 2008 for hip data</i>		pre-school (<i>free living</i>)	Direct observation- Activity levels	SB: CARS 6.2(5.8); Actigraph wrist 7.2 (3.9) and Actigraph hip 12.2 (7.2). LPA: CARS 27.3(10.6); Actigraph wrist 16.9 (6.6) and Actigraph hip 17.2(6.6). MVPA: CARS 2.2 (2.2); Actigraph wrist 11.5 (6.3) and Actigraph hip 6.1 (4.5) TPA: CARS 29.5(10.4); Actigraph wrist 28.4(9.8) and Actigraph hip 23.3 (9.8). Mean difference and 95% levels of agreement (LOA): Standing and LPA: Actigraph wrist (3.1, -14.3, 19.9) Actigraph hip (2.8, -10.3, 16.4) MVPA: Actigraph wrist (-9.3, -20.0, 1.5)

					Actigraph hip (-3.8, -10.7, 3.1) TPA: Actigraph wrist (1.1, -9.9, 12.1) Actigraph hip (6.3, -8.8, 21.4) Large differences in the count data from the hip and wrist placements.
Kahan et al (2013) – Study 1: USA; n=57; 4-5 years, 4.7 years, 25 male, 32 female 70.2% White, 12/3% Asian/Pacific Islander, 10.5% Black/African American, 7% Hispanic/Latino.	Actigraph (GT3X) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: Evenson et al., 2008, Pate et al., 2006, Sirard et al., 2005 and Van Cauwenberghe et al., 2011b</i>	Direct observation (OSRAC-P)	At least 15 minutes of activity in pre-school (<i>free living</i>)	Actigraph- % time spent in SB and MVPA determined from activity counts Direct observation- Activity intensity levels	Z-test: SB- Sirard followed by Van Cauwenberghe cut points most closely matched OSRAC-P estimates; both overestimated SB by less than 9%. MVPA- Sirard followed by Van Cauwenberghe and Evenson cut points most closely matched OSRAC-P estimates. Sirard cut points underestimated and Van Cauwenberghe and Evenson cut points overestimated OSRAC-P estimates by less than 5% and by about 10% respectively.

					Bland Altman: Biases between OSRAC-P and Actigraph Sirard cut points were $-8.4\% \pm 18.3\%$ (SB) and $5.7\% \pm 9.7\%$ (MVPA). Limits of agreement were -45.0% to 28.1% (SB) and -13.7% to 25.1% (MVPA). Comparatively less was observed when SB means exceeded about 70%. For MVPA, the greatest clustering around zero difference was observed when mean MVPA was very low.
Kahan et al (2013)- Study 2 : USA; n=12; 4-5 years, mean age not reported; 3 male, 9 female	Actigraph (GT3X) <i>Right hip</i> <i>5 sec epoch</i> <i>Cut points: Evenson et al., 2008, Pate et al., 2006, Sirard et al., 2005 and Van Cauwenberghe et al., 2011b</i>	Direct observation (OSRAC-P)	At least 15 minutes of activity in pre-school (<i>free living</i>)	Actigraph- % time spent in SB and MVPA determined from activity counts Direct observation- Activity intensity levels	Kappa coefficients: Sirard ($\kappa = 0.48$) and Van Cauwenberghe ($\kappa = 0.48$) cut points showed lowest kappa values, and Evenson ($\kappa = 0.58$) showed highest. Sensitivity: SB: Evenson – 75.2%; Pate- 50%; Sirard- 81.4%; Van Cauwenberghe- 80% MVPA: Evenson – 54.8%; Pate-62.7%; Sirard-35.8%; Van Cauwenberghe- 53.9%

					Across cut points, sensitivity was higher for SB than for MVPA, for all except Pate cut points. Specificity: SB: Evenson – 66.4%; Pate- 85.7%; Sirard- 50.7%; Van Cauwenberghe- 52.4% MVPA: Evenson –78.9%; Pate-72.2%; Sirard-87.6%; Van Cauwenberghe- 79.1% Across cut points, specificity was higher for MVPA than for SB, for all except Pate cutpoints. Positive Predictive Value: SB: Evenson – 46.1%; Pate- 74.2%; Sirard- 57.5%; Van Cauwenberghe- 58% MVPA: Evenson –42.1%; Pate-62.8% Sirard-44.8%; Van Cauwenberghe- 40.3% Positive predictive value was higher for SB than for MVPA. Negative Predictive Value:

					SB: Evenson – 87.6%; Pate- 67.6%; Sirard- 76.9%; Van Cauwenberghe- 76.1% MVPA: Evenson –87%; Pate-88%; Sirard-84%; Van Cauwenberghe- 86.7% Negative predictive value was higher for MVPA than SB, except for Evenson. Accuracy: SB: Evenson – 68.9%; Pate- 69.6%; Sirard- 64.5%; Van Cauwenberghe- 64.8% MVPA: Evenson –73.9%; Pate-73%; Sirard-77.5%; Van Cauwenberghe- 73.8% Accuracy was higher for MVPA than SB across all cut points. Sirard and Pate cut points were most accurate for MVPA and SB, respectively.
Janssen et al. (2013a) ; Australia N=40; 4-6 years, mean age 5.3 years; 22 male, 18 female.	Actigraph (GT3X) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: Activity energy expenditure and VO₂: Puyau et al., 2002, Pate et al.,</i>	Direct observation (CARS)	150 minute structured activity laboratory protocol (<i>lab based</i>)	Actigraph- Activity counts converted to activity energy expenditure, VO ₂ and activity intensity (SB, LPA, MVPA).	Classification accuracy determined using ROC-AUC, sensitivity, specificity: PA intensity-

	2006; and activity intensity: <i>Puyau et al., 2002, Pate et al., 2006, Evenson et al., 2008, Sirard et al., 2005, Van Cauwenberghe et al., 2011b and Reilly et al., 2003</i>			Direct observation- Activity intensity: SB, LPA, MVPA.	Using direct observation alone as the criterion measure: SB: Classification accuracy significantly higher for Evenson cut point (ROC-AUC = 0.80, sensitivity = 86.7%, specificity= 72.9%) compared to all other cut points (p<0.05). LPA: Classification accuracy was significantly higher for Evenson cut point (ROC-AUC = 0.65, sensitivity = 54.8%, specificity= 74.8%) compared to all other cut points (P<0.05). MVPA: using the Pate cut point (ROC-AUC = 0.72, sensitivity = 54.2%, specificity= 88.9%) cut-point resulted in classification accuracy which was significantly higher compared to all other cut points (p<0.05). Both PA intensity and EE: Classification accuracy for Evenson cut point for SB

					(ROC-AUC = 0.90, sensitivity = 90.7%, specificity= 89.7%), LPA (ROC-AUC = 0.76, sensitivity = 69.9%, specificity= 82.2%) and MVPA (ROC-AUC = 0.76, sensitivity = 60.5%, specificity= 91%); higher accuracy compared to all others expect Pate cut point for MVPA (ROC-AUC = 0.78, sensitivity = 69.6%, specificity= 86.4%).
Hislop et al.(2012a); Scotland, UK; N=31; 3-5 years, mean age 4.4 years; 15 male, 16 female	Actigraph (GT1M) <i>Waist</i> <i>1 sec epoch</i> <i>Cut points: SB, LPA, MVPA -Puyau et al.,2002, Van Cauwenberghe et al., 2011, Sirard et al., 2005; MVPA only- Pate et al., 2006; SB only – Reilly et al., 2003, Evenson et al., 2008</i>	Direct observation (CARS)	60 minutes activity in pre-school setting (<i>free living</i>)	Actigraph- SB, LPA, MVPA Direct observation- Activity levels	Friedman's repeated measures ANOVA, followed by Wilcoxon paired t-tests: Significant difference between the mean time in SB, LPA and MVPA estimated by the different cut points and the CARS (p=0.00). Actigraph compared with direct observation categories, using the different cut points: Puyau → No significant difference between SB (p=0.8); LPA (p=0.06), MVPA (p=0.06) and CARS.

					Sirard→ No significant difference for MVPA and CARS ($p=0.2$). Significant difference for SB and LPA and CARS ($p<0.01$). Van Cauwenberghe→Significant difference for SB, LPA and MVPA, with CARS ($p<0.01$). Pate →Significant difference for MVPA and CARS ($p<.01$). Reilly and Evenson→Significant difference for SB and CARS ($p<0.01$). Bland Altman: SB- There was a bias towards a greater difference in estimation of time spent in SB between accelerometer and CARS when using Sirard, Van Cauwenberghe and Reilly cut points, lower mean difference in time in SB for the Evenson cut points, Puyau SB cut point were close to 0, but had wide limits of agreement (13.2 to 14.0 mins).

					Sirard (-7.2; -20.2,5.7) Puyau (0.4; -13.25, 13.99) Van Cauwenberghe (-8.2; -22.7,6.2) Evenson (13.2; -2.2,28.6) Reilly (-3.6; -17.6,10.4) LPA-All cut points had lower estimation of time spent in LPA in relation to CARS. Sirard (6.3; -6.0,18.6) Puyau (-2.2; -15.0,10.5) Van Cauwenberghe (11.4; 1.4,24.2) MVPA- Positive bias, with underestimation in MVPA when compared with CARS when using Sirard and Puyau cut points, overestimation when using Van Cauwenberghe and Pate cut points. Sirard (0.8; -6.2,7.8) Puyau (1.7; -8.0, 11.5) Van Cauwenberghe (-3.3; -11.8,5.3) Pate (-8.7; -19.9,2.5)

Alhassan et al. (2017); USA; N=33; mean age 4.4 years; 21 male, 12 female	Actiwatch (Spectrum) <i>Non dominant wrist 15 sec epoch Cut points: Ekblom et al., 2012</i> Actigraph (GT3X) <i>Waist, centre of lower back 15 sec epoch Cut points: Pate et al., 2006 and Sirard et al., 2006</i>	Direct observation (OSRAC-P)	30 minutes of daily activity in pre-school (<i>free living</i>)	Accelerometers- activity counts to determine SB, LPA, MPA, VPA, MVPA Direct observation- Activity intensity categories	Spearman's correlation: Association between direct observation intensity categories and the accelerometers $r=0.47$. Wilcoxon rank sum: Actiwatch and Actigraph able to differentiate between median values for adjacent direct observation intensity levels. All comparisons between accelerometers and direct observation intensity categories significant ($p=0.03$). Bland Altman: Show that both Actiwatch and Actigraph similarly underestimate mins of MVPA compared with direct observation. % Agreement between devices and direct observation: SB: Actiwatch = 65.4%; Actigraph = 61.5-62.7% LPA: Actiwatch = 53.8%; Actigraph = 60.7-65%

					MPA: Actiwatch = 82.8%; Actigraph= 80.3-86.3% VPA: Actiwatch = 86.3%; Actigraph = 86.1-86.5% MVPA: Actiwatch = 77.5%; Actigraph = 75.3-76.9%
Kelly et al. (2004); UK; n=78; 3-4 year olds, mean age 3.5 years; 30 male, 48 female	Actiwatch (AW16) <i>Right hip</i> <i>1 min epoch</i> Actigraph (CSA/MTI) <i>Right hip</i> <i>1 min epoch</i>	Direct observation (Children's physical activity form (CPAF))	39-45 minutes structured play class <i>(usual activity)</i>	Accelerometers- Activity counts per minute Direct observation- Activity levels	Spearman's correlations: Actiwatch not significantly correlated with direct observation $r=0.16$. Actigraph significantly correlated with direct observation $r=0.72$ ($p<0.001$). Pearsons's correlation: Within child significant positive correlations between Actigraph and direct observation $r=0.52$ ($p<0.01$; range -0.22 to $+0.87$), and for Actiwatch and direct observation $r=0.55$ ($p<0.01$; range $-$ 0.41 to $+0.88$).
Finn & Specker (2000); USA; n=40; 3-4 years, mean age not reported; 16 male, 24 female 95% Caucasian	Actiwatch (MinMitter) <i>Waist lower back</i> <i>1 min epoch</i>	Direct observation (CARS)	5-6 hours activity in child care setting <i>(free living)</i>	Actiwatch- Activity counts per minute Direct Observation- Activity level	Correlations: 3 min CARS scores with 3 min activity counts had a median correlation coefficient of $r=0.74$ (range 0.03 to 0.92).

					Mixed model repeated measures: Positive relationship between within-child correlation coefficients and the mean 3 min CARS $r=0.37$ ($p=0.02$) or mean activity counts $r=0.31$ ($p=0.05$).
Djafarian et al. (2013); Scotland, UK; N=42; 3-5 years, mean age 4.06 years; 22 male, 20 female 100% Caucasian	Actiwatch-L <i>Non dominant wrist</i> <i>1 min epoch</i>	Direct observation (CARS)	2 hours usual structured and unstructured activity in preschool setting (<i>usual activity, not free living</i>)	Actiwatch-L: Activity counts per minute Direct observation: Activity level	Correlations: Correlation between Actiwatch-L counts and CARS scores for varying intensities, range from $r=0.41$ to $r=0.63$ ($p<0.001$).
Davies et al. (2012) ; UK; n=32; 3.1-4.9 years, mean age 4.1 years; 11 male, 21 female	ActivPAL <i>Thigh</i> <i>1 sec epoch</i>	Direct observation (not specific tool)	1 hour activity in nursery setting (<i>free living</i>)	Time spent in the different postures (posture allocation): Sit/ lie, stand, walk	Paired t-test: ActivPAL underestimated total time spent sitting (mean difference -4.4%, $p<0.01$) and overestimated time standing (mean difference 7.1%, $p<0.01$). No difference in overall % time categorised as walk ($p=0.2$). Bland Altman: Direct observation with ActivPAL output showed that bias was not

					associated with the amount of time detected in each category ($r=-0.17$, $p=0.4$ for sit/lie; $r=-0.03$, $p=0.9$ stand). Sensitivity, specificity, positive predictive value (PPV): Sit/lie validation: individual, sensitivity = 92.8%; specificity = 97.3% and PPV = 97%. Group, specificity = 99.5% and PPV = 99.4% Stand validation: individual, sensitivity = 91.8%; specificity = 86.5% and PPV = 70.4%. Group, specificity = 87.9% and PPV = 72.4% Walk validation: individual, sensitivity = 77.9%; specificity = 96.5% and PPV = 73.4%. Group, specificity = 96.7% and PPV = 77.6%
Janssen et al. (2013b); Australia; n=38; 4-6 years, mean age 5.3 years; 20 male, 18 female	ActivPAL <i>Thigh</i> <i>1 sec epoch</i>	Direct observation (not specific tool)	150 minute protocol (laboratory based)	Posture allocation (sit/lie; stand; walk)	Classification accuracy determined by ROC-AUC, sensitivity and specificity: Sit/Lie: ROC-AUC = 0.84-0.88 (including and excluding 'other' postures

					respectively). Classification accuracy increased significantly when 'other' postures were excluded ($p < 0.05$). Standing: ROC-AUC = 0.76-0.77 (including and excluding other postures respectively). Walking: ROC-AUC = 0.73-0.74 (including and excluding other postures respectively). Dependent samples t-test: Differences between time spent in postures between ActivPAL and direct observation: Mean difference and 95% confidence intervals between the methods (ActivPAL-observation) were: Sit/Lie: +5.9% (0.6% to 11.1%) Stand: +14.8% (11.6% to 17.9%) Walk: -10.0% (-12.9% to -7.0%)

					No significant difference was found between time spent in sit/lie for activPAL and direct observation ($p=0.58$). Significant difference between time spent in stand/walk for activPAL and direct observation ($p<0.05$).
Janssen et al. (2014); Australia; N=18; 4-6 years, mean age 5.2 years; 9 male, 9 female	ActivPAL <i>Thigh</i> <i>15 sec epoch</i> <i>Cut points: produced during ROC curve analysis</i>	Direct Observation (CARS)	150 minute room calorimeter protocol (<i>laboratory based</i>)	ActivPAL- activity counts to determine MVPA and METs equation Direct observation- Activity levels	Classification accuracy determined by ROC-AUC, sensitivity, specificity: ActivPAL METs equation overestimated METs during SB (+6.0%) and underestimated METs for LPA (-15.3%), MVPA (-32.8%) and total METs (-13.6%) (all $p<0.001$). Classification accuracy for activPAL determined MVPA and direct observation: sensitivity = 88.3%, specificity = 88.2%, ROC-AUC = 0.88. Classification accuracy for activPAL determined MVPA

					when using both EE and direct observation as criterion: Sensitivity = 94.8%, Specificity = 84.8%, ROC-AUC = 0.90.
Ettienne et al. (2016); USA; n=30; Mean age =3.5 years; 17 male, 13 female 46% Native Hawaiian, 14% Other Pacific Islander- All of mixed ethnicities. Children recruited from Head Start sites	Actical (Respironics, Philips) <i>Non dominant wrist</i> <i>15 sec epoch</i> <i>Cut points: Schaefer et al., 2014</i> <i>Non-wear time: >1300 minutes of SB per day.</i>	Direct observation (System for observing fitness instruction time for preschoolers SOFIT-P))	7 days habitual activity examined by accelerometer, up to 3 hours of observation during pre-school hours (<i>free living</i>)	Actical- activity counts to determine activity intensity- SB, LPA, MPA, VPA. Direct observation – activity intensity categories	Cohen's Kappa: Proportion of agreement between Actical and direct observation =74%, weighed kappa coefficient was 0.17 (p<0.001).
Janssen et al. (2015); Australia; N=40, 4-6 years, mean age 5.3 years; 22 male, 18 female	Actical (Respironics, Philips) <i>Right hip</i> <i>15 and 60 sec epoch</i> <i>Cut points: Pfeiffer et al., 2006, Adolph et al., 2012 and Evenson et al., 2008</i>	Direct observation (CARS)	150 minute room calorimeter protocol (<i>laboratory based</i>)	Actical- AEE; activity intensity - SB, LPA, MVPA Direct observation- PA Intensity categories	Actical compared with direct observation Kappa: Adolph (60s): ($\kappa = 0.64$, 95% CI = 0.64–0.65) exhibited substantial agreement. Adolph(15s): ($\kappa = 0.58$, 95% CI = 0.57–0.58) and Evenson ($\kappa = 0.52$, 95% CI = 0.52–0.53) exhibited moderate agreement for predicting PA and SB determined by direct observation.

					Classification accuracy determined by ROC-AUC, sensitivity and specificity : Broken down by intensity: SB: Adolph (15s): ROC-AUC=0.80, sensitivity = 89.1%, specificity= 71.4% Adolph (60s): ROC-AUC=0.82, sensitivity = 82.5%, specificity=81.1% Evenson (15s): ROC-AUC=0.79, sensitivity = 91%, specificity= 68% Classification accuracy was significantly higher for Adolph compared to Evenson (p< 0.05). LPA: Adolph (15s): ROC-AUC=0.68, sensitivity = 51.9%, specificity= 84.3% Adolph (60s): ROC-AUC=0.73, sensitivity = 66.2%, specificity=80.2% Evenson (15s): ROC-AUC=0.65, sensitivity = 51%, specificity= 79.1% Adolph showed a significantly higher sensitivity (66.2%), which resulted in significantly

					higher classification accuracy compared to Evenson ($P < 0.05$). MVPA: Adolph (15s): ROC-AUC=0.82, sensitivity = 71%, specificity= 93.6% Adolph (60s): ROC-AUC=0.85 sensitivity = 74.3%, specificity=95.3% Evenson (15s): ROC-AUC=0.75, sensitivity = 54.9%, specificity= 96% Pfeiffer (15s): ROC-AUC=0.70, sensitivity = 43.9%, specificity= 96.5% Classification accuracy was significantly higher for Adolph compared to all others ($p < 0.05$). Actical compared with direct observation and EE: Kappa: κ values were slightly increased compared to using direct observation only. Adolph ($\kappa = 0.72$, 95% CI = 0.70–0.74), Adolph(15s) ($\kappa = 0.65$, 95% CI = 0.64–0.65) and Evenson ($\kappa = 0.61$, 95% CI = 0.60–0.62)

					exhibited substantial agreement predicting PA and SB determined by direct observation and calorimetry. Classification accuracy determined by ROC-AUC, sensitivity and specificity : ROC-AUC values when using direct observation combined with EE were slightly higher but similar to using direct observation only. Adolph was significantly better than all others ($p < 0.05$) when classifying LPA and MVPA. For SB, classification accuracy was good for the Adolph and Evenson (ROC-AUC = 0.82–0.85) with Adolph performing significantly better than Evenson. However, when comparing Evenson and Adolph(15s) this difference disappeared.
Byun et al. (2018b); USA;	Fitbit (Flex 1) <i>Non-dominant wrist</i> <i>60 sec epoch</i>	Direct observation (not a specific tool)	Children performed a set of unstructured and structured	Fitbit - Activity counts to define	Spearman's rank correlation:

n=28; 3-5 years, mean age 4.8 years; 15 male, 13 female			activities of different intensities for 34 minutes (<i>free living</i>)	activity intensities: SB, LPA, MPA, VPA Direct observation-activity intensity score: SB, LPA, MPA, VPA	Correlations between direct observation and Fitbit were statistically significant for SB $r=0.81$ ($p<0.05$), MVPA $r=0.62$ ($p<0.05$), and TPA $r=0.81$ ($p<0.05$), but not for LPA $r=0.21$. Relative agreement (in minutes): Fitbit recorded higher minutes of SB (mean difference = 2.3mins), and LPA (mean difference = 4.6mins), but lower mins of MVPA (mean difference = 6.9 mins) and total PA (mean difference = 2.3mins). Mean absolute % error: Lower for SB (28.8%) and TPA (11.5%), than for MVPA (46%) and LPA (92%). Kappa: Levels of agreement determined by the kappa statistic for SB ($k=0.78$) and TPA ($k=0.78$), weaker for

					LPA ($k=0.18$) and MVPA ($k=0.51$). ROC-AUC, sensitivity, specificity: SB: ROC-AUC: = 0.92; Sensitivity= 96.8%; Specificity=88.6% LPA: ROC-AUC = 0.63; Sensitivity= 55%; Specificity=70% MVPA: ROC-AUC = 0.77; Sensitivity= 53.3%; Specificity=99.4% TPA: ROC-AUC = 0.92; Sensitivity= 88.6%; Specificity=96.8%
Saris & Binkhorst (1977); Netherlands; n=4 ; 4.8-6.1 years; mean age and sex not reported	Actometer <i>Right ankle and wrist</i>	Direct observation (not specified)	Children engaged in free play activity in their pre-school setting (<i>free living</i>)	Actometer units Direct observation - Activity category (sitting, standing, walking, running) and intensity (high, low)	Correlations: Significant correlations between direct observation and actometer ankle $r=0.97$ ($p<0.01$), actometer wrist $r=0.71$ ($p<0.05$), actometer total $r=0.97$ ($p<0.01$).
PA (n=4)					
Liggett et al. (2012); New Zealand; N=14; 3-4 years old, mean age 3.9 years; 7 male, 7 female.	New Lifestyles NL-1000 <i>Right side of waistband 4 sec epoch</i>	Direct observation (CARS)	15-20 minutes activity at pre-school (<i>usual activity-children stopped every few minutes so</i>)	New Lifestyles NL-1000-MVPA Direct observation- Activity intensity level	Linear mixed model: Clear linear association between CARS and NL-1000. Most of the variation in CARS (95.4%, with 95% confidence intervals (CI

	<i>Cut points: Internal categorisation</i>		<i>researcher could check device)</i>		92.6–97.5) could be explained by variation in the NL-1000 measurements, with 2.4% of the variation being participant-specific.
Fairweather et al. (1999); Scotland, UK; n=11; mean age 3.7 years; 3 male, 8 female	Actigraph (CSA 7164) <i>Right hip</i> <i>1 min epoch</i> <i>Cut points: not specified</i>	Direct observation (CPAF)	40-50 minute structured play session (<i>usual activity</i>)	Actigraph- Counts per minute Direct observation- Activity level	Spearman's correlation: Rank order correlation between activity as determined by Actigraph and CPAF was high and statistically significant $r=0.79$ ($p<0.01$). Linear correlation: Mean Actigraph counts per min significantly correlated with mean CPAF counts per min $r=0.87$ ($p<0.01$).
Hislop et al. (2012b); Scotland, UK; N=31; 3-5 years, mean age 4.4 years; 15 males, 16 females	Actigraph (GT1M) <i>Right hip</i> <i>1 sec epoch</i> <i>Cut points: Sirard et al., 2005, Freedson et al., 2005, Pate et al., 2006, Evenson et al., 2008, Van Cauwenberghe et al., 2011, Puyau et al., 2002</i> RT3 (Stayhealthy)	Direct observation (CARS)	60 min free play in nursery setting (free living)	Accelerometers- Activity counts, MVPA Direct observation- MVPA	Spearman's correlation: Significant correlation between counts per minute and % of time spent in MVPA as recorded by direct observation for the GT1M $r=0.56$ ($p<0.01$) and for the RT3 $r=0.39$ ($p<0.03$). Bland Altman: Mean difference between number of minutes of MVPA recorded by GT1M

	<i>Right hip 1 sec epoch Cut points: Vanhelst et al., 2011, Rowlands et al., 2004; Sun et al., 2008, Chu et al., 2007</i>				and direct observation was 0.8 min (15s epoch) and 3.2 min (60s epoch). Mean difference for RT3 (walking) was 12.2 min (15s epoch) and 13.6 min (60s epoch) and for the RT3 (light jog) was 0 min (15s epoch) and 1.7 (60 s epoch).
Sharp et al. (2017); North Wales, UK; N=56; 3-4 years, mean age 3.7 years; 29 female, 27 male	FitBit (Zip) <i>Both right hip</i>	Direct observation (not specific tool)	5 min structured walking task (<i>nursery setting, not free living</i>)	Step counts	Bland Altman: Bias between observer counts with identified steps from devices: Fitbit 1: 22.8±19.1 (-14.7 to 60.2) steps and Fitbit 2: 25.2±23.2 (-20.2 to 70.5) steps. Absolute percent error: Fitbit 1: 6.44% (SE=0.66) Fitbit 2: 7.27% (SE=0.94) The majority of Fitbit 1 and Fitbit 2 step counts were in the exact-counting range (80.40% and 76.70%, respectively), with the remaining in the undercounting range.
SB (n=3)					
Reilly et al. (2003); Scotland, UK;	Actigraph (CSA/MTI AM7164)	Direct observation (CPAF)	<i>40 minutes of free play activity in</i>	Actigraph- Activity counts	Kruskal Wallis followed by Mann Whitney-U tests:

N=52; 3-4 years, mean age 3.5 years; 21 male, 31 female	<i>Right hip</i> <i>1 min epoch</i> <i>Cut point: Reilly et al., 2003</i>		<i>nursery setting (free living)</i>	CPAF scores - Activity level	Median accelerometer counts per minute differed significantly between CPAF categories 1 to 4 ($p < 0.001$). However, accelerometer output counts did overlap between categories.
De Decker et al. (2013); Belgium; n=44; mean age 5.5 years; 22 male, 22 female	ActivPAL <i>Thigh</i> <i>15 sec epoch</i> Actigraph (GT1M) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: Evenson et al., 2008</i> Both devices- <i>Wear time: First day of data deleted, minimum of 6 hrs per day required (this included data between 7am and 9pm on weekdays, 8am and 8pm on weekend days)</i> <i>Non wear time: ≥ 10 min of consecutive zero activity counts.</i>	Direct observation (not specific tool)	Monitors to be worn for 5 consecutive days whilst being directly observed for 1 hour during pre-school time (<i>free living</i>)	ActivPAL: Posture allocation, time spent in: sit/ lie, stand or stepping. Actigraph: SB or non-sedentary behaviour Direct observation: SB or non-sedentary behaviour	ROC-AUC, sensitivity and specificity: Sitting/Lying: ActivPAL: ROC-AUC: 0.61; Sensitivity = 53.8%; Specificity = 67.5%. Actigraph: ROC-AUC: 0.60; Sensitivity = 58.5%; Specificity = 61.2%. Sitting/Lying/Standing: ActivPAL: ROC-AUC: 0.52; Sensitivity = 27.8%; Specificity = 75.8% Actigraph: ROC-AUC: 0.61; Sensitivity = 46.3%; Specificity = 75.8%

	<i>Valid number of days: 3 days required to be included</i>				
Alghaeed et al. (2013); Scotland, UK; N=30; mean age 4.1 years; 10 male, 20 female	ActivPAL <i>Thigh</i> <i>1, 2, 5, 10 sec epochs</i>	Direct observation (not specific tool)	1 hour during usual activity in nursery (free living)	<i>Sitting bouts</i>	Bland Altman: 2s minimum sitting upright position (MSUP): minimised bias and showed no significant difference relative to direct observation (limits of agreement -14 to +17 bouts per hour, mean difference 1.83, p=0.02) 5s MSUP: underestimated number of sitting bouts as measured by direct observation (limits of agreement -23 to 8, mean difference -7.27, p=0.001.) 10s MSUP: underestimated number of sitting bouts as measured by direct observation (limits of agreement -29 to 4, mean difference -12.57, p=0.001) - 90% of sitting bouts were identical between activPAL

					and direct observation when using 2s MSUP. 5s and 10s underestimated number of sitting bouts.
Pedometers (n=8)					
PA and SB (n=4)					
McKee et al. (2005); Northern Ireland, UK; N=30; 3-4 years, mean age not reported; 13 male, 17 female	Yamax Digiwalker (SW-200) <i>Waist</i> <i>3 min sampling period</i>	Direct observation (CARS)	1 hour during usual activity during school time (<i>free living</i>)	Yamax Digiwalker SW-200 - step counts Direct observation: Physical activity level	Correlations: Correlation between the two methods: $r = 0.86$. Multilevel linear regression: CARS predicted an estimate of 88.60 (± 2.48) of pedometer step counts.
Hands et al. (2006); Australia; N=23; 5-6 years, mean age 5.5 years; 12 male, 11 female	Yamax Digiwalker SW-200 <i>Waist</i>	Direct observation (CARS)	Children monitored for 30 minutes over 5 days during free playing in their school setting (<i>free living</i>)	Yamax Digiwalker SW-200- Step count Direct observation- Activity levels	Pearson's correlations: Pedometer and direct observation $r = 0.90$ ($p < 0.01$). Pedometer shared 81% of variance with direct observation One way ANOVA: When the children were grouped into low, moderate, and high activity levels using observation, there was statistically significant differences between groups for the pedometer, with the pedometer data better able

					to separate the groups than accelerometer data.
Nishikido et al. (1982); Japan; N=49; 5-6 years, mean age not reported; 25 male, 24 female	Yamasa AM-5 Pedometer <i>Hip</i>	Direct observation (not specific tool)	2 days of children engaging in usual activities in pre-school setting (<i>free living</i>)	Yamasa AM-5 Pedometer- Step count Direct Observation- 'activity appearance rate' - running, walking, standing and sitting.	Pearson's correlation: Correlation coefficients between step rate and 1) running activity $r=0.69$ ($p<0.001$), $r=0.83$ ($p<0.001$); 2) walking $r=0.36$ ($p<0.05$), $r=0.42$, ($p<0.05$); 3) standing $r=-0.37$ ($p<0.05$), $r=-0.21$ (ns) and; 4) sitting $r=-0.50$ ($p<0.01$), $r=-0.77$ ($p<0.001$), in school 1 and 2 respectively.
Saris & Binkhorst (1977); Netherlands; N=11; 4.8-6.1 years, mean age and sex not reported	Pedometer (type not specified) <i>Right side of waist</i>	Direct observation (not specified)	Children engaged in free play activity in their pre-school setting (<i>free living</i>)	Pedometer units Direct observation - Activity category (sitting, standing, walking, running) and intensity (high, low)	Correlations: Significant correlations between direct observation and pedometer $r=0.93$ ($p<0.01$).
PA (n=4)					
Louie & Chan (2003); Hong Kong; N=145; 3-5 years, mean age 4.2 years; 84 male, 61 female	Yamax Digiwalker (SW-200) <i>Right hip (some invited to wear on both right and left hip)</i>	Direct observation (CARS)	25 minute physical activity class in normal school day (<i>free living</i>)	Yamax Digiwalker SW-200- step count Direct observation- Activity level	Pearson's correlation: Significant correlation was found between the total pedometer counts and CARS scores $r=0.64$ ($p<0.01$).

Oliver et al. (2007); New Zealand; N=13; 3-4.8 years, mean age 4.2 years; 7 male, 6 female	Yamax Digiwalker (SW-200) <i>Left hip or right hip (at anterior superior iliac spine), randomly assigned</i>	Direct observation (CARS)	35 minutes free- living activity (<i>crèche setting, not free living</i>)	Yamax Digiwalker SW-200- step count Direct observation- Activity level	Spearman's correlation: Significant correlation between direct observation and pedometry $r=0.59$ ($p = 0.04$). Wide 95% prediction intervals for the pedometer data against CARS.
Fotini, Antonis and Dimitra, 2015 Greece; N=45; mean age 5.5 years; sex not reported	Omron Walking Style Pro Pedometer (HJ-720IT-E2) and the Walking Style II (HJ-113) <i>Right and left hip, right and left front trouser pocket, back.</i>	Direct observation- Actual number of steps using hand counter (Basch SJ- 504).	Treadmill walking and track walking (<i>structured protocols</i>)	Step count	Standard error of measurement: Pedometers underestimated actual number of steps for treadmill walking by -0.5% to -38%. Accuracy improved at higher speeds. Pedometers accurate (- 3.5% to -0.6%) in measuring steps on over ground walking test. Back and right or left hip placements yielded most accurate information.
Murray (2009); USA; N=75; 3-5 years, mean age 4.4 years; 31 male, 44 female	Pedometer- MVP 4 Walk4Life Digital <i>Right or left hip</i>	Direct Observation (SOFIT-P)	Children were observed in pre- school for 4 consecutive minutes or 12 observations	MVP 4 Walk4Life Digital Pedometer- step count, MVPA, activity time	Pearson's correlations: No significant correlations at pre-test $r = 0.06$ ($p = 0.77$) and post-test $r = 0.30$ ($p = 0.36$)

34.3% African American, 64.4% Hispanic, 1.4% Native American Children recruited from Head Start Centres			whilst wearing pedometer at 2 time points (<i>free living</i>)	Direct observation- Activity level, type and context Results presented as: step counts, MVPA, activity time.	
Proxy reported measurement tools (n=2)					
PA and SB (n=1)					
Nishikido et al. (1982); Japan; n=49; 5-6 years, mean age not reported; 25 male, 24 female	Proxy report of children's Habitual Physical activity (Questionnaires to mothers and teachers)	Direct observation (not a specific tool)	Questionnaire assesses habitual activity. Remaining protocol for 2 days of children engaging in usual activities in pre-school setting (<i>free living</i>)	Proxy report: Children rated as: Inactive; relatively inactive; medial; relatively active; active. Direct Observation: 'activity appearance rate' - running, walking, standing and sitting.	Kendall's rank order correlation: Teacher's evaluations significant correlations with direct observation running activity appearance rate $r=0.27$ ($p<0.01$) and significant negative correlations with direct observation sitting activity $r=-0.19$ ($p<.05$). No other significant correlations.
PA (n=1)					
Noland et al. (1990); USA; n=21; 3.8-5.6 years, mean age 4.7 years; 11 male, 10 female 90% White, 10% Black	Parent and teacher report	Direct observation (CARS)	20 minute free play in school setting, sub-sample (8 <i>children</i>) observed for 6 hours at both home and school (<i>free-living</i>)	Parent and teacher report- activity level Direct observation- activity intensity score	Correlations: No significant correlations between parent and teacher report with direct observations.

Families were primarily middle and upper class.					

Abbreviations in the table: PA= physical activity; SB=sedentary behaviour; LPA= light physical activity; MPA=moderate physical activity; VPA= vigorous physical activity; MVPA=moderate to vigorous physical activity; TPA=total physical activity; MET= metabolic equivalent of task; HR= heart rate; EE=energy expenditure; AEE=activity energy expenditure; MSUP= minimum sitting upright position; CARS= Children's Activity Rating Scale; OSRAC-P= Observation System for Recording Physical Activity in Children- Preschool; SOFIT-P= System for Observing Fitness Instruction Time for Preschoolers; CPAF= Children's Physical Activity Form; DO= direct observation; RT3= Triaxial Research Tracker 3; ECG= electrodiagram; HRM = heart rate monitor; AUC-ROC=area under the receiver operating curve; CI= confidence intervals; LoA= limits of agreement; USA=United States of America; UK=United Kingdom

Table 3.9. Summary table of level 2 validity evidence of the measurement tools

Method	Outcome measures							References
	SB	Posture allocation	LPA	MVPA	TPA	Levels of activity	Step counts	
Direct observation								
Fargo Activity Timesampling Survey (FATS-continuous sampling)								Klesges, Woolfrey and Vollmer, 1985
Combined heart rate and accelerometer								
Actiheart								Adolph et al., 2012
Heart rate monitor (HRM)								
Polar Vantage XL Monitor								Bar-Or et al., 1996
Accelerometers								
Actigraph (GT3X+)	1, 2		1, 2	1, 2	1, 2			Hislop et al., 2016
Actigraph (GT3X)	1,3-7		1,3-6	1,3-6				Alhassan et al., 2017; Dobell et al., 2019; Janssen et al., 2013a; Kahan, Nicaise and Reuben, 2013
Actigraph (GT1M)	1,4-7	1	4-6	1,3-6,8				De Decker et al., 2013; Hislop et al.,

								2012a; Hislop et al., 2012b
Actigraph (MTI/CSA)	 4		 4	 4	 4	 7		Fairweather et al., 1999; Hands and Larkin, 2006; Kelly et al., 2004; Reilly et al., 2003; Sirard et al., 2005
Actical	 1,9		 1,9	 1,9-10		 11		Adolph et al., 2012; Ettienne et al., 2016; Janssen et al., 2015
ActivPAL								Alghaeed et al., 2013; Davies et al., 2012; De Decker et al., 2013; Janssen et al., 2013b; Janssen et al., 2014
Fitbit (Flex)								Byun et al., 2018b
Fitbit (Zip)								Sharp et al., 2017

NewLifestyles NL-1000								Liggett et al., 2012
Triaxial Research Tracker 3 (RT3)				 12-15				Adolph et al., 2012; Hislop et al., 2012b
Actiwatch (Spectrum)	 16		 16	 16				Alhassan et al., 2017
Actiwatch (MiniMitter)								Finn and Specker, 2000
Actiwatch (AW16)								Kelly et al., 2004
Actiwatch L								Djfarian et al., 2013
Pedometers								
Yamax Digiwalker (SW-200)								Hands and Larkin, 2006; Louie and Chan, 2003; McKee et al., 2005; Oliver et al., 2007
Yamasa AM-5 Pedometer								Nishikido, Kashiwazaki and Suzuki, 1982
MVP 4 Walk4Life Digital								Murray, 2009
Proxy reported measurement tools								

Teacher/mother reported habitual PA								Nishikido, Kashiwazaki and Suzuki, 1982; Noland et al., 1990

Published cut points used:¹ Evenson et al., 2008; ² Johansson et al., 2015; ³ Pate et al., 2006; ⁴ Sirard et al., 2005; ⁵ Van Cauwenberghe et al., 2011b; ⁶ Puyau, Vohra and Butte, 2002; ⁷ Reilly et al., 2003; ⁸ Freedson Pober and Janz, 2005; ⁹ Adolph et al., 2012; ¹⁰ Pfeiffer et al., 2006; ¹¹ Schaefer, Nace and Browning, 2014; ¹² Vanhelst et al., 2011; ¹³ Rowlands et al., 2004; ¹⁴ Sun, Schmidt and Teo-Koh, 2008; ¹⁵ Chu, McManus and Yu, 2007; ¹⁶ Ekblom et al., 2012

***Methodology used to assess the ability of the tool is detailed in the methods above and is indicated in the summary table as:**

Good =  Moderate =  Weak = 

Key for colour of boxes:  = evidence from ≥3 studies

 = evidence from <3 studies

This table shows a summary of the results of studies where they aimed to compare a particular measurement tool (e.g. Actigraph GT3X accelerometer) against direct observation (or electrodiagram for the heart rate monitor). The summary ratings were based on the quality of the tools for this specific measurement property. Where the measurement tool was deemed 'good' in the majority of the studies, the summary assessment was deemed 'good'. Where the measurement tool was deemed 'moderate' in the majority of the studies, the summary assessment was deemed 'moderate'. Where the measurement tool was deemed 'weak' in the majority of the studies, the summary assessment was deemed 'weak'. In instances where the measurement tool had mixed evidence in the studies, such as studies with outcomes of 'weak' and 'moderate', or 'moderate' and 'good', the overall assessment was deemed to be the most positive of the two outcomes. All tools of reasonable quality where any evidence was available are included in this table, including where only one or two studies reported that result.

3.5.4.3 Level 3 Validity

Level 3 validity evidence mainly consisted of proxy reported measurement tools, which were most frequently compared to accelerometry. Proxy reported measurement tools were generally poor at determining PA and SB outcomes. However, the PRE-PAQ was shown to be moderately accurate at determining stationary behaviour, LPA and VPA (Dwyer et al., 2011) and the *leisure time report* was able to determine MVPA (Manios, Kafatos and Markakis, 1998). The Netherland's physical activity questionnaire and *nursery teacher's report (based on Toyama Cohort Study survey questions)* could distinguish between different levels of activity (Chen et al., 2002; Janz, Broffitt and Levy, 2005). Although relatively few proxy reported tools demonstrated reasonable criterion or convergent validity, this could be due to a lack of face and content validity testing during the development of these tools (Hidding et al., 2020; Terwee et al., 2018a). It is also worth noting that this evidence is based on very few, low quality studies. An advantage of the proxy reported tools over the other measurement tools is that they were able to capture the context and type of the behaviours, such as screen time, rather than just determine movement.

The direct observation protocol Observation System for Recording Physical Activity in Children- Preschool (OSRAC-P) showed promising agreement with a heart rate monitor and pedometer for determining different levels of activity (Larson, Normand and Hustyi, 2011). Whilst the System for Observing Fitness Instruction Time for Preschoolers (SOFIT-P) did not demonstrate strong correlations with the output of the Actigraph (GT3X) (Sharma et al., 2011).

The Fitbit (Flex) showed excellent ability to determine SB and TPA, but not MVPA, when compared with the Actigraph (GT3X+) (Byun, Kim and Brusseau, 2018a). For

pedometers, activity counts from the Actigraph (CSA/MTI/GT1M) accelerometer were moderately correlated with step counts of the pedometer Yamax Digiwalker SW-200 (Cardon and De Bourdeaudhuij, 2007; Pagels, Boldemann and Raustorp, 2011) and the Omron Walking Style Pro Pedometer (HJ-720IT-E2) (De Craemer et al., 2015) (values ranging from $r=0.64$ to 0.92); suggesting that pedometers may be a plausible cheaper alternative to accelerometers in some instances (Hands and Larkin, 2006).

Table 3.10 provides detailed information and table 3.11 provides a summary table of the quality of included studies that examined level 3 validity evidence of measurement tools.

Table 3.10. Detailed information of studies of level 3 validity studies

Study details	Methods			Units of measure	Validity results
	Measurement tool(s) under study <i>Placement, epoch, cut points, wear time, non-wear time and number of valid days</i>	Comparison tool <i>Placement, epoch, cut points</i>	Study protocol <i>(lab/free living)</i>		
Direct observation (n=2)					
Physical activity (PA) and sedentary behaviour (SB) (n=2)					
Sharma et al. (2011); USA; n=27; 3-6 years, mean age 3.8 years; 11 male, 16 female 23.1% African American, 65.4% Hispanic, 7.7% Other (White/Native American) Children recruited from Head Start Centres	System for Observing Fitness Instruction Time for Preschoolers (SOFIT-P)	Accelerometer-Actigraph (GT3X) <i>Right hip 5 sec epoch Cut points: Sirard et al., 2005</i>	Children wore accelerometer for 1 entire school day, no reports on length of observations (<i>free living</i>)	SOFIT-P- % time in each physical activity (PA) level (Sedentary behaviours (SB), light physical activity (LPA), moderate physical activity (MPA), moderate to vigorous physical activity (MVPA)) Accelerometer- activity counts to determine SB, LPA, MPA, vigorous physical activity (VPA), MVPA	Pearson’s correlations: Correlation between % time in MPA using SOFIT-P with accelerometer data for % time in MPA r=0.506 (p=0.007) and MVPA r=0.532 (p=0.004) and SB r=−0.532 (p=0.004). Correlation with % time in MVPA on the SOFIT-P with accelerometer data for % time in MVPA r=0.541 (p=0.004), MPA r=0.530 (p=0.004) and SB r=−0.541 (p=0.003).
Larson et al. (2011);	Observation System for Recording	Heart Rate Monitor-Polar F-4	25 min of structured activity protocol	OSRAC-P – Activity levels (1 – stationary or	Agreement:

USA; N=4; 3-5 years, mean age not reported; 2 male, 2 female	Physical Activity in Children- Preschool (OSRAC-P)	<i>Chest and wristwatch</i> Pedometer- New Lifestyles -NL-2000 <i>Not reported</i>	under researchers instruction (<i>usual activity, but not free living</i>)	motionless; 2 – stationary with limb or trunk movement; 3 – slow, easy movements; 4 – moderate movements; 5 – fast movements) Heart rate monitor – beats per minute (bpm) Pedometer- step count	Activity levels 1 and 2 – step total were at or near 0. Most (3/4) participants had highest step total during activity level 5. Step total increased in activity level 3, dropped slightly during activity level 4. Pedometers registered increased steps during activity level 2 even though children were stationary. Heart rate bpm systematically increased from activity level 1 through to activity level 5.
Accelerometers (n=3)					
PA and SB (n=3)					
Martin et al. (2011); Scotland, UK; N=23; 3-5 years, mean age 4.5 years; 9 male, 14 female	ActivPAL <i>Thigh</i> <i>Wear time:</i> <i>>6 hours per day</i> <i>Non-wear time:</i> <i>Parents to complete</i> <i>daily log to state</i> <i>when devices were</i> <i>removed with</i> <i>reasons. Data</i> <i>between 7am and</i> <i>9pm eligible for</i> <i>inclusion.</i>	Actigraph (GT1M and GT3X) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: <1100 cpm=</i> <i>SB; >1100 cpm = PA</i> <i>Wear time:</i> <i>>6 hours per day</i> <i>Non-wear time:</i> <i>Parents to complete daily</i> <i>log to state when devices</i> <i>were removed with</i> <i>reasons. Data between</i>	Monitors worn simultaneously for 7 consecutive days during waking hours, except during water based activities (<i>free living activity</i>)	Actigraph- SB and PA ActivPAL- Sedentary behaviour (time spent sitting/lying plus 'quiet standing') and PA (any time not classified as SB)	Rank order correlations: Correlations between the two devices $r = 0.676$ ($p < 0.001$) and $r = 0.663$ ($p = 0.001$), both when Actigraph data were corrected and when uncorrected, respectively. Paired t-test: Differences in % of time spent sedentary between the monitors were significant when Actigraph data was corrected ($p < 0.001$) and not corrected ($p = 0.04$).

	Valid number of days: At least 3 days	7am and 9pm eligible for inclusion. Valid n of days: At least 3 days			Bland Altman: Difference for percentage of time spent sedentary of -4.3% $\pm$ 4.8 of daily time (and limits of agreement -14.0–5.4%) when Actigraph data were corrected, and -2.1% $\pm$ 4.6 (limits of agreement -11.4–7.2%) when Actigraph data were uncorrected.
Byun et al. (2018a); USA; n=27; 3-5 years, mean age 4.9 years; 16 male, 11 female	Fitbit (Flex) Non dominant wrist Non-wear time: ≥ 90 minutes consecutive zero counts Parents asked to report times for non- wear period, unusual PA and sleep.	Actigraph (GT3X+) Right hip 15 sec epoch Cut points: Evenson et al., 2008; Pate et al., 2006 Non-wear time: ≥ 90 minutes consecutive zero counts Parents asked to report times for non-wear period, unusual PA and sleep.	Children wore both monitors simultaneously for 24 hours (free living)	Fitbit- SB, LPA, MPA, VPA Actigraph- Activity counts to determine SB, MVPA, TPA	Pearson's correlation: Correlations between Actigraph and Fitbit were all significant. SB $r=0.85$ [Evenson cut point], $r=0.87$ [Pate cut point] ($p<0.01$), TPA $r=0.69$ [Evenson cut point], $r=0.71$ [Pate cut point] ($p<0.01$), MVPA $r=0.59$ [Pate cut point], $r=0.58$ [Evenson cut poing] ($p<0.01$). Overall correlations stronger between the Fitbit and Actigraph using Pate cut points $r=0.71-0.87$ ($p<0.01$), than Evenson cut points $r=0.69-0.85$ ($p<0.01$) Mean absolute percent errors (MAPE): No difference in the estimated

					SB when the Fitbit was compared with the Actigraph, although the Fitbit recorded slightly higher minutes of SB than the Actigraph (Pate) and the Actigraph (Evenson): (Fitbit, 673 minutes; Actigraph (Pate), 631 minutes; Actigraph (Evenson), 602 minutes)). The MAPE was lower when the Fitbit was compared with the Actigraph (Pate) (9.1%) than with the Actigraph (Evenson) (13%). The Fitbit produced significantly lower MVPA estimates than the Actigraph (Pate) and the Actigraph (Evenson) (Fitbit, 32 minutes; Actigraph (Pate), 99 minutes; and Actigraph (Evenson), 62 minutes); the MAPE was greater for the comparison with the Actigraph (Pate) (70.1%) than with the Actigraph (Evenson) (55.7%). No significant difference in total PA and MAPE was found between the Fitbit and Actigraph (Pate) (Fitbit, 337 minutes vs Actigraph (Pate),

					379 minutes; MAPE, 14.3%), but the difference and MAPE was slightly greater between the Fitbit and the Actigraph (Evenson) (Actigraph (Evenson), 407 minutes; MAPE, 18.1%).
Shin (2015); USA; N=19; 3-5 years, mean age 4.6 years; 10 male, 9 female	Best Fit Friend (BFF) accelerometer <i>Left and right wrist, back, waist, chest, right ankle, upper arm</i> <i>Cut points: determined using ROC-AUC analysis</i>	Actigraph (GT3X and GT3X+) <i>Waist and right wrist</i> <i>1 sec epoch</i>	Children followed structured activity protocol at the pre-school under researchers instruction (<i>usual activity, not free living</i>)	BFF Accelerometer- Raw signal data Actigraph Accelerometer- Vector magnitude counts	Pearson's Correlation: No significant correlations between BFF and Actigraph at either the wrist or waist, for all activities.
Pedometers (n=6)					
PA and SB (n=2)					
Pagels et al. (2011); Sweden/USA; N=55; 3.4-5.7 years, mean age 4.5 years; 28 male, 27 female	Pedometer – Yamax Digi-Walker (SW-200) <i>Right side of waist</i>	Accelerometer- Actigraph (GT1M) <i>Waist belt</i> <i>15 sec epoch</i> <i>Cut points: Sirard et al., 2005</i>	Children wore monitors simultaneously for 5 days during nursery time (<i>free living</i>)	Pedometer- step counts Accelerometer- activity counts	Linear regression analyses: Correlation between: 1) Mean step counts and mean accelerometer counts $r=0.67$ ($p<0.001$), accounting for 45% of variance and minutes of light to vigorous PA $r=0.76$ ($p=0.001$). Correlation between step counts and MVPA in 3 year olds $r=-0.19$ ($p=0.191$) and 4-5 year olds combined $r=0.50$ ($p=0.001$).

Saris & Binkhorst (1977); Netherlands; N=4; 4-6 year, mean age not reported; sex not reported	Pedometer (type not reported) <i>Right side of waist</i>	Actometer <i>Right ankle and right wrist</i>	Children engaged in free play activity in their pre-school setting (<i>free living</i>)	Pedometer and Actometer units	Correlations: Correlations between pedometer and actometer ankle $r=0.89$ ($p<0.001$); actometer wrist $r=0.78$ ($p<0.05$); actometer total $r=0.95$ ($p<0.001$). Correlations between actometer total and 1) actometer ankle $r=0.96$ ($p<0.001$) and 2) actometer wrist $r=0.86$ ($p<0.05$). No significant correlation between actometer ankle and wrist $r=0.69$.
PA (n=4)					
De Craemer et al. (2015); Belgium; N=41; 4-6 years, mean age 5.4 years; 21 male, 20 female	Pedometer - Omron Walking Style Pro (HJ-720IT-E2) <i>Right hip</i> <i>Wear time:</i> ≥ 5 hours data (recorded between 7am and 8pm weekdays, 7am and 9pm weekend days) <i>Non wear time:</i> Manually checked, if both monitors	Actigraph (GT1M) <i>Left hip</i> <i>15 sec epoch</i> <i>Wear time:</i> ≥ 5 hours data (recorded between 7am and 8pm weekdays, 7am and 9pm weekend days) <i>Non wear time:</i> Manually checked, if both monitors registered 0 per hour, the hour was deleted.	Children wore monitors simultaneously for four days (<i>free living</i>)	Pedometer- step count Accelerometer- step count and activity count	Pearson's correlations: Pedometer based steps correlated with accelerometer activity counts per hour $r=0.65$ ($p<0.001$) and per day $r=0.64$ ($p<0.001$), and with accelerometer step count per hour $r=0.92$ ($p<0.001$) and per day $r=0.89$ ($p<0.001$). Independent samples t-test No significant difference on group level between hourly steps ($p=0.359$) and daily steps ($p=0.592$) from both devices during the four days.

	<i>registered 0 per hour, the hour was deleted.</i>	<i>Valid number of days: At least 1 day of data.</i>			Paired sample t-tests: Significant difference on the individual level between hourly steps from both devices ($p=0.001$), but no significant difference was found between daily steps ($p=0.122$) Bland Altman: Agreement between step count on the devices. Bias of 221.81 (± 1679.78), limits of agreement ranged from -3070.57 to 3514.18 steps per day (accelerometer-pedometer).
Lee et al. (2014); Korea; N=131; 3-6 years, mean age not reported; sex not reported	Pedometer-Omron HJ-720ITC <i>Not reported</i>	Accelerometer-Actigraph (GT3X) <i>Not reported</i> <i>No cut points reported</i>	7 days activity (<i>free living</i>)	Accelerometer-MVPA Pedometer- step count	Pearson's Correlation: Correlations between total steps and MVPA were $r=0.59$ for 7 days, $r=0.55$ for weekdays and $r=0.37$ for weekend days. Younger age groups (3–4 years) showed higher correlation coefficients than older age groups (5–6 years): 7 days, $r=0.67$ vs. 0.54; weekday, $r=0.60$ vs. 0.52; and weekend, $r=0.56$ vs. 0.38, respectively.
Bikchu (2014); Hong Kong;	Pedometer-Yamax Digi-Walker (SW-700) <i>Left side of waist</i>	Accelerometer-RT3 (StayHealthy) <i>Right side of waist</i> <i>60 sec epoch</i>	30 minutes of structured activity protocol whilst wearing monitors	Accelerometer- activity units presented as vector magnitude score	Pearson's correlations: Correlations for various activities between devices ranged from $r=0.24$ to $r=0.78$,

N=143; 4-5 years, mean age not reported; 80 male, 63 female			simultaneously (<i>usual activity, but not free living</i>)	Pedometer- step count	with sliding ($r=0.16$ for girls) and cycling ($r=0.26$ for girls) having the lowest correlations. While correlations were all higher than $r=0.57$ for the other 8 play activities (crawl, throw/catch, kick, roll, agility run, walk balance, rhythmic, hopscotch).
Cardon & De Bourheaudhiuj (2007); Belgium; N=76; 4-5.9 years, mean age 4.9 years; 37 male, 39 female 55% from high SES; 45% from low SES	Pedometer- Yamax Digiwalker (SW-200) <i>Left hip</i> <i>Wear time: Pedometer data satisfactory when missing data (reported by parent) did not exceed 1 hour on 1 or more days.</i> <i>First day of data omitted (resulting in data being collected on 2 weekdays and 2 weekend days). Diary for parents to record daily step count, as well as time frame and reasons for removing the devices.</i>	Accelerometer- Actigraph (MTI AM 7164) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: Sirard et al., 2005</i> <i>Wear time: >8.6 hours per day</i> <i>First day of data omitted (resulting in data being collected on 2 weekdays and 2 weekend days). Diary for parents to record daily step count, as well as time frame and reasons for removing the devices.</i>	Monitors worn during waking hours for 5 consecutive days (<i>free living</i>)	Pedometer- step counts Accelerometer- Vector magnitude accelerometer output to determine minutes engaged in MVPA	Pearson's correlation: Correlation between mean step counts and minutes of MVPA engagement $r=0.73$ ($p<0.001$).

Proxy reported measurement tools (n=12)					
PA and SB (n=7)					
Bacardi-Gascón et al. (2012); Mexico; N=35; 3-5 years, mean age 4.4 years; 17 male, 18 female	Questionnaire developed for parents of pre-schoolers in Mexico Broad question on PA level	Accelerometer-Actigraph (GT1M) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: Pate et al., 2006 and Sirard et al., 2005</i> <i>Wear time: 10 hours per day</i> <i>Valid number of days: 3 weekdays and 1 weekend day</i>	No specific time frame- questionnaire assesses usual activity, accelerometer worn for 4 days (<i>free living</i>)	Questionnaire: Minutes and % of time in different intensities of PA (low-reported as SB, moderate, vigorous) PA level category: inactive to very active Actigraph accelerometer: % of time spent in different intensity of activities	Spearman's correlation: Significant correlation between parents' answers to broad PA question and accelerometer (Sirard cut points) for time in MPA, VPA $r=0.57$ and MVPA $r=0.62$ (both $p<0.001$). Correlations between % SB time reported by questionnaire and accelerometer; Sirard cut point: $r=0.35$ ($p<0.05$) and Pate cut points: $r=0.34$ ($p<0.05$). Correlations for % time spent in VPA and MVPA as reported by questionnaire were significant correlated with VPA as assessed by accelerometry; Sirard cut point: $r=0.53$ ($p=0.001$) and Pate cut point: $r=0.41$ ($p=0.01$)
Dwyer et al. (2011); Australia; N= 67; 3-5 years, mean age 3.8 years; 35 male, 32 female	Pre School Physical Activity Questionnaire (PRE-PAQ)	Accelerometer-Actigraph (MTI 7164) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: Sirard et al., 2005 and Reilly et al., 2003</i>	3 day activity (<i>free living</i>)	Pre-PAQ: 1-Stationary (no movement); 2- Stationary (limb/trunk movement); 3- moving slowly; 4- moving at a medium or moderate	Bland Altman: Agreement between PRE-PAQ and accelerometer (mean difference in mins per day): Stationary: 7.6 (Limits of agreement (LoA) -141.3 to 156.4)

91% White, 3% Mediterranean, 6% other ethnicity. 9% low SES, 34% middle, 57% high		<i>Wear time:</i> <i>>6 hours of recorded activity per day (between 8am and 6pm)</i> <i>Non wear time:</i> <i>>20 minutes of 0 counts not reported by parent log explaining when device had been removed</i> <i>Valid number of days:</i> <i>First day data excluded and accelerometer had to have 3 days of monitoring.</i>		pace; 5- moving at a fast pace. Accelerometer: Time in different intensities: Stationary, SB, LPA, MPA, VPA, MVPA, light MVPA (or non-sedentary)	SB: -208.6 (LoA -349.8 to -67.5 using Reilly cut points); -235.4 (LoA -383.1 to -87.7 using Sirard cut points) LPA: -4.8 (LoA -105.5 to 96.0 mins per day) MPA: 48.2 (LoA -24.9 to 121.3 mins per day) VPA: 1.9 (LoA -37.5 to 41.3 mins per day) MVPA: 50.1 (LoA -42.9 to 143.1 mins per day) Non-Sedentary: 20.9 (LoA -121.9 to 163.7 mins per day) LMVPA: 45.2 (LoA -103.6 to 194.1 mins per day) Wide limits of agreement. Pearson's Correlations: Stationary: $r=0.25$ Sedentary: $r=0.28$ (Reilly cut points), $r=0.19$ (Sirard cut points) LPA: $r=-0.07$ MPA: $r=0.13$ VPA: $r=0.17$ MVPA: $r=0.17$ Non-sedentary: $r=0.16$ LMVPA: $r=0.05$
Janz et al. (2005); USA;	Netherland's Physical Activity	Accelerometer-Actigraph (MTI 7164)	Questionnaire asks about activity over	NPAQ: Total activity score	Spearman's correlation:

N=204; 4-7 years, mean age 5.7 years; 91 male, 113 female 96% Caucasian Almost all had families of relatively high SES; only 10% listed a family income below 20,000 dollars.	Questionnaire (NPAQ)	<i>Waist Cut point for vigorous activity: ≥ 2818</i> <i>Wear time: At least 8 hours per day</i> <i>Valid number of days: At least 3 days of data.</i>	past 6 months, remaining protocol 4 days including 1 weekend day (<i>free living</i>)	Actigraph: activity counts to determine TPA and VPA	NPAQ and accelerometer derived variables were $r=0.33$ ($p<0.01$) for total activity and $r=0.36$ ($p<0.01$) for vigorous activity. TV viewing and accelerometry-derived variables were $r= -0.16$ ($p <0.05$) for total activity and $r= -0.18$ for vigorous activity ($p <0.05$). Associations between individual NPAQ items and accelerometer variables were lower than associations between total NPAQ score and accelerometry. Logistic Regressions: Children with high NPAQ scores were 2.7 times more likely to be in the upper tertile for total activity than those with low or moderate NPAQ scores, and 3.2 times more likely for vigorous activity. Children with high TV viewing score were 10% less likely to be in the upper tertile of total activity when compared to

					children with low or moderate TV viewing scores.
Manios et al. (1998); Greece; N=39; 6 years, 17 male, 22 female	Proxy report measures (Teacher Activity Rating / 3 Day Leisure Time Report –parental reported)	Heart rate monitoring – Sport Tester 3000 <i>Chest and wrist watch</i>	‘Real time’ activity- 5 day reports for teacher, 3 days for leisure reports, 3 days heart rate monitoring (<i>free living</i>)	Teacher rating: time spent in MVPA Parent rating: Time spent in different intensities of PA (SB, LPA, MVPA- focus on MVPA) Heart rate monitor: Activity intervals to determine MVPA (based on >149 bpm or above for at least 10 minutes)	Spearman’s correlation: Teacher Report- Significant correlations between reports and activity intervals during school hours $r=0.587$ ($p<0.001$), and during the 3 days of heart rate recording $r=0.406$ ($p<0.05$), not significant for weekend days $r=0.072$ ($p>0.05$). Leisure Time Report- Significant correlation between report and activity intervals for the 3 day report $r=0.715$ to 0.815 ($p<0.001$), significant correlations between MVPA reported on leisure report and activity intervals during the same days $r=0.68$ ($p<0.001$).
Nishikido et al. (1982); Japan; n=49; 5-6 years, mean age not reported; 25 male, 24 female	Proxy report of children’s Habitual Physical activity (Questionnaires to mothers and teachers)	Pedometer- Yamasa AM-5 <i>Hip</i>	Questionnaire assesses habitual activity. Remaining protocol for 2 days of children engaging in usual activities in pre-school setting (<i>free living</i>)	Proxy report: Children rated as: Inactive; relatively inactive; medial; relatively active; active. Yamasa AM-5: Step count	Kendall’s rank order correlation: Teacher’s evaluations significant correlations with pedometer step rate $r=0.25$ ($p<0.05$). No other significant correlations.

Wen et al. (2010); Australia; n=31 ; 3-5 years, mean age 3.5 years; 19 male, 12 female Children recruited from centers said to vary in SES.	Brief Survey of Activity Preferences (based on NPAQ) 7 day activity diary (adapted from CLASS)	Accelerometer- Actigraph (GT1M) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: Sirard et al., 2005</i> <i>Wear time:</i> <i>At least 8hr wear time per day</i> <i>Non wear time:</i> <i>Consecutive 0 counts for at least 60 minutes</i> <i>Valid number of days:</i> <i>At least 4 days of data</i>	7 day activity measured by diary and accelerometer; brief survey of habitual activity (<i>free living</i>)	Brief survey: activity preferences Diary: Total time in PA and SB (PA defined as walking and other MVPA activities.) Actigraph: activity counts to determine SB, MPA, VPA, and steps count	Spearman's correlation: Brief survey was not significantly correlated with any of the accelerometer outcomes (correlations ranged from -0.11 to 0.24). Significant correlations between the diary outcomes for PA (excluding walking) $r=0.42$ ($p=0.02$) and total PA $r=0.44$ ($p=0.014$) in comparison with vigorous PA determined by accelerometer. No other significant correlations, including no significant correlations for screen time or sedentary time.
Ettienne et al. (2016); USA; n=45; Mean age =3.5 years; 27 male, 18 female Children were primarily native Hawaiian or other Pacific Islander- all of mixed ethnicities.	Parental reported activity logs (report all children's activities for 3 days)	Actical (Respironics, Philips) <i>Non dominant wrist</i> <i>15 sec epoch</i> <i>Cut points: Schaefer et al., 2014</i> <i>Non-wear time: >1300 minutes of SB per day.</i>	7 days habitual activity examined by accelerometer and log (<i>free living</i>)	Accelerometer- activity counts to determine activity intensity- SB, LPA, MPA, VPA. Parental logs- Metabolic equivalent of task (MET) values and activity categories of behaviours	Cohen's Kappa: Proportion of agreement between the accelerometers and PA logs was 40%, weighed kappa coefficient was 0.15 ($p<0.001$). PA logs compared to accelerometer underestimated moderate and vigorous types of activity.

Children recruited from Head Start sites					
PA (n=4)					
Chen et al. (2002); Japan; n=21; 3-4 years, mean age 3.8 years; 12 male, 9 female	Nursery teacher's proxy report (based on Toyama Cohort Study survey questions)	Accelerometer-Actiwatch-L <i>Ankle</i> <i>1 min epoch</i> Caloriecounter- <i>Waist</i> <i>4 sec epoch</i>	Nursery teachers report of activity during the past week, remaining protocol for 3 days (<i>free living</i>)	Proxy report: frequency, level and preferences of activity Accelerometer Actiwatch-L: Activity counts Caloriecounter: total energy expenditure (TEE); energy expenditure from physical activity (EEPA); step count.	ANOVA: Children rated 'very often' activity had a significantly higher level of TEE, mean activity counts and daytime activity counts per day, compared with peers rated 'not often' (1257.1 $\pm$ 101.9 kcal vs. 1148.1 $\pm$ 71.4 kcal, p=0.019; 570.5 $\pm$ 192.8 counts vs. 334.9 $\pm$ 123.4 counts, p=0.026; 796.1 $\pm$ 194.9 vs. 486.7 $\pm$ 191.7 counts, p=0.014, respectively). Children rated 'very active' had a significantly higher TEE, EEPA, steps, mean activity counts and daytime activity counts, compared with 'inactive' peers (1257.1 $\pm$ 101.9 kcal vs. 1126.7 $\pm$ 124.4 kcal, p=0.024; 140.6 $\pm$ 17.5 kcal vs. 78.2 $\pm$ 14.7 kcal, p=0.03; 16103 $\pm$ 1896 steps vs. 10038 $\pm$ 320 steps, p=0.031; 570.5 $\pm$ 192.8 counts vs. 215 $\pm$ 165.8, p=0.002; 796.1 $\pm$ 194.9 vs. 396.6 $\pm$ 336.2 counts, p=0.02, respectively)

					Children with PA rated as "like very much" had higher activity counts than those with PA rated as "like" ($p < 0.05$). No other significant correlations between likeability of PA and indices of PA.
Chow & Au (2009); Hong Kong; n=23; 5-6 years, mean age not reported; 12 male, 11 female Majority of the children from lower to lower- middle families.	Habitual Activity Estimation Scale (HAES)– Parent Proxy report	Pedometer- Yamax SW-700 <i>Right side of waist</i> <i>Parents asked to report time of child waking up in the morning, sleeping at night and taking naps during the pedometer data collection period. Step count reading shown by the pedometer recorded by parent every day of use.</i>	Questionnaire examining usual activity, protocol involved 4 days of habitual activity (free living)	HAES- % MVPA Pedometer- daily step count (recorded each night by parent)	Pearson's correlation: No significant correlation between parent report of %MVPA and daily step counts of children for weekday $r = -0.33$ ($p = 0.13$) or weekend $r = 0.17$ ($p = 0.44$).
Corder et al. (2009); UK; N=27; 4-5 years, mean age 4.9 years; 17 male, 10 female	Children's physical activity questionnaire (CPAQ)	Accelerometer- Actigraph (MTI 7164) <i>Hip</i> 60 sec epoch <i>Cut points: >3000 counts per minute (cpm)</i> <i>Wear time: >600 minutes of valid data per day</i>	Questionnaire for past 7 days, remaining protocol for 11 days assessing habitual activity (free living)	CPAQ- Physical activity energy expenditure (PAEE) and time spent in MVPA Actigraph- Time spent in MVPA	Spearman's correlations: MVPA assessed by CPAQ significantly correlated with MVPA obtained from accelerometer $r = 0.42$ ($p = 0.04$). Bland Altman: Mean bias (MVPA) = 235.9 ± 362.0 (95% confidence intervals (CI) 86.5, 385.3).

		<i>Non wear time: 20 minutes of consecutive 0</i>			
Telford et al. (2004); Australia; n=58; 5-6 years, mean age 5.3 years; 37 male, 21 female	Children's Leisure Activities Study Survey (CLASS)	Accelerometer- Actigraph (MTI AM7164) <i>Waist 1 min epoch Cut points: Freedson et al., 1997</i> <i>Wear time: 10,000 movement counts per day</i> <i>Non wear time: When 10,000 movement counts per day were not collected it was assumed that the device had not been worn</i> <i>Valid number of days: At least 4 days of data</i>	Usual activity, accelerometer 8 days during waking hours (<i>free living</i>)	CLASS- units of time spent in MPA, VPA and MVPA (total) PA. Actigraph- activity counts to determine time (minutes per week) spent in moderate and vigorous intensity activities	Spearman's correlation: No significant correlation between proxy and accelerometer for MPA $r=-0.06$, VPA $r=-0.04$, TPA $r=-0.04$ or raw movement counts per day $r=0.05$ (individual level). Mean difference between proxy reported and accelerometer for MPVA was - 135.8 min per day (95% CI - 155.3 to -116.2), VPA -5.0 min per day (95% CI -13.3 to 3.4), TPA -140.7 min per day (95% CI -164.9 to -116.6).
SB (n=1)					
Mendoza et al. (2013); USA; N=96; 3-5 years, mean age 4.7 years;	TV Diary	Accelerometer- Actigraph (GT1M) <i>Hip 15 sec epoch Cut points: Reilly et al., 2003</i>	Measurement tools administered simultaneously for 7 days (<i>free living</i>)	Diary, TV allowance and EMA- TV and video viewing (mins per day) Accelerometer - Activity counts to determine SB	Spearman's Correlation: TV diary was not significantly correlated with accelerometer, but was significantly correlated with TV allowance (time 1 $r=0.45$, $p<0.001$ and time 2 $r=0.55$, $p<0.001$) and EMA

53 male, 41 female (sex of 2 participants not reported)		<i>Wear time:</i> ≥ 8 hours per day <i>Valid number of days:</i> ≥ 1 day TV allowance Ecological Momentary Assessment (EMA)			(time 1 $r=0.47$, $p<0.001$ and time 2 $r=0.51$, $p<0.001$).

Abbreviations in this table: PA= physical activity; SB=sedentary behaviour; LPA= light physical activity; MPA=moderate physical activity; VPA= vigorous physical activity; MVPA=moderate to vigorous physical activity; LMVPA= light moderate to vigorous physical activity; TPA=total physical activity; MET= metabolic equivalent of task; TEE =total energy expenditure; EEPA= energy expenditure from physical activity; bpm= beats per minute; cpm= counts per minute; OSRAC-P= Observation System for Recording Physical Activity in Children- Preschool; SOFIT-P= System for Observing Fitness Instruction Time for Preschoolers; AG= Actigraph; FF =Fitbit Flex; BFF= Best fit friend; MAPE= mean absolute percent errors; RT3= Triaxial Research Tracker 3; HRM = heart rate monitor; PRE-PAQ= Pre School Physical Activity Questionnaire; NPAQ= Netherland's Physical Activity Questionnaire; HAES= Habitual Activity Estimation Scale; CPAQ= Children's physical activity questionnaire; CLASS = Children's Leisure Activities Study Survey; EMA= Ecological Momentary Assessment; ROC-AUC=area under the receiver operating curve; CI= confidence intervals; LoA= limits of agreement; VM=vector magnitude; SES= socioeconomic status; USA=United States of America; UK=United Kingdom

Table 3.11. Summary table of level 3 validity evidence of the measurement tools

Measurement tools under study	Outcome measures						Step count	References
	SB	Posture allocation	LPA	MVPA	TPA	Levels of activity		
Direct observation								
OSRAC-P (Observation System for Recording Physical Activity in Children-Preschool)								Larson, Normand and Hustyi, 2011
SOFIT-P (System for Observing Fitness Instruction Time for Preschoolers)								Sharma et al., 2011
Accelerometers								
Fitbit (Flex)								Byun, Kim and Brusseau, 2018a
ActivPAL								Martin et al., 2011
Best Fit Friend								Shin, 2015
Pedometers								
Yamax Digi-Walker (SW-200)								Cardon and De Bourdeau

								dhuij, 2007; Pagels, Boldema nn and Raustorp, 2011
Yamax Digi-Walker (SW-700)								Bikchu, 2014
Omron Walking Style Pro (HJ-720IT-E2)								De Craemer et al., 2015; Lee et al., 2014
Proxy reported measurement tools								
Nursery teacher's report (<i>based on Toyama Cohort Study survey questions</i>)								Chen et al., 2002
Leisure time report								Manios, Kafatos and Markakis, 1998
Pre School Physical Activity Questionnaire (PRE-PAQ)								Dwyer et al., 2011

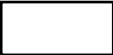
Netherland's Physical Activity Questionnaire (NPAQ)								Janz, Broffitt and Levy, 2005
Questionnaire developed for parents of pre-schoolers in Mexico								Bacardi-Gascón et al., 2012
Teacher activity rating								Manios, Kafatos and Markakis, 1998
7 day activity diary (adapted from CLASS)								Wen et al., 2010
Habitual Activity Estimation Scale (HAES)								Chow and Au, 2009
Children's physical activity questionnaire (CPAQ)								Corder et al., 2009
Children's Leisure Activities Study Survey (CLASS)								Telford et al., 2004

TV Diary								Mendoza et al., 2013
Teacher/mother reported habitual PA								Ettienne et al., 2016; Nishikido, Kashiwazaki and Suzuki, 1982

***Methodology used to assess the ability of the tool is detailed in the methods above and is indicated in the summary table as:**

Good =  Moderate=  Weak= 

Key for colour of boxes:  = evidence from ≥ 3 studies

 = evidence from < 3 studies

This table shows a summary of the results of studies where they aimed to compare a particular measurement tool (e.g. Pre School Physical Activity Questionnaire (PRE-PAQ)) against a measure which is considered to have *relatively low validity* such as a device based measurement tool (see Table 3.3 for a full explanation)). The summary ratings were based on the quality of the tools for this specific measurement property. Where the measurement tool was deemed 'good' in the majority of the studies, the summary assessment was deemed 'good'. Where the measurement tool was deemed 'moderate' in the majority of the studies, the summary assessment was deemed 'moderate'. Where the measurement tool was deemed 'weak' in the majority of the studies, the summary assessment was deemed 'weak'. In instances where the measurement tool had mixed evidence in the studies, such as studies with outcomes of 'weak' and 'moderate', or 'moderate' and 'good', the overall assessment was deemed to be the most positive of the two outcomes. All tools of reasonable quality where any evidence was available are included in this table, including where only one or two studies reported that result.

3.5.4.4 *Level 4 Validity*

A selection of studies included a comparison between two different makes of accelerometer each with unknown validity. These studies demonstrated that the combined heart rate with accelerometer, Actiheart, was shown to be similar in the activity count outcome to the accelerometers, Actical and RT3 ($r=0.80$ to 0.95) (Adolph et al., 2012). Similarly, the Actigraph (GT1M) and RT3 showed reasonable similarity in activity count outcome ($r=0.72$) (Hislop et al., 2012b). However, the majority of the studies did not show reasonable convergence, demonstrating that the various types of accelerometer can produce different outcomes (Alhassan et al., 2017; De Decker et al., 2013; Kelly et al., 2004; Van Cauwenberghe et al., 2012; Vanderloo et al., 2016). The increased availability of accelerometers and differences in the outcomes of these studies demonstrates the importance of assessing validity of different devices simultaneously, alongside comparison measures (Oliver, Schofield and Kolt, 2007). There was weak comparisons between parental and teacher reported habitual physical activity (Nishikido, Kashiwazaki and Suzuki, 1982; Noland et al., 1990).

Table 3.12 provides detailed information and table 3.13 provides a summary table of the quality of included studies that examined level 4 validity evidence of measurement tools.

Table 3.12. Detailed information of studies of level 4 validity evidence

Study details	Methods		Units of measure	Validity results
	Measurement tools under study <i>Placement, epoch, cut points, wear time, non-wear time and number of valid days</i>	Study protocol <i>(lab/free living)</i>		
Combined heart rate and accelerometry and accelerometers (n=1)				
Physical activity (PA) and Sedentary Behaviour (SB) (n=1)				
Adolph et al. (2012); USA; n=64; 3-5 years, mean age 4.5 years; 37 male, 27 female 50% White, 27% Black, 22% Hispanic, 2% Asian	Actiheart (MiniMitter) <i>Chest</i> <i>15 sec epoch</i> Actical (Respironics, MiniMitter) <i>Right hip</i> <i>15 sec epoch</i> Triaxial Research Tracker 3 (RT3- Stayhealthy) <i>Right hip</i> <i>1 sec epoch</i> <i>Cut points: Piecewise linear regressions determined thresholds into sedentary, light, moderate and vigorous levels of PA</i>	All accelerometers worn during 3 hour whole room calorimeter protocol <i>(laboratory based)</i>	Accelerometers- activity counts per minute	Pearson's correlations: Accelerometer counts from each of the devices were significantly correlated with one another $r=0.80-0.95$ ($p=0.001$).
Accelerometers (n=6)				
PA and SB (n=3)				
Alhassan et al. (2017);	Actiwatch (Spectrum) <i>Non dominant wrist</i>	Monitors worn 24 hours per day for up	Accelerometers- activity counts to determine sedentary	Spearman's Correlation:

USA; N=33; mean age 4.4 years; 21 male, 12 female	<i>15 sec epoch</i> <i>Cut points: Ekblom et al., 2012</i> Actigraph (GT3X) <i>Waist, centre of lower back</i> <i>15 sec epoch</i> <i>Cut points: Pate et al., 2006</i> <i>and Sirard et al., 2006</i>	to 16 days (<i>free living</i>)	behaviour (SB), light physical activity (LPA), moderate physical activity (MPA), vigorous physical activity (VPA), moderate to vigorous physical activity (MVPA)	Correlation between Actiwatch and Actigraph $r=0.41$.
Kelly et al. (2004); UK; n=78; 3-4 year olds, mean age 3.5 years; 30 male, 48 female	Actiwatch (AW16) <i>Right hip</i> <i>1 min epoch</i> Actigraph (CSA/MTI – WAM 7164) <i>Right hip</i> <i>1 min epoch</i>	39-45 minutes structured play class (<i>usual activity, reflective of free living</i>)	Accelerometers- activity counts per minute	Correlation: Correlation between Actigraph and Actiwatch counts $r = 0.36$ ($p < 0.01$).
Vanderloo et al. (2016); Canada; N=23; 4-5 years, mean age 5.08 years; 12 male, 16 female (sex sample size prior to exclusion of participants)	Actical (B series) <i>Right hip</i> <i>15, 60 sec epoch</i> <i>Cut points: Pfeiffer et al., 2006</i> Actigraph (GT3X+) <i>Right hip</i> <i>15, 60 sec epoch</i> <i>Cut points: Pate et al., 2006</i> Both devices- <i>Wear time:</i> <i>8 hours per day</i> <i>Non-wear time:</i> <i>>60 minutes of consecutive zeroes</i>	Both monitors worn simultaneously for 7 consecutive days (<i>free living</i>)	Accelerometers- Activity counts to determine SB, total physical activity (TPA), MVPA	Paired t-tests, with Bonferroni correction: Lower rates of both MVPA and TPA were measured with the Actical method compared with Actigraph ($p < 0.001$). Higher SB by Actical compared with Actigraph ($p < 0.001$), at both 15 and 60 second epoch. At 15s epoch: significantly lower rates of MVPA ($t[22] = -12.75$, $p < 0.00$, Cohen's $d = -2.93$) and TPA ($t[22] = -5.75$, $p < 0.00$, Cohen's $d = -1.52$) measured with the Actical compared with the ActiGraph. Significantly higher level of SB via the Actical in comparison with

	Valid number of days: >3 days valid data All wear time information recorded in daily log by parents/guardians.			the ActiGraph ($t[22] = 11.00$, $p < 0.00$, Cohen's $d = 1.73$). At 60s epoch: significantly lower rates of MVPA ($t[22] = -11.57$, $p < 0.00$, Cohen's $d = -2.87$) and TPA ($t[22] = -12.50$, $p < 0.00$, Cohen's $d = -2.54$) measured via the Actical in comparison with the ActiGraph. Significantly higher level of SB with the Actical in comparison with the ActiGraph ($t[22] = 12.41$, $p < .00$, Cohen's $d = 2.14$). Bland Altman (mean difference in minutes): 15s epoch: SB (9.18 ± 7.84), MVPA (-6.61 ± 4.87), TPA (-9.41 ± 15.37). 60s epoch: SB (12.70 ± 9.62), MVPA (-6.78 ± 5.50) and TPA (-14.19 ± 10.67). The difference between Actical and Actigraph: 95.7% (MVPA), 95.7% (TPA) and 95.7-100% (SB) of the values were within 2SD of the difference between the devices.
PA (n=1)				
Hislop et al. (2012b); Scotland, UK; N=31; 3-5 years, mean age 4.4 years; 15 male, 16 female	Actigraph (GT1M) Waist band 1 sec epoch Cut points: Sirard et al., 2005, Freedson et al., 2005, Pate et al., 2006, Evenson et al., 2008, Van Cauwenberghe et al., 2011, Puyau et al., 2002	60 min free play in nursery setting (<i>free living</i>)	Accelerometers- Activity counts	Pearson's Correlation: Significant correlation in total counts between the two accelerometers $r=0.72$ ($p < .001$).

	RT3 (Stayhealthy) <i>Waist band</i> <i>1 sec epoch</i> <i>Cut points: Vanhelst et al., 2011, Rowlands et al., 2004, Sun et al., 2008, Chu et al., 2007</i>			
SB (n=2)				
De Decker et al. (2013); Belgium; n=52; mean age 5.5 years; 26 male, 26 female	ActivPAL <i>Thigh</i> <i>15 sec epoch</i> Actigraph (GT1M) <i>Right hip</i> <i>15 sec epoch</i> <i>Cut points: Evenson et al., 2008</i> Both devices- <i>Wear time: First day of data deleted, minimum of 6 hrs per day required (this included data between 7am and 9pm on weekdays, 8am and 8pm on weekend days)</i> <i>Non wear time:</i> <i>≥10 min of consecutive zero activity counts.</i> <i>Valid number of days: 3 days required to be included</i>	Children wore monitors simultaneously for 5 days (<i>free living</i>)	ActivPAL: Posture allocation- Time spent in sit/ lie, stand or stepping. Actigraph: Activity counts to determine sedentary or non-sedentary behaviour	ANOVA: Significant differences between the Actigraph and ActivPAL data ($p<0.05$). Bland Altman: Mean time defined as SB was lower for ActivPAL compared to Actigraph (mean bias $-7.71\% \pm 10.87\%$). No systematic bias for SB and wide limits of agreement in 95% confidence (-29.01% to 13.6%). Mean time defined as SB including standing still indicated a mean bias of $-7.26\% \pm 21.15\%$ and wide limits of agreement (-48.75% to 34.21%) were found in 95% confidence analysis.

Van Cauwenberghe et al. (2012); New Zealand; N=49; 3-4 years, mean age 4 years; 22 male, 27 female	Actical (Respironics, MiniMitter) <i>Hip (above the right iliac crest)</i> 15 sec epoch <i>Cut points: Evenson et al., 2008</i> ActivPAL <i>Thigh</i> 15 sec epoch <i>Wear time:</i> At least 3 hours of data from both devices	Monitors worn simultaneously during 1 day at nursery (<i>free living</i>)	Both monitors: time spent in SB and non-SB	Bland Altman: Group mean time in SB as detected by the Actical and activPAL was similar and that the 95% limits of agreement for time in SB were wide. No obvious systematic bias. K statistic: The κ statistic = 0.46 (95% CI: 0.45–0.47; $p < 0.001$). Percent agreement: The difference in SB between both monitors for each participant showed 73% agreement and 27% disagreement ActivPAL classified 15.4% of all epochs as standing still, with 54.9% of these defined as SB by Actical. Epochs defined as SB by Actical and non-SB by ActivPAL, 61.6% regarded as standing still by ActivPAL.
Proxy reported measurement tools (n=2)				
PA and SB (n=1)				
Nishikido et al. (1982); Japan; n=49; 5-6 years, mean age not reported; 25 male, 24 female	Proxy report of children's habitual physical activity (Mothers and teachers)	Questionnaire assesses habitual activity (<i>free living activity</i>)	Proxy report- Children rated as: Inactive; relatively inactive; medial; relatively active; active.	Kendall's rank order correlation: Correlation between mother and teacher evaluations were $r = 0.41$ ($p < 0.01$) and $r = 0.55$ ($p < 0.001$) in school A and B, respectively.
PA (n=1)				

Noland et al. (1990); USA; n=21; 3.8-5.6 years, , mean age 4.7 years; 11 male, 10 female 90% White, 10% Black Families were primarily middle and upper class.	Parent, Teacher and Child Report	Habitual overall activity level (<i>free-living</i>)	Parent and teacher report-activity level Child report- activity preference	Correlations: Significant correlation between behaviourally anchored rating from parents and teachers $r=0.40$ ($p<0.05$). Children's activity preference not significantly correlated with parents or teachers behaviourally anchored rating.

Abbreviations in this table: PA= physical activity; SB= sedentary behaviour; LPA= light physical activity; MPA=moderate physical activity; VPA=vigorous physical activity; MVPA= moderate to vigorous physical activity; TPA= total physical activity; RT3= Triaxial Research Tracker 3

Table 3.13. Summary table of level 4 validity evidence of the measurement tools

Measurement tools under study	Outcome measure					Reference
	SB	MVPA	TPA	Activity counts	Habitual physical activity reports	
Combined heart rate and accelerometer and accelerometers						
Actiheart, Actical and Triaxial Research Tracker 3 (RT3)						Adolph et al., 2012
Accelerometers						
Actigraph (GT1M) ¹⁻⁶ and RT3 ⁷⁻¹⁰						Hislop et al., 2012b
Actigraph (GT1M) ⁴ and ActivPAL						De Decker et al., 2013
Actigraph (GT3X+) ³ and Actical ¹¹						Vanderloo et al., 2016
Actigraph (GT3X) ^{1,3} and Actiwatch (Spectrum) ¹²						Alhassan et al., 2017
Actigraph (CSA/MTI) and Actiwatch (AW16)						Kelly et al., 2004
Actical ⁴ and ActivPAL						Van Cauwenbergh et al., 2012
Proxy reported measurement tools						
Parental versus teacher reports						Nishikido, Kashiwazaki and Suzuki,

						1982; Noland et al., 1990

Cut points: ¹ Sirard et al., 2005; ² Freedson Pober and Janz, 2005; ³ Pate et al., 2006; ⁴ Evenson et al., 2008; ⁵ Van Cauwenberghe et al., 2011b; ⁶ Puyau, Vohra and Butte, 2002; ⁷ Vanhelst et al., 2011; ⁸ Rowlands et al., 2004; ⁹ Sun, Schmidt and Teo-Koh, 2008; ¹⁰ Chu, McManus and Yu, 2007; ¹¹ Pfeiffer et al., 2006; ¹² Ekblom et al., 2012.

***Methodology used to assess the ability of the tool is detailed in the methods above and is indicated in the summary table as:**

Good = ● Moderate= ● Weak= ●

Key for colour of boxes:  = evidence from ≥3 studies

 = evidence from <3 studies

This table shows a summary of the results of studies where they aimed to compare two (or more) of the same type of measurement tools where *neither tool is considered to have known higher validity* (e.g. comparison between Actical and ActivPAL). The summary ratings were based on the quality of the tools for this specific measurement property. Where the measurement tool was deemed 'good' in the majority of the studies, the summary assessment was deemed 'good'. Where the measurement tool was deemed 'moderate' in the majority of the studies, the summary assessment was deemed 'moderate'. Where the measurement tool was deemed 'weak' in the majority of the studies, the summary assessment was deemed 'weak'. In instances where the measurement tool had mixed evidence in the studies, such as studies with outcomes of 'weak' and 'moderate', or 'moderate' and 'good', the overall assessment was deemed to be the most positive of the two outcomes. All tools of reasonable quality where any evidence was available are included in this table, including where only one or two studies reported that result.

3.5.4.5 Face and content validity

Only two studies commented on the face and content validity of the proxy reported measurement tools (Bacardi-Gascón et al., 2012; Dwyer et al., 2011). Face and content validity was determined by focus groups with parents and pre-school staff and consulting experts during questionnaire development (Dwyer et al., 2011); and a pilot study with 21 parents, to determine the comprehension and reproducibility of the measure (Bacardi-Gascón et al., 2012). The Pre-PAQ included individual response options for both weekend days, due to parents indicating that children's PA varied more on a weekend than during the week (Dwyer et al., 2011). No further information was reported on the level of face and content validity within these studies; however, no major comprehension concerns were reported. There was minimal information about the procedures in which face/content validity was assessed in these studies, and so it was not possible to assess the quality or provide firm conclusions on the content validity.

3.5.4.6 Test-retest reliability

Several of the proxy reported tools showed reasonable test-retest reliability (values ranging from $r=0.76$ to 0.94) (Bacardi-Gascón et al., 2012; Manios, Kafatos and Markakis, 1998; Mendoza et al., 2013; Telford et al., 2004) or variable test-retest results, whereby good test-retest for some items on the questionnaire but poor for other items (Dwyer et al., 2011; González-Gil et al., 2014; Janz, Broffitt and Levy, 2005; Manios, Kafatos and Markakis, 1998). There was very limited evidence showing the test-retest of accelerometry and pedometry, with good test-retest for some activities but poor for others (ICC range from 0.34 to 0.87) (Bikchu, 2014). Table 3.14 provides detailed information and table 3.17

provides a summary table of the quality of included studies that examined test-retest reliability evidence of measurement tools.

3.5.4.7 Intra-Instrument Reliability

One accelerometer, the Fitbit (Zip) was the only tool to be examined for intra-instrument reliability in the included studies; showing excellent intra-instrument reliability (ICC=0.91) when two devices were worn simultaneously on the right hip during a 5 minute structured walking task in the nursery (Sharp et al., 2017). Table 3.15 provides detailed information and table 3.17 provides a summary table of the quality of included studies that examined intra-instrument reliability evidence of measurement tools.

Table 3.14. Detailed information of studies examining test-retest reliability of measurement tools

Study details	Methods		Units of Measure	Reliability Results
	Measurement tool	Test Retest period		
Accelerometers and pedometers (n=2)				
Physical activity (PA) (n=2)				
Bikchu (2014); Hong Kong; N=143; 4-5 years, mean age not reported. 80 male, 63 female	Pedometer- Yamax Digi-Walker SW-700 <i>Left side of waist</i>	Same structured activity protocol whilst wearing monitors simultaneously completed at 2 time points	Step count	Intaclass correlation coefficient (ICC): ICC= 0.34 to 0.87.
	Accelerometer- TriTrac RT3 triaxial (StayHealthy) <i>Right side of waist</i> <i>60 sec epoch</i>		Vector magnitude score	
Murray (2009) ; USA; n=75; 3-5 years, mean age 4.4 years; 31 male, 44 female 34.3% African American, 64.4%	MVP 4 Walk4Life Digital Pedometer <i>Left or right hip</i>	Each child was observed by two measurements at baseline and two measurements at post-test whilst wearing	Step counts Moderate to vigorous physical activity (MVPA) Activity time	Pearson’s Correlation: No significance at pre-test, yet show significant results for all physical activity (PA) outcomes at post-test: Step counts r=0.52 (p<0.001). MVPA r=0.57 (p<0.001) Activity time r=-0.55 (p<0.001).

Hispanic, 1.4% Native American Children recruited from Head Start Centres		pedometer and engaging in free play activities.		
Proxy reported tools (n=8)				
PA and Sedentary behaviour (SB) (n=5)				
Bacardi-Gascón et al. (2012); Mexico; N=21; 3-5 years, mean age not reported; sex not reported	Questionnaire developed for parents of pre-schoolers in Mexico	7 days	Duration of activities in: Low Moderate Vigorous	Correlations: Low $r=0.86$ ($p=0.01$) Moderate $r=0.79$ ($p=0.04$) Vigorous $r=0.94$ ($p=0.002$)
	Broad question on PA level			$r=0.97$ ($p=0.001$)
Dwyer et al. (2011); Australia; N=103; 3-5 years, mean age 3.8 years; 54 males, 49 females 87% Caucasian, 6% Mediterranean, 7% other ethnicity 7% low socioeconomic status (SES), 29% middle, 64% high	Pre School Physical Activity Questionnaire (PRE-PAQ)	1-2 weeks	Pre-PAQ levels 1-2 Pre-PAQ levels Pre-PAQ levels 4 Pre-PAQ levels 5	Intraclass correlation coefficient: ICC= 0.44 ICC =0.53 ICC =0.44 ICC =0.64

Janz et al. (2005) ; USA; n=72; 4-7 years, mean age 5.7 years; sex not reported 96% Caucasian Almost all had families of relatively high SES; only 10% listed a family income below 20,000 dollars.	Netherland's Physical Activity Questionnaire (NPAQ)	2-8 weeks	NPAQ items	Spearman's Correlation: All individual NPAQ items were significant ($p < 0.05$) and ranged from $r=0.30$ to 0.66. Outdoor play preferences had the lowest correlation $r=0.30$. Intraclass correlation coefficient: $ICC=0.70$ (95% CI = $0.56, 0.80$, $p < 0.01$). Cohen's Kappa: $\kappa=0.39$ (95% CI = $0.22, 0.56$). Individual items ranged from $\kappa=0.21$ to 0.48. Outdoor play preference had lowest kappa score $\kappa=0.21$.
	TV viewing		TV viewing	Intraclass correlation coefficient: $ICC=0.70$(95% CI=$0.55, 0.80$ $p < 0.01$). Cohen's Kappa: $\kappa=0.53$ (95% CI= $0.35, 0.74$).
Manios et al. (1998); Greece; N=39; 6 years; 17 male, 22 female	Proxy report measures - 3 Day Leisure Time Report –parental reported	2 weeks	Time spent in: Sedentary Light physical activity (LPA) Moderate to vigorous physical activity (MVPA)	Spearman's Correlation: SB- $r=0.53$ ($p=0.002$) LPA- $r=0.27$ ($p>0.05$). MVPA- $r=0.64$ ($p<0.001$).
	Teacher activity rating		Time spent in MVPA	$r=0.84$ ($p<0.001$).
González-Gil et al. (2014); Belgium, Bulgaria, Germany, Greece, Poland and Spain;	'Toybox' Primary Caregivers Questionnaire on Lifestyle Behaviours	2 weeks	PA SB	Intaclass correlation coefficient: PA- ICC range = $0.304-0.966$ (25.9% of questions excellent reliability; 22.2% good; 40.7% moderate; 11.1% poor)

n=93; 3.5-5.5 years; sex not reported 29 children attended kindergartens from the low SES, 26 from middle, 38 from high.				SB- ICC range = 0.058–0.959 (10.5% excellent; 44.7% good; 31.6% moderate; 13.2% poor)
PA (n=2)				
Corder et al. (2009); UK; n=20; 4-5 years; 12 male, 8 female	Children's physical activity questionnaire (CPAQ)	7 days	MVPA Physical activity energy expenditure (PAEE)	Intaclass correlation coefficient: MVPA- ICC = 0.39 (p<0.05). PAEE- ICC =0.25
Telford et al. (2004); Australia; n=58; 5-6 years, mean age 5.3 years; 37 male, 21 female	Children's Leisure Activities Study Survey (CLASS)	2 weeks	Frequency (Q1+Q2) Duration (Q1+Q2)	Intaclass correlation coefficient: ICC=0.83 (p<0.001) ICC=0.76 (p<0.001).
SB (n=1)				
Mendoza et al. (2013); USA; N=80; 3-5 years; sex not reported 100% Latino or Hispanic. Children recruited from Head Start Centers	TV Diary	3-4 weeks	Minutes per day of TV and video viewing	Intaclass correlation coefficient: ICC=0.82 (p<0.001)

Table 3.15. Detailed information of studies examining intra-instrument reliability of measurement tools.

Study details	Methods		Units of Measure	Results
	Measurement tool(s) examined <i>Placement</i>	Study protocol <i>(lab/free living)</i>		
Accelerometers (n=1)				
PA (n=1)				
Sharp et al. (2017); North Wales, UK; n=56; 3-4 years, mean age 3.7 years; 27 male, 29 female	Fitbit Zip <i>Both right hip</i>	Two Fitbit Zip monitors worn simultaneously on the right hip during a 5 minute structured walking task in the nursery (<i>lab- structured activity protocol</i>)	Step counts	Intraclass correlation coefficient (ICC) ICC between the 2 Fitbits =0.91 (95% confidence intervals (CI)= 0.85, 0.95)

Abbreviations in these tables: PA=physical activity; SB=sedentary behaviour; ICC= Intraclass correlation coefficient; k=kappa; RT3=; MVPA=moderate to vigorous physical activity; LPA= light physical activity; PAEE= physical activity energy expenditure; PRE-PAQ=Pre School Physical Activity Questionnaire; NPAQ= Netherland's Physical Activity Questionnaire; CPAQ= Children's physical activity questionnaire; CLASS= Children's Leisure Activities Study Survey; TV=television; SES=socioeconomic status; UK=United Kingdom; USA=United States of America

3.5.4.8 Feasibility

Whole room calorimetry was shown to be accepted by pre-school aged children as a way of measuring PA and SB (Janssen et al., 2013c; Oortwijn et al., 2009). However, this method is expensive, cannot examine free living activity, may only be feasible for smaller scale projects and is highly burdensome on researchers due to the required training and expertise; thereby not being viable for surveillance and the majority of research projects (Fulton et al., 2001). Similarly, there were promising results for the feasibility of direct observation protocols (Puhl et al., 1990; Sharma et al., 2011). However, due to the intensive and demanding nature of direct observation, there are limits on the practicality of this method. Observations usually take place in just one location and for a short period of time, impacting the viability of this method in large samples and to identify habitual activity (Fulton et al., 2001).

Feasibility and acceptability of device based measurement was generally high (Bar-Or et al., 1996; Cardon and De Bourdeaudhuij, 2007; De Decker et al., 2013; Ellison et al., 1992; Van Cauwenberghe et al., 2012), even when more than one device was worn simultaneously (Cardon and De Bourdeaudhuij, 2007; Van Cauwenberghe et al., 2012). Although high acceptability of the ActivPAL accelerometer was reported (Van Cauwenberghe et al., 2012), there was evidence of concerns of irritability of the ActivPAL due to these being attached directly to the skin (Davies et al., 2012; De Decker et al., 2013). This may also be a concern when using other devices that attach in a similar way (Costa et al., 2013; Shin, 2015). The proportions of missing and excluded data, for reasons such as device malfunction or children not wearing the device for a sufficient period of time, should be considered when calculating the sample size for studies.

There were no studies that determined the feasibility of proxy reported measurement tools through qualitative research. Table 3.16 provides detailed information and table 3.17 provides a summary table of the quality of included studies that examined feasibility of measurement tools.

Table 3.16. Detailed information of studies with an element of feasibility of measurement tools

Study details	Methods		Study protocol (free living/lab)	Results
	Measurement Tool(s) <i>Placement</i>	Aspect of feasibility		
Calorimetry (n=5)				
Physical activity (PA) and sedentary behaviour (SB) (n=4)				
Janssen et al. (2013c); Australia, mean age 5.3 years; n=40; 4-6 years; 22 male, 18 female	Whole room calorimetry	Practical utility of whole room calorimetry, assessed by compliance with the graded activity protocol, refusal to participate and data loss.	150 minute room calorimeter protocol (laboratory based)	74% of those eligible agreed to participate. 5% missing data due to calorimeter malfunction (n=2) and 25% did not comply with full protocol (n=10).
Oortwjin et al. (2009); Australia; N=5; mean age 5.2 years; 3 male, 2 female.	Whole room calorimetry	Acceptability, assessed by adherence to study protocol	150 minute room calorimeter protocol (laboratory based)	80% (n=4/5) children completed the protocol All able to follow instructions, none felt anxious, claustrophobic, or terminated the study early.
Janssen et al. (2015); Australia; N=40, 4-6 years; 22 male, 18 female	Whole room calorimetry	Reasons for drop out/exclusion	150 minute room calorimeter protocol (laboratory based)	5% exclusion due to calorimeter malfunction (n=2)
Janssen et al. (2013a); Australia N=40; 4-6 years, mean age 5.3 years; 22 male, 18 female.	Whole room calorimetry	Reasons for drop out/exclusion	150 minute room calorimeter protocol (laboratory based)	14% exclusion due to: illness (n=1), inability to schedule second visit (n=1), refusal to participate in activity protocol (n=2) and calorimeter malfunction (n=2)
PA (n=1)				

Steenbock et al. (2019); Germany; n=41; 3yrs – 6.3 years; 22 male, 19 female 100% Caucasian	Indirect calorimetry: MetaMax3b Doubly labelled water	Reasons for drop out/exclusion	75 minute protocol of both pre- determined and free play activities (both free living and structured)	34% exclusion due to: withdrawal consent on day of measurement (n=11), illness (n=5), did not fast prior to doubly labelled water (n=2), did not complete activity protocol (n=1) and calorimeter malfunction (n=2).
Accelerometers and calorimetry (n=3)				
PA and SB (n=1)				
Janssen et al. (2014); Australia; N=18; 4-6 years, mean age 5.2 years; 9 male, 9 female	Accelerometer: ActivPAL <i>Thigh</i> Whole room calorimetry	Reasons for drop out/exclusion	150 minute room calorimeter protocol (laboratory based)	10% exclusion due to: accelerometer failure (n=2) and calorimeter failure (n=2)
PA (n=2)				
Lopez-Alarcon et al. (2004); USA; N=29; 4-6 years; 17 male, 12 female 58% White 42% African American 6 of the children recruited from Head Start Center	Accelerometer- Actiwatch (AW16) <i>Right ankle</i> Doubly labelled water	Reasons for drop out/exclusion	7 days of habitual activity (free living)	7% exclusion due to: accelerometer failure (n=1) and no urine sample collected by parent (n=1)
Sijtsma et al. (2012); Netherlands; N=30; 3.1-4.1 years; 12 male, 18 female	Accelerometer: Direct life triaxial accelerometer (Tracmor_D)- <i>Middle of lower back</i>	Reasons for drop out/exclusion	3-5 days of habitual activity (free living)	17% exclusion due to: accelerometer battery problems (n=1) and no valid energy expenditure data (n=4).

	Doubly labelled water			
Direct observation (n=3)				
PA and SB (n=3)				
Puhl et al. (1990); USA; n=192, 3-4 year old, mean age not reported; sex not reported	Children's Activity Rating Scale (CARS)	Reactivity of the children to the presence of an observer (observer reported using brief form).	Children observed 1-4 times a year, for approx. 10-12 hours per day (observers alternating 2 hour shifts) (<i>free living</i>)	922 reports filed based on the 491 observations-no child reactivity problems reported in 93.3% of the reports. 153 (16.6%) reactive behaviours were reported where observers thought that the child's behaviour seemed modified for approx. 30 mins or less (38% of reactive reports), 3 hours or less (30%), half or full day (29%), and an undetermined amount of time (3%).
Sharma et al. (2011); USA; n= 67; 3-6 years, mean age 4.3 years; 27 male, 40 female 34.8% African American, 63.6% Hispanic, 1.5% Other (White/Native American) Children recruited from Head Start Centres	System for Observing Fitness Instruction Time for Preschoolers (SOFIT-P)	Feasibility of method determined by feedback from trained observers	1 day of protocol training, followed by two days of implementation and coding (<i>free living</i>)	Observers stated that SOFIT-P was relatively easy to use and feasible for both indoor and outdoor settings. Training was not overwhelming. No teachers/children refused to participate in measurement, reacted negatively or voiced any discomfort or unease with the observation.

Klesges et al. (1985); USA; n=30; 41-77 months; 10 male, 20 female	Fargo Activity Timesampling Survey (FATs)– Time sampling vs continuous observation.	Completion rate	Child observed for 1 hour during freeplay in nursery setting (free living)	100% parents contacted agreed to participate, 100% of these completed the protocol and had usable data.
Direct observation and accelerometers (n=3)				
PA and SB (n=1)				
Alhassan et al. (2017); USA; N=33; Mean age 4.4 years; 21 male, 12 female	Actiwatch (Spectrum) <i>Non dominant wrist</i> Actigraph (GT3X) <i>Waist, centre of lower back</i> Direct observation (Observational system recording physical activity in children – preschool version (OSRAC- P))	Reasons for drop out/exclusion	30 minutes of daily activity in pre-school (free living)	55% exclusion due to: Actiwatch reported as 'off- wrist' (n=11- 15%), Actigraph registered continuous zeros (n=15- 20%), participants absent from direct observation (n=15- 20%)
PA (n=1)				
Sharp et al. (2017); North Wales, UK; N=56; 3-4 years, mean age 3.7 years; 29 female, 27 male	FitBit (Zip) <i>Both right hip</i> Direct observation	Reasons for drop out/exclusion	5 min structured walking task (nursery setting)	15% exclusion due to: not performing task as requested (n=5) and technical errors in video recordings (n=5)

SB (n=1)				
Alghaeed et al. (2013); Scotland, UK; N=30; Mean age 4.1 years; 10 male, 20 female	ActivPAL <i>Thigh</i> Direct observation	Reasons for drop out/exclusion	1 hour during usual activity in nursery <i>(free living)</i>	6% exclusion due to not having complete datasets for accelerometer and direct observation (n=2)
HRM (n=1)				
PA and SB (n=1)				
Bar-Or et al. (1996); Canada; n=23; 3-5 years, mean age not reported; 11 female, 12 male	Polar Vantage XL Monitor <i>Chest and wrist watch</i>	Acceptability of device (researcher reported)	Child observed twice for 60-90 mins wearing heart rate monitor, researcher evaluated child's acceptance based on 6 point scale <i>(free living)</i> .	90% of children rated between enthusiastic and agreed. Decrease by 4% in refusal ratings by observation 2.
Accelerometers (n=20)				
PA and SB (n=13)				
Bacardi-Gascón et al. (2012); Mexico; N=35; 3-5 years; 17 male, 18 female	Actigraph (GT1M) <i>Right hip</i>	Reasons for drop out/exclusion	4 days habitual activity-accelerometer used as comparison method against proxy report tool <i>(free living)</i>	22% exclusion due to not wearing accelerometer for sufficient amount of time (n=10)
Dobell et al. (2019); UK; N=62; 3-4 years, mean age 3.5 years; 36 males, 30 females (sex of participants prior to exclusion of 4 data points)	Actigraph (GT3X) <i>Right hip and non-dominant wrist</i>	Reasons for drop out/exclusion	60 minutes of activity in pre-school setting <i>(free living)</i>	6% exclusion due to: accelerometer failure (wrist worn) (n=1) and child withdrawal while wearing accelerometers (n=3)

Hands & Larkin (2006); Australia; N=23; 5-6 years; 12 male, 11 female	Actigraph (MTI AM7164) <i>Waist</i>	Reasons for drop out/exclusion	Children monitored for 30 minutes over 5 days during free playing in their school setting (free living)	4% exclusion due to child spending majority of the time on a swing which inflated the accelerometer count (n=1)
Sirard et al. (2005); USA; N=269; 3-5 years, mean age not reported; 125 male, 144 female 3yr; 37.7% white 4yr; 25.6% white 5yr; 38.7% white	Actigraph (MTI) <i>Right hip</i>	Reasons for drop out/exclusion	Children wore Actigraph for the entire time they were at school for up to 10 consecutive weekdays, direct observation for 1-3 hours of activity in preschool (free living)	4% exclusion due to incomplete or missing data (n=12)
Dwyer et al. (2011); Australia; N= 67; 3-5 years; 36 male, 31 female 91% White, 3% Mediterranean, 6% other ethnicity. 9% low SES, 34% middle, 57% high	Actigraph (MTI 7164) <i>Right hip</i>	Reasons for drop out/exclusion	3 day activity (<i>free living</i>)	36% exclusion due to: not wearing accelerometer for sufficient amount of time (n=38)
Wen et al. (2010); Australia; n=31 ; 3-5 years; 19 male, 12 female	Accelerometer- Actigraph (GT1M) <i>Right hip</i>	Did not complete measures for required period of time.	7 day accelerometer wearing protocol (free living)	30% exclusion due less than 4 days accelerometer data (n=13).

Children recruited from centers said to vary in SES.				
Martin et al. (2011); Scotland, UK; N=23; 3-5 years; 9 male, 14 female	Actigraph (GT1M and GT3X) <i>Right hip</i> ActivPAL <i>Thigh</i>	Reasons for drop out/exclusion	Monitors worn simultaneously for a 7 consecutive days during waking hours, except during water based activities (free living activity)	28% exclusion due to : lost/missing or malfunctioning accelerometer (n=2), non-compliance with study protocol (n=6) or illness (n=1)
Davies et al. (2012); UK; n=20; 3.2-4.9 years, mean age 4.4 years; 6 male, 14 female	ActivPAL <i>Thigh</i>	Acceptability of devices (parent reported); reasons for missing or invalid data	Children wore activPAL monitor for up to 7 days. Parents completed 10 item questionnaire to assess acceptability of devices on 5 point Likert scale (free living).	Acceptability: Support for practicality of using activPAL in preschool children. One parent reported that they stopped using the overlying Tegaderm dressing because it was uncomfortable for their child. Missing or invalid data: 206 hours of data loss over the week, equivalent to 10.3 hours of missing data per participant per week.
Janssen et al. (2013b); Australia; n=38; 4-6 years, mean age 5.3 years; 20 male, 18 female	ActivPAL <i>Thigh</i>	Reasons for drop out/exclusion	150 minute structured activity protocol (laboratory based)	5% exclusion due to accelerometer failure (n=2)
Vanderloo et al. (2016); Canada; N=23; 4-5 years; 12 male, 16 female	Actical <i>Right hip</i> Actigraph (GT3X+) <i>Right hip</i>	Reasons for drop out/exclusion	Both monitors worn simultaneously for 7 consecutive days (free living)	18% exclusion due to not wearing accelerometer for required amount of time (n=5)
Ettienne et al. (2016); USA;	Actical	Reasons for drop out/exclusion	7 days habitual activity (free living)	10% exclusion due to not completing both phases of work (39/49) and/or having at least one day of

n=30; Mean age =3.5 years; 17 male, 13 female 46% Native Hawaiian, 14% Other Pacific Islander- All of mixed ethnicities. Children recruited from Head Start sites	<i>Non dominant wrist</i>			data deleted due to the total number of mins of SB being greater than 1300 minutes per day (23/49)
Kelly et al. (2004); UK; n=78; 3-4 year olds, mean age 3.5 years; 30 male, 48 female	Actiwatch (AW16) <i>Right hip</i> Actigraph (CSA/MTI) <i>Right hip</i>	Reasons for drop out/exclusion	39-45 minutes structured play class <i>(usual activity, reflective of free living)</i>	7% exclusion due to technical failures and lack of cooperation (n=6)
Shin (2015); USA; N=19; 3-5 years; 10 male, 9 female	BFF <i>Left and right wrist, back, waist, chest, ankle, upper arm</i> Actigraph (GT3X/GT3X+) <i>Waist and Right Wrist</i>	Acceptability of device	Children followed structured activity protocol at the pre- school under researchers instruction – participant feedback used to determine feasibility <i>(usual activity, not free living)</i>	Acceptability: According to their feedback, participants agreed to accept the size and features of devices in use. However, the comfort level of the devices was in question- participants sometimes requested to stop testing when they felt uncomfortable wearing devices on their chests, especially when they slipped. Many participants strongly opposed attaching the devices on their chests directly, but accepted them attached over clothes. (Also reported low response rate – 36% as part of feasibility aspect)
PA (n=4)				

Nyström et al. (2017); Sweden; N=40; 5.2- 5.7 years, mean age 5.5 years; 22 male, 18 female	Actigraph (wGT3X-BT) <i>Non-dominant wrist 1 sec epoch No cut points reported for PA</i> <i>Non-wear time: Determined using raw acceleration data. SD of each axis calculated over a 30-min period. If SD of acceleration of any two axes less than 0.002 g for the same window= non- wear time.</i> <i>Valid day based on wear time: When the awake wearing time (non-sleeping data) was ≥600 minutes</i> <i>Number of days: Minimum of 3</i>	Compliance with study protocol	7 days activity (free living)	100% compliance with activity protocol and accelerometer wearing; Thirty-four children (85%) wore the ActiGraph for 7 days, 4 children (10%) for 6 days, 1 child (2.5%) for 5 days and 1 child (2.5%) for 4 days

	<i>days of valid data.</i>			
Corder et al. (2009); UK; N=27; 4-5 years; 17 male, 10 female	Actigraph (MTI 7164)	Reasons for drop out/exclusion	Questionnaire for past 7 days, remaining protocol for 6-10 days assessing habitual activity (free living)	7% participants had accelerometer malfunction (n=2)
Ellison et al. (1992); USA; N=100; 3-5 years, mean age and sex not reported Predominantly white, middle class families.	Caltrac accelerometer <i>Right hip</i>	Reasons for lack of usable data; Cost of devices (including replacement for malfunctions)	Children asked to wear accelerometer for 5 consecutive days, at 2 time points, 6 months apart (free living).	82% of schedule days were available, all families provided usable data. Looking at a random sub-set of 500 reports, it was identified that on days when data were not obtained: 52% of the time it was due to the parent forgetting to attach/refusal from child to wear the device in the morning; 28% of the time information parent record was insufficient in regards to when monitor was attached/removed; 20% of the time it was due to instrument malfunction. With overall costs for devices, batteries and repairs, the average cost per day of successful monitoring = £1.31.
Chen et al. (2002); Japan; n=21; 3-4 years; 12 male, 9 female	Accelerometer- Actiwatch-L Caloriecounter	Reasons for drop out/exclusion	3 days of accelerometer wearing (free living activity)	8% exclusion due to measurement failure from instruments (n=2)
SB (n=3)				
De Decker et al. (2013); Belgium;	ActivPAL <i>Thigh</i>	Practical utility of the accelerometers (parent	Monitors worn for at least 3 days simultaneously,	Practical utility: 71% (n=28) parents indicated on a 5 point scale that their child found it pleasant to wear both monitors.

n=52; mean age 5.5 years; 26 male, 26 female	Actigraph (GT1M) <i>Right hip</i>	reported); Reasons for drop out/exclusion.	direct observation for 1 hour during pre-school hours. Parents asked to indicate on 5 point scale whether their child found it pleasant or unpleasant to wear the devices (free living)	4 parents indicated that their child found it very unpleasant to wear the ActivPAL during measurements; 38% (n=15) reported that their child had some skin irritation to on the thigh due to wearing the waterproof ActivPAL for 5 consecutive days. Reasons for exclusion/drop out: 15% children did not wear ActivPAL (n=8), 13% did not have adequate data from both devices for at least 3 days (n=7).
Van Cauwenberghe et al. (2012); New Zealand; N=49; 3-4 years; 22 male, 19 female	Actical <i>Waist, above the right iliac crest</i> ActivPAL <i>Thigh</i>	Acceptability of devices (researcher reported); Reasons for drop out/exclusion	Monitors worn simultaneously during 1 day at nursery. Researchers coded children's acceptability to wearing monitors on 5 point scale (free living)	Acceptability: 81.6% of children had high or very high acceptability of wearing both monitors, very low acceptability was not reported for any children. Reasons for drop out/exclusion: 36% excluded data due to: did not attend the centre on the day of measurement (n=8); did not assent to wear one or both monitors (n = 7); technical failure when downloading the raw activPAL data (n = 5); monitors not worn simultaneously (n= 3); monitors removed during the measurement period (n = 2); and monitors worn for less than 3 h (n = 2).
Mendoza et al. (2013); USA; N=96; 3-5 years; 41 female, 53 male, 2 missing 100% Latino or Hispanic.	Accelerometer- Actigraph (GT1M) <i>Hip</i>	Compliance; Retention rate	Measurement tools administered for 7 days at 2 time periods (free living)	100% compliance at time 1 and 82.3% at time 2. Retention rate- 85.4%

Children recruited from Head Start Centers				
Pedometers (n=3)				
PA (n=3)				
Murray (2009); USA; N=75; 3-5 years; 21 male, 44 female 34.3% African American; 64.4% Hispanic; 1.4% Native American Children recruited from Head Start Centres	MVP 4 Walk4Life Digital Pedometer <i>Right or left hip</i>	Process evaluation from data collectors and teacher input, to determine feasibility.	Children were observed in pre-school whilst wearing pedometer at 2 time points (free living).	Reports of inability to wear the pedometer for the full day – failing off, restroom breaks, loose pants, nap time, play time. At pre-test, measurement 1 – 45% (n=34) of the children did not achieve sufficient wear time and at post-test, measurement 1 49% (n=37) did not achieve sufficient wear time. Lacked feasibility due to logistical barriers in design.
Chow & Au (2009); Hong Kong n=23; 5-6 years; 12 male, 11 female Majority of the children from lower to lower-middle families.	Yamax Digiwalker (SW-700) <i>Right side of waist</i>	Reasons for exclusion/drop out	Pedometer worn for 4 days, as comparison method for proxy report (free living)	23% (n=7) missing data as parents forgot to complete pedometer record.
Louie & Chan (2003); Hong Kong; N=145; 3-5 years, mean age 4.2 years; 84 male, 61 female	Yamax Digiwalker (SW-200) <i>Right hip (some invited to wear on both right and left hip)</i>	Reasons for exclusion/drop out	25 minute physical activity class in normal school day (free living)	2% exclusion of data due to pedometer battery problems (n=3)
Accelerometers and pedometers (n=3)				
PA and SB (n=1)				

Pagels et al. (2011); Sweden/USA; N=55; 3.4-5.7 years; 28 male, 27 female	Accelerometer- Actigraph (GT1M) <i>Waist belt</i> Pedometer – Yamax SW-200 <i>Right side of waist</i>	Reasons for exclusion/drop out	Children wore monitors simultaneously for 5 days during nursery time (<i>free living</i>)	5% (n=3) excluded due to incomplete data.
PA(n=2)				
Cardon & De Bourdeaudhuij (2007); Belgium; N=76; 4-5.9 years; 37 male, 39 female 55% from high SES; 45% from low SES	Accelerometer- Actigraph (MTI AM 7164) <i>Right hip</i> Pedometer- Yamax Digiwalker (SW- 200) <i>Right hip</i>	Receptivity to measurement tools (parent reported); Reasons for drop out/exclusion	Monitors worn simultaneously for waking hours up to 5 days. Parents asked to report child's receptivity to wearing the instruments on 5 point scale (<i>free living</i>).	Receptivity/Acceptability: 83% of children found it very pleasant (59,46%) or pleasant (50 children, 38%) to wear accelerometer or accelerometer and pedometer; 8% (11 children) found it not pleasant or not unpleasant, 2% (2 children) found it unpleasant Receptivity of wearing the instruments was higher in children concurrently wearing the accelerometer and the pedometer (1.7 ± 0.8) than in children wearing only the pedometer (1.4 ± 0.5 ; $t = 2.6$, $p = 0.01$). Reasons for drop out/exclusion: 5% excluded: Incomplete data due to not wearing pedometer more than 1hr on 1 or more days (n=7), pedometer lost (n=1), monitor damaged (n=1). 15% excluded due to satisfactory measurements not obtained from accelerometer (n=13)
De Craemer et al. (2015); Belgium; N=41; 4-6 years;	Actigraph (GT1M) <i>Left hip</i>	Reasons for exclusion/drop out	Children wore monitors simultaneously for	44% missing data (n=23), as children did not have valid data for both measurement devices

21 male, 20 female	Pedometer - Omron Walking Style Pro (HJ- 720IT-E2) <i>Right hip</i>		at least 1 day (<i>free living</i>)	

Abbreviations in this table: CARS= Children's Activity Rating Scale; SOFIT-P= System for Observing Fitness Instruction Time for Preschoolers; OSRAC-P= Observation System for Recording Physical Activity in Children- Preschool; USA=United States of America; UK=United Kingdom

Table 3.17. Summary table of the reliability and feasibility of the measurement tools

Method	Reliability	Feasibility	Reference
Calorimetry			Janssen et al., 2015; Janssen et al., 2013a; Janssen et al., 2014; Janssen et al., 2013c; Lopez-Alarcon et al., 2004; Oortwijn, 2009; Sijtsma et al., 2013; Steenbock et al., 2019
Direct observation			Alghaeed et al., 2013; Alhassan et al., 2017; Klesges, Woolfrey and Vollmer, 1985; Puhl et al., 1990; Sharma et al., 2011; Sharp et al., 2017
Heart rate monitors			
Polar Vantage XL Monitor			Bar-Or et al., 1996
Accelerometers			
Fitbit (Zip)			Sharp et al., 2017
Triaxial Research Tracker 3 (RT3)			Bikchu, 2014
Actigraph (GT3X/+)			Alhassan et al., 2017; Dobell et al., 2019; Martin et al., 2011; Shin, 2015; Vanderloo et al., 2016
Actigraph (CSA/MTI)			Cardon and De Bourdeaudhuij, 2007; Corder et al., 2009; Dwyer et al., 2011; Hands and Larkin, 2006; Kelly et al., 2004; Sirard et al., 2005
Actigraph (wGT3X-BT)			Delisle Nyström et al., 2017
Actigraph (GT1M)			Bacardi-Gascón et al., 2012; De Craemer et al., 2015; De Decker et

			al., 2013; Martin et al., 2011; Mendoza et al., 2013; Pagels, Boldemann and Raustorp, 2011; Wen et al., 2010
ActivPAL		●	Alghaeed et al., 2013 Davies et al., 2012; De Decker et al., 2013; Janssen et al., 2013b; Janssen et al., 2014; Martin et al., 2011; Van Cauwenberghe et al., 2012
Actical		●	Ettienne et al., 2016; Van Cauwenberghe et al., 2012; Vanderloo et al., 2016
Actiwatch (AW16)		●	Kelly et al., 2004; Lopez-Alarcon et al., 2004
Actiwatch (Spectrum)		●	Alhassan et al., 2017
Actiwatch-L		●	Chen et al., 2002
Caloriecounter		●	Chen et al., 2002
Tracmor _D		●	Sijtsma et al., 2013
Caltrac		●	Ellison et al., 1992
Best Fit Friend		●	Shin, 2015
Pedometers			

Yamax (SW-700)			Bikchu, 2014; Chow and Au, 2009
Yamax Digiwalker (SW-200)			Cardon and De Bourdeaudhuij, 2007; Louie and Chan, 2003; Pagels, Boldemann and Raustorp, 2011
MVP 4 Walk4Life Digital			Murray, 2009
Omron Walking Style Pro (HJ-720IT-E2)			De Craemer et al., 2015
Proxy reported measurement tools			
Children's Leisure Activities Study Survey (CLASS)			Telford et al., 2004
Questionnaire developed for parents of pre-schoolers in Mexico			Bacardi-Gascón et al., 2012
Teacher activity rating			Manios, Kafatos and Markakis, 1998
TV Diary			Mendoza et al., 2013
Leisure time report			Manios, Kafatos and Markakis, 1998
Pre School Physical Activity Questionnaire (PRE-PAQ)			Dwyer et al., 2011
Netherland's Physical Activity Questionnaire (NPAQ)			Janz, Broffitt and Levy, 2005
'Toybox' Primary Caregivers Questionnaire			González-Gil et al., 2014

Children's physical activity questionnaire (CPAQ)			Corder et al., 2009

***Methodology used to assess the ability of the tool is detailed in the methods above and is indicated in the summary table as:**

Good =  Moderate =  Weak = 

Key for colour of boxes:  = evidence from ≥ 3 studies

 = evidence from < 3 studies

This table shows a summary of the results of studies where they tested the reliability or feasibility of the measurement tool. The summary ratings were based on the quality of the tools for the specific measurement property. Where the measurement tool was deemed 'good' in the majority of the studies, the summary assessment was deemed 'good'. Where the measurement tool was deemed 'moderate' in the majority of the studies, the summary assessment was deemed 'moderate'. Where the measurement tool was deemed 'weak' in the majority of the studies, the summary assessment was deemed 'weak'. In instances where the measurement tool had mixed evidence in the studies, such as studies with outcomes of 'weak' and 'moderate', or 'moderate' and 'good', the overall assessment was deemed to be the most positive of the two outcomes. All tools of reasonable quality where any evidence was available are included in this table, including where only one or two studies reported that result.

3.5.4.9 Generalisability of results and health inequality implications

The majority of the studies reported the age (n=74) and sex (n=69) of the included children. However, only ten studies (Dwyer et al., 2011; Ellison et al., 1992; Ettienne et al., 2016; Janz, Broffitt and Levy, 2005; Lopez-Alarcon et al., 2004; Mendoza et al., 2013; Murray, 2009; Noland et al., 1990; Sharma et al., 2011) reported other key attributes that help determine the generalisability of the results to the wider population, including: ethnic origin and SES. Therefore, there is insufficient evidence to suggest that results from the individual studies are generalisable across other populations.

3.5.4.10 Ethnicity

Eighteen studies described the ethnicity of the sample in which the measurement property of the tool was examined. The majority of the studies that reported on ethnicity had samples with children whom were primarily white or Caucasian (Adolph et al., 2012; Djfarian et al., 2013; Dwyer et al., 2011; Ellison et al., 1992; Finn and Specker, 2000; Janz, Broffitt and Levy, 2005; Kahan, Nicaise and Reuben, 2013; Lopez-Alarcon et al., 2004; Noland et al., 1990; Sirard et al., 2005; Steenbock et al., 2019). Three of the studies reported primarily Hispanic populations samples (Mendoza et al., 2013; Murray, 2009; Sharma et al., 2011), followed by primarily African American samples (Pate et al., 2006; Pfeiffer et al., 2006). There was no indication of lower measurement properties of the tools in any specific ethnic group in these studies, however, this was not examined directly.

3.5.4.11 Socioeconomic status

Fourteen studies reported on the SES of their sample. Some studies reported that at least some of the participants were recruited from pre-schools characterised to have

individuals with lower SES, such as Head Start Centres, which require proof of income to demonstrate that families are at or below poverty level (Ettienne et al., 2016; González-Gil et al., 2014; Lopez-Alarcon et al., 2004; Mendoza et al., 2013; Murray, 2009; Sharma et al., 2011; Wen et al., 2010). The remaining studies that reported on SES were based on individual level demographics. These studies reported that the samples were primarily made up of individuals of higher SES (Dwyer et al., 2011; Ellison et al., 1992; Janz, Broffitt and Levy, 2005; Noland et al., 1990). Only one study reported that the children in their sample were from lower to lower-middle SES families (Chow and Au, 2009). Whilst one study also reported an equal amount of participants from both high and low SES families (Cardon and De Bourdeaudhuij, 2007). None of the studies directly examined whether SES affected validity or reliability, or whether there was reduced feasibility of the tools in different SES groups. There was no evidence to suggest that SES was affecting the validity or reliability of the measurement tools being evaluated. The majority of the studies found no indication of reduced feasibility amongst the different SES, however, two studies conducted with participants of lower SES reported a lack of feasibility when using pedometer-based measurement tools (Chow and Au, 2009; Murray, 2009).

3.6 Discussion

This systematic review identified 69 articles, describing 75 studies that examined the measurement properties of measurement tools used to assess PA and/or SB in pre-school aged children. In this review, I have provided an overview on what measurement tools have been examined for what outcome measures, with an indication on whether these have been shown to be valid, reliable, or feasible.

The heterogeneity of the studies included in this review emphasises the complexity of measurement of PA and SB, identified previously by others (Rowe, 2011; Thompson et al., 2015), alongside the additional challenges associated with measurement in a pre-school aged population (Cliff, Reilly and Okely, 2009; Welk, Corbin and Dale, 2000). The review demonstrates that different measurement tools often examine different dimensions of PA and SB (e.g. time spent in different intensities of activity, posture allocation, step count, energy expenditure) in line with previous literature (Kelly, Fitzsimons and Baker, 2016; Trost, 2007). Measurement tools all have a specific use depending on their derived purpose, desired measurement outcomes and the context in which the tool is being used. However, when selecting an appropriate and useful measurement tool to assess PA and SB amongst children, there is often a trade-off between the three main utilities (validity, reliability, feasibility), alongside further considerations, such as the sample size of the study, budget, and availability of resources (Trost, 2007). As such, it is not a question of which measurement tool is 'best' for assessing the PA and SB of pre-school aged children, but rather, what measure or combination of measures, are most appropriate in the given context for the desired outcome measures (Kowalski et al., 2012; Rennie and Wareham, 1998; Rowe, 2011; Troiano, 2009; Trost, 2007). Ultimately, any study assessing these behaviours needs to choose a measurement tool, and as such, recommendations have been made from the findings of this review, that may be of interest to those involved in research based on the frequently used PA and SB outcomes.

Overall, based on the current evidence, which included a limited number of studies of varying quality, of the measurement properties of measurement tools used to examine PA and SB of pre-school aged children, multiple accelerometers, including the Actigraph (in particular GT3X versions), Actical, ActivPAL and Fitbits (Flex and Zip), can provide valid

measures, with some evidence of feasibility, of movement behaviours that would be of interest in a range of research where resources and capacity allow. However, a limited number of devices were examined for reliability; the Fitbit Zip showed excellent intra-instrument reliability, but this was based on one study only (Sharp et al., 2017). More research studies examining the intra-instrument reliability of measurement tools in pre-school aged children are warranted.

There was a discrepancy in the amount of studies examining each measurement tool; with an emphasis placed on examining the measurement properties of device based measurement tools, primarily accelerometers, making up 70% of all studies, with multiple studies examining the same type of accelerometer. Accelerometers were also frequently used as a comparison measurement tool. Continuous technological advances and development of new and improved activity monitoring devices is a disadvantage of such tools that increases difficulty for standardisation within and between studies (Bauman, Pedišić and Bragg, 2016; Lee and Shiroma, 2014). Alongside this, a major limitation of the included studies utilising accelerometers was the considerable variation in the interpretation of data, due to differences in the subjective decisions on epochs, cut points, placement, wear time, non-wear time and valid number of days (Cain et al., 2013; Cliff, Reilly and Okely, 2009; Janssen and Cliff, 2015; Migueles et al., 2017; Oliver, Schofield and Kolt, 2007). At least sixteen different published cut points were used in the included studies (Adolph et al., 2012; Chu, McManus and Yu, 2007; Ekblom et al., 2012; Ekelund et al., 2001; Evenson et al., 2008; Freedson, Pober and Janz, 2005; Johansson et al., 2015; Pate et al., 2006; Pfeiffer et al., 2006; Puyau, Vohra and Butte, 2002; Reilly et al., 2003; Rowlands et al., 2004; Schaefer, Nace and Browning, 2014; Sirard et al., 2005; Sun, Schmidt and Teo-Koh, 2008; Vanhelst et al., 2011), with some studies applying their own. Such methodological

inconsistencies can be problematic, as applying different cut points to the same data can result in statistically and biologically significant differences in the outcomes of PA (Banda et al., 2016; Ojiambo et al., 2011; Van Cauwenberghe et al., 2011b). The published cut points used for each study are provided in the summary tables (Tables 3.7, 3.9, 3.11, 3.13), but the overall aims of the studies included in this review were not to show which data interpretation choices were most valid, and so the outcomes of this review cannot determine which data interpretation choices are 'best' to use. Therefore, it is important to note that the devices shown to be valid can only be said to be valid with the data interpretation choices in which they were tested. If readers of this work decide a device based measurement tool is best for their research study, it is important to ensure use of the respective epochs, cut points and monitor placement outlined in the validation studies. This information is provided in each study in the tables outlined in this chapter (Tables 3.7, 3.9, 3.11, 3.13).

Proxy reported measurement tools, in particular the PRE-PAQ and *leisure time report*, also show some promise in terms of validity and test-retest reliability, for some dimensions of PA and SB. Although, the evidence quality is weak, therefore, much more evaluation of the measurement properties of these types of tools is needed. These measurement tools have advantage in terms of identifying contextual behaviour, cost, and can be consistently used in a standardised way, as they are not reliant on upgrades in technology (Hallal et al., 2012). With this, content validity has been suggested to be an essential first step in determining the appropriateness of a self/proxy reported measure, to determine if the measurement tool is assessing what it intends to measure and to identify whether the measure is understandable to the target population (Mokkink et al., 2010b; Terwee et al., 2010; Terwee et al., 2018a). However, only two of the included studies

reported on an element of face and content validity of their measures (Bacardi-Gascón et al., 2012; Dwyer et al., 2011).

Similarly, very few studies explicitly examined the feasibility, including acceptability, of the measurement tools (Bar-Or et al., 1996; Cardon and De Bourdeaudhuij, 2007; Davies et al., 2012; De Decker et al., 2013; Ellison et al., 1992; Janssen et al., 2013c; Mendoza et al., 2013; Murray, 2009; Oortwijn et al., 2009; Puhl et al., 1990; Sharma et al., 2011; Shin, 2015; Van Cauwenberghe et al., 2012), although, a proportion of the studies did report exclusion of participants and missing data. The measurement properties of measurement tools may be comprised if they are not feasible in the context and with the population in which they are to be used. The ability to identify reasons for missing data, reasons for non-compliance and overall acceptability and feasibility of the measurement tools, to determine end user experience, will help with understanding which tools are most applicable for use (Evenson, Goto and Furberg, 2015). The included studies often evaluated the tools over a short space of time, usually around one hour, due to use of an intensive reference method (such as whole room calorimetry or direct observation). The feasibility of the devices for longer term measurement cannot be determined from this.

Whole room calorimetry and direct observation were shown to be accepted and feasible methods for measuring PA and SB of pre-school aged children. However, this was based on a very small number of studies, small sample sizes, in controlled environments. As these methods are expensive, highly burdensome and intensive, require extensive training, and can only capture a short period of time, they are only feasible for smaller scale projects and for only some dimensions of PA and SB. Whole room calorimetry and direct observation would not be feasible for use at scale or with large population samples. In line with this, the

measurement properties of the tools are only reflective of the context in which they were tested, and cannot necessarily be generalised to other contexts. For example, these tools were shown to be feasible in laboratory/controlled settings over a short period of time, and as such the methods can only be said to be feasible in this context, and the assessment does not apply to free living longer term measurement.

The majority of studies included in this review were conducted in a free living (habitual) context, including at the pre-school and/or at home. However, the results of the studies were often lacking in potential for generalisability and translation. As is the case for much research, identified studies have traditionally been conducted in high income countries. The review demonstrated that a very small proportion of studies examining measurement properties of tools used to examine PA and SB of pre-school children included participants from low SES backgrounds (12% studies) and ethnic minority group (7% included studies) representation. With this, there was no available evidence to suggest differences in the measurement properties of the measurement tools across different ethnicities or SES, however, a very small proportion of the studies explored these factors, and none directly. Similarly, there was minimal evidence of the measurement properties of the tools being evaluated with large sample sizes, so the capability of the tools at scale are unknown.

A major limitation in the field of PA and SB measurement is the lack of criterion methods for the majority outcomes, and indeed a lack of consensus on what might be considered a criterion. The level of evidence scheme was employed in this review to distinguish between differing levels of validity of the comparison measurement tools, but

caution should always be applied when validity is measured against non-criterion methods, especially when the validity of these methods themselves are not well established.

3.6.1 Strengths and limitations of the review, including limitations on the methodology used

A primary strength of this review is that international literature for evidence on this topic was assessed, which included extensive searching of multiple platforms; published academic research articles through database searching, searches of the grey literature, and manual searches. Another strength of this review is that various measurement properties of measurement tools were examined; including the feasibility of the measures, a measurement property that is usually not included in the scope of reviews of measurement properties (e.g. Adamo et al., 2009; Helmerhorst et al., 2012; Kowalski et al., 2012).

A potential weakness of this review was the use of the Downs and Black risk of bias assessment (Downs and Black, 1998). Although through extensive searching this was seen to be the best available tool, that has also been used in similar systematic reviews previously (Adamo et al., 2009; Prince et al., 2008), it was not ideal because it was not devised specifically for studies examining the measurement properties of measurement tools. However, an additional risk of bias assessment on proxy reported measurement tools was conducted due to a tool being available for these types of studies. Further, although two reviewers screened a 10% sample of the title and abstracts with very high agreement, only one reviewer screened the remaining studies and so it is plausible that studies were missed through human error. It is also plausible that studies where examining the utility of the measurement tool was not the primary aim were not identified for inclusion in the review. However, finding such studies where the information on assessment of measurement

properties of the tools is lacking in the key information (e.g. title and abstracts) is a limitation of reviews of this kind (e.g. Hidding et al., 2017).

A further limitation of the research is that we did not explore the measurement property of 'responsiveness' in our search terms. Responsiveness, defined as '*The ability of a patient reported outcome measure to detect change over time in the construct to be measured*' (Terwee et al., 2016) is a measurement property that is not well established in the field of PA and SB measurement. This is apparent from the literature (Hidding et al., 2017; Hidding et al., 2018) and when conducting initial scoping searches for this review, no studies exploring responsiveness were identified. No studies were identified, also, through our additional hand searches. It is, therefore, currently not possible to determine from this review whether any tools are able to detect changes in behaviour over time (e.g. in response to an intervention) and it is essential that this is the focus of future research.

Although there was a broad age range for inclusion in this review (studies with children aged 3-7 years), it is plausible that the behaviours that children engage in vary between the ages of 3 and 7 due to different stages of development, and in some cases differences in routine (e.g. nursery through to starting reception). However, this broad age range was used to ensure international relevance of the present study and to ensure that studies across the globe were detected where the age of 'pre-school' child may include older children than the classification used in the UK (of children aged 3-4 years old).

Systematic reviews are classified as the optimum review method when collating and synthesising evidence on a given topic, however, a limitation of the current review was the heterogeneity of included studies meaning that it was not possible to conduct a meta-analysis. Meta analyses are thought to provide an advantage over narrative synthesis

approaches due to the ability to combine statistics from the results of separate studies to provide more precise overall effects in answering research questions (Deeks, Higgins and Altman, 2022).

3.6.2 Implications and recommendations for future research

This review presents an overview of the utility of the measurement tools used to assess PA and SB in the pre-school population, based on the currently available evidence by providing an indication on how valid, reliable or feasible the range of measurement tools have been shown to be for what outcome measures. The main implication of this review is that the work can be a useful resource for researchers, practitioners, and others wanting to assess these behaviours of pre-school children, when selecting a measurement tool for use, alongside acting as a criterion for ensuring confidence in the outcomes of studies examining these behaviours. Further, this review can be used by policy makers for recommendations for the surveillance of physical activity and sedentary behaviour, both in the UK and internationally.

The lack of multiple studies examining the same measurement tool makes it difficult to produce firm conclusions on the measurement properties of some tools. Differences in the number of studies examining each measurement tool (e.g. multiple examining certain types of accelerometer, with only one study examining each proxy reported tool), the ways in which the studies were conducted, the differences in comparison measurement tools used, and reported outcome measures can make direct comparisons between the results of the studies difficult. Future research in this area should, where possible, assess the utility of multiple methods simultaneously to help make comparisons that are more meaningful; only where the purpose of the tools are to examine the same dimensions of PA or SB. When

measurement tools are multi-dimensional, in that their purpose is to examine varying dimensions of PA and SB (e.g. frequency, duration), they should be assessed against varying comparison measures, dependent on what will constitute a reasonable reflection of the validity or reliability of the tool. However, this would require more extensive funding to purchase varying tools.

Future research should focus on the measurement properties of measurement tools being examined in different populations to ensure external validity of the measures and to extend the generalisability of the findings from these types of studies. This should include using large sample sizes, individuals from varying SES backgrounds, different ethnicities, and conducted in lower income countries; this would help to evaluate the measurement properties, including feasibility and acceptability, of the measures in different contexts. However, only after a measurement tool is shown to be valid, reliable, and feasible in the population for which it was originally developed (Terwee et al., 2016).

In addition to this, research projects examining the PA or SB of pre-school aged children should look to identify the context in which the measurement properties of tools have been examined prior to choosing which tool may be most appropriate for their purpose. If the measurement properties of the selected tool are unknown in the context in which they are to be used, researchers should aim to conduct a validation check even if just on a sub-sample of children included in the study. This would be useful to identify whether the measurement tools are working as intended and to ensure greater confidence in the results of the study.

This review highlights the importance of reporting on internal validity, such as missing data and non-completion within studies. Studies that reported such data revealed

various important implications of using measurement tools, including that malfunctions of measurement tools, primarily with device based tools and calorimetry, can have sufficient impact on the included final sample (Alhassan et al., 2017; De Craemer et al., 2015; Dwyer et al., 2011). Examining and reporting missing data is important, as these considerations can substantially impact on the utility of a tool (Dössegger et al., 2014).

Further qualitative work in this area is needed, with two main purposes: 1) to determine the feasibility and acceptability of measurement tools and 2) to examine face and content validity. In some instances in this review, an indication of feasibility of the tools was provided based on numerical scores due to this being the only available feasibility data within the studies. Previous research has highlighted the importance of conducting qualitative research to examine the feasibility of measurement tools in the target population prior to use (Costa et al., 2013; Hidding et al., 2020; Kowalski et al., 2012); with such work being important to understand recruitment bias and reasons for missing data or non-completion (Dössegger et al., 2014; Evenson, Goto and Furberg, 2015). Additionally, qualitative work to determine validity is rarely conducted, demonstrated by only two studies included in the current review that commented on these aspects (Bacardi-Gascón et al., 2012; Dwyer et al., 2011). Face and content validity conducted using qualitative methods have been highlighted as an essential first step in examining the validity of measurement tools for PA and SB, to determine whether tools are valid for their intended purpose (Hidding et al., 2017; Kelly, Fitzsimons and Baker, 2016). In line with this, there is a future research need to focus on further development and evaluation of proxy report methods together with the target population and ensure representativeness and feasibility of the measurement tool in the context in which it is intended to be used.

3.7 Conclusion

The measurement tools used to measure PA and SB in pre-school aged children show mixed measurement properties, and were generally based on minimal studies providing variable quality of evidence. Based on currently available evidence, the Actigraph (in particular GT3X versions), Actical and ActivPAL, have the greatest measurement properties for assessing common movement related outcomes (e.g. SB, MVPA, total PA) for free living activity of pre-school aged children, and should be the tool of choice where resources allow and where logistically possible. The Fitbit (Flex and Zip) also shows very promising results; however, these were based on a very limited sample of studies. Where measurement of a large sample is required and where budgets are limited, some proxy reported measures including the Pre-PAQ, *Netherland's physical activity questionnaire* and *nursery teachers report* can provide some valid data particularly for distinguishing between different levels of activity, alongside useful contextual information not captured by device-based measurement tools. The review demonstrated that there is insufficient evidence to conclude which tools are most suitable for low SES and ethnic minority groups. The review highlighted a clear need for further and more in-depth evaluation work, particularly of proxy reported measurement tools. Additionally, there is a need to determine the qualitative acceptability and feasibility of measurement tools used for free living activity measurement.

3.8 Summary and link to next chapter

The findings of this review provide us with information on which measurement tools have reasonable measurement properties for examining the PA and SB of pre-school children. However, as I continued with my research, it became apparent that 'sleep' completed the picture, which was particularly evident through the inclusion of sleep in the

latest WHO guidelines (as described in section 2.6.2.1, p.63). The review presented in this chapter was already underway with a published protocol, however, it was also evident that tools used to assess sleep can vary from those used in PA and SB research. Whilst it was important to identify measurement properties of available tools used to assess sleep, it was felt that a separate review was an appropriate next step of the research. As such, this is the focus of the next part of this PhD, *Chapter 4*, which was to review the evidence base and determine the validity, reliability, and feasibility of measurement tools used to assess the sleep of pre-school aged children. In addition to this, further findings of this review included the lack of evidence on the acceptability and feasibility of measurement tools, which is important to examine and understand to ensure that a tool can be successfully used. As such, the work described in *Chapter 5* was to assess the acceptability and feasibility of a range of measurement tools from the views of parents and carers. This review also highlighted that there is a need for more in-depth development and evaluation work of proxy reported tools, with the target population, which will be addressed in *Chapter 6* of this thesis.

4 [Chapter Four: The validity, reliability, and feasibility of measurement tools used to assess sleep of pre-school aged children: a rapid review](#)

The study in this chapter has been published in *Frontiers in Pediatrics*:

Phillips SM., Summerbell C., Ball HL., Hesketh KR., Saxena S., and Hillier-Brown, FC. (2021)

The Validity, Reliability, and Feasibility of Measurement Tools Used to Assess Sleep of Pre-school Aged Children: A Systematic Rapid Review. *Frontiers in Pediatrics*, 9:770262. doi:

10.3389/fped.2021.770262

4.1 Prelude

In the endeavour to assess the best way of measuring movement behaviours of pre-school children at scale, knowing about the measurement properties of tools used to assess PA and SB was fundamental. However, as I continued with my research, it became apparent that ‘sleep’ completed the picture, which was particularly evident through the inclusion of sleep in the latest WHO guidelines (as described in section 2.6.2.1 p.63). The review presented in *Chapter 3* was already underway with a published protocol, however, it was also evident that tools used to assess sleep can vary from those used in PA and SB research. Whilst it was important to identify measurement properties of available tools used to assess sleep, it was felt that a separate review was an appropriate next step of the research. On extensive searching of the literature, it was apparent that there were currently no reviews of this kind. On this basis, I conducted a study to fill this knowledge gap, to review the evidence base and determine the validity, reliability, and feasibility of measurement tools used to assess the sleep of pre-school aged children. In this review, I deliberately extracted information relating to the SES and ethnicity of participating families, to determine representation of different groups in the pre-existing measurement tool literature.

4.2 Introduction

This chapter presents a rapid review of the validity, reliability, and feasibility of measurement tools used to examine sleep of pre-school aged children. As I continued to research the movement behaviour ‘space’ in pre-school aged children it became apparent that sleep completed the picture, and that it was being introduced in the public health literature which had not been the case in previous years (Chaput, 2019; Chaput and Shiau, 2019). There was a growing interest in assessing the whole 24 hours of the day, and the

movement behaviours that this encompasses from sleep to PA (Galland et al., 2014; Quante et al., 2015). This was pertinently highlighted by the announcement of the WHO movement behaviours guidelines in April 2019, which incorporated sleep (alongside PA and SB) for the first time (WHO, 2019). At this point, the systematic review of the validity, reliability, and feasibility of PA and SB measurement tools was underway, including a published protocol (see appendix 3.2, online file accessed [here](#)). Additionally, until the recent shift in thinking, the measurement of sleep was generally examined independently from PA and SB (Quante et al., 2015). The range of tools used to assess sleep (as described in *Chapter two*, section 2.8, p.99), differ from those used to examine PA and SB. However, similarly to PA and SB measurement tools, there were no available overviews on the evidence of measurement properties (validity, reliability, feasibility) of pre-existing tools used to measure healthy sleep parameters (sleep duration and quality) of pre-school aged children. Several existing reviews have provided an overview of available sleep assessment methods with information on measurement properties where available (Erwin and Bashore, 2017; Ibáñez, Silva and Cauli, 2018; Sen and Spruyt, 2020; Spruyt and Gozal, 2011) or have identified tools available for assessing parental knowledge of their child's sleep (McDowall et al, 2017). Reviews examining the measurement properties of sleep measurement tools have focused on a specific type of tool, including accelerometers (Meltzer et al., 2012b; Schoch, Kurth and Werner, 2021), questionnaires (Lewandowski, Toliver-Sokol and Palermo, 2011; Nascimento-Ferreira et al., 2016), or proxy reported tools (Bauer and Blunden, 2008; Matricciani, 2013).

However, up to this point, no review had examined the measurement properties of the full range of measurement tools used to assess sleep, nor had they explored a range of measurement properties (validity, reliability and feasibility), and nor had they focused specifically on pre-school aged children. As such, as I had continued with my PhD research,

and gained the understanding that sleep ‘completed the picture’ with measuring movement behaviours of young children, I felt it was important to thoroughly examine the evidence base on the measurement properties of tools that already exist used to measure healthy sleep parameters of pre-school aged children, which I later conducted. Similar to the review outlined in *Chapter three* of this thesis, this included to understand the validity, reliability, and feasibility of the full range of tools specifically for the pre-school population. I felt this was important to understand the pre-existing evidence base, not only to identify questions that may have been applicable for the development of the movement behaviour tool, but also to examine which tools may be most appropriate for use as a comparison tool in further validation work. A rapid review was the method of choice, and I explain the reasons for this choice of methodology in the methods section of this chapter (section 4.4, p310-311).

4.3 Aims and objectives

The aim of this study was therefore to:

- Review studies examining the validity, reliability, and feasibility, of measurement tools used to assess sleep in pre-school aged children.

The objectives of this study were to:

- Identify, collate and synthesise all pre-existing tools (that had been evaluated) used to examine sleep of pre-school aged children
- Determine which tools were valid, reliable, and feasible for measuring sleep of pre-school aged children

4.4 Methods

This review is reported in line with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria (Page et al., 2021) (see appendix 4.1). I developed a detailed protocol for the review, which was registered with PROSPERO, registration number CRD42021230900 (see appendix 4.2, online file accessed [here](#)).

This review took the form of a rapid review (Grant and Booth, 2009). A rapid review has been defined as *“a form of knowledge synthesis that accelerates the process of conducting a traditional systematic review through streamlining or omitting a variety of methods to produce evidence in a resource-efficient manner”* (Hamel et al., 2021, p.7). Rapid reviews often employ similar, but modified, systematic reviewing methods. I conducted all phases of the review process using the same systematic methods that were used in the systematic review outlined in *Chapter three*. The modifications in the present review included employing only one primary reviewer for screening, risk of bias assessment and data extraction (Ganann, Ciliska and Thomas, 2010). Acknowledgement of the potential bias this may introduce, with ways in which these were overcome are presented in the discussion of this chapter. This included uncertainty on study eligibility resolved through consultation with a second reviewer (Dr Frances Hillier-Brown, supervisor and co-author), and a further reviewer (Professor Carolyn Summerbell, supervisor and co-author) checking the outcomes of the data extraction and risk of bias assessment against all of the original studies. I also followed the ‘A measurement tool to assess systematic reviews’ (AMSTAR) protocol as closely as possible (Shea et al., 2007) to ensure that the review was conducted to a high quality standard (Kelly, Moher and Clifford, 2016)

A rapid review was selected as the method of choice to identify, synthesise and appraise research on the validity, reliability, and feasibility of measurement tools used to assess the sleep of pre-school aged children. The rationale for using a rapid review was due to understanding that the literature for sleep measurement was less extensive than that of the review outlined in *Chapter three*; knowledge obtained through scoping searches and conversations with sleep experts. As such, it was decided that whilst it was important to conduct a review on the sleep measurement literature in pre-school aged children to determine the current evidence base for tools used for this purpose, a less intensive review method was appropriate to complement the first review and to map out the sleep measurement tool literature.

4.4.1 Search strategy

Systematic searches of six major electronic databases (Scopus (Science Direct) Web of Science (Web of Science), MEDLINE (Ovid), APA PsycARTICLES (EBSCOhost), APA PsycINFO (EBSCOhost), and SPORTDiscus (EBSCOhost)), the grey literature (opengrey.eu) and trial registries (International Clinical Trials Registry Platform) were conducted in April 2021 to identify relevant studies. Searches of topic specific journals (*Sleep* and *Sleep Health*), reference and citation searches of the included studies, as well as reference checking of selected existing reviews (Bauer and Blunden, 2008; Erwin and Bashore, 2017; Ibáñez, Silva and Cauli, 2018; Lewandowski, Toliver-Sokol and Palermo, 2011; Matricciani, 2013; McDowall et al, 2017; Meltzer et al., 2012b; Nascimento-Ferreira et al., 2016; Schoch, Kurth and Werner, 2021; Sen and Spruyt, 2020; Spruyt and Gozal, 2011) were conducted for completeness on 4th May 2021.

The search strategy included combinations of the construct (sleep); population (pre-school, early years, early childhood, young children and kindergarten); and measurement properties (assessment, measurement, method, valid, reliable, feasible). Searches were adapted to each database, alongside the use of appropriate Boolean operators and database specific filters (see Table 4.1). No restrictions were placed on language or year of study. Multiple preliminary searches were conducted to ensure that search strategy could identify a selection of clearly eligible studies (e.g. Duraccio et al., 2018; Iwasaki et al., 2010; Sekine et al., 2002; Werner et al., 2008).

Table 4.1. Full search strategy and outcomes

Initial searches were conducted on PROSPERO, electronic databases (Scopus, Web of Science, PsycARTICLES, PsycINFO, MEDLINE, CINAHL and SPORTDiscus), and of the topic specific journals to identify pre-existing reviews. Citation searches of relevant reviews were conducted to identify further reviews to ensure that a similar review had not already been conducted.

Source of Search Conducted	Date of Search	Searches Conducted	Search Outcome (n)
Web of Science (<i>Web of Science</i>)	14/04/21	ALL=(sleep) AND ALL= (valid* OR reliab* OR feasib*) AND ALL= (pre\$school* OR "early years" OR "early childhood" OR "young children" OR kindergarten*) AND ALL= (assess* OR measur* OR method*)	464
Medline (<i>Ovid</i>)	14/04/21	(sleep and (valid* or reliab* or feasib*) and (pre\$school* or "early years" or "early childhood" or "young children" or kindergarten*) and (assess* or measur* or method*)).af.	788
APA PsycARTICLES, APA PsycINFO, SPORTdiscus (<i>EBSCOhost</i>)	14/04/21	TX SLEEP AND TX (assess* OR measur* OR method*) AND TX (valid* OR reliab* OR feasib*) AND TX (pre#school* OR "early childhood" OR "early years" OR "young children" OR kindergarten*)	2425
Scopus (<i>Science Direct</i>)	14/04/21	(TITLE-ABS-KEY ("sleep") AND TITLE-ABS-KEY ((pre*school* OR "early years" OR "early childhood" OR "young children" OR kindergarten*)) AND TITLE-ABS-KEY ((assess* OR measure* OR method*)) AND TITLE-ABS-KEY ((valid* OR reliab* OR feasib*)))	976
Trial registries- International Clinical Trials Registry Platform	14/04/21	SLEEP AND CHILD*	703

Grey literature- opengrey.eu	14/04/21	sleep AND "preschool*" OR "early years" OR "early childhood" OR "young children" OR "kindergarten"	1068
Sleep Health	04/05/21	pre*school* OR "early years" OR "early childhood" OR "young children" OR kindergarten*	153
Sleep	04/05/21	"preschool" OR "pre-school" OR "early years" OR "early childhood" OR "young children" OR "kindergarten"	746
Reference lists of existing reviews		Reference lists	1029
Citation searching/reference list searching of included studies	During data extraction	Reference lists/citation searching	

4.4.2 Eligibility criteria

4.4.2.1 *Inclusion Criteria*

Articles were eligible for inclusion if their aim was to examine the measurement properties (validity and/or reliability and/or feasibility) of a tool used to measure the sleep of a typically developing sample of pre-school children aged 3-7 years old. For the same reasons as those outlined in chapter three (see section 3.4.2, p.147) and to ensure consistency between the work, we used the broad age range for a 'pre-school child', referring to any child who has not yet reached the age of formal schooling (World Bank Data, 2020). There were no restrictions on study design or setting.

4.4.2.2 *Exclusion Criteria*

Articles were excluded if:

- a) The measurement tool was examined in children outside of the pre-school age range (aged <3 or >7 years old), and did not include independent analysis of children within this age group (e.g. did not include an analysis of the measurement tool for 3-5 year olds only).
- b) The measurement tool was examined in children with clinically diagnosed conditions that may impact on sleep (e.g. autism, attention deficit hyperactivity disorder (ADHD), cerebral palsy, sleep apnoea).
- c) The measurement tool had a primary purpose of determining clinical sleep problems and disturbances, such as sleep apnoea or sleep disordered breathing (studies were included where they report on behavioural 'sleep problems')

including variables such as night waking or delayed sleep onset, but not when diagnosing clinical sleep problems).

- d) The article was a book or review.
- e) There was not a full text article or abstract with sufficient information.

4.4.3 Screening for relevant studies

All identified articles were imported into a referencing manager software (Endnote X20) and duplicates removed. First, I screened titles and abstracts of identified articles, followed by full texts of potentially relevant articles, for inclusion. Articles where the eligibility was uncertain were independently double screened by Dr Frances Hillier-Brown; consensus on the eligibility of these articles was reached through discussion. In some cases, where eligibility was unclear (e.g. the study population was not clear or a full text could not be accessed), authors of the studies were contacted to ensure appropriate inclusion or exclusion of the study (n=4). This included one instance where the authors of a potentially relevant article written in Japanese could not be accessed.

4.4.4 Data extraction

A pre-piloted (on a subsample of included studies) Excel data extraction form was used to extract data from the included studies. I extracted data from all relevant studies, and Professor Carolyn Summerbell checked all extracted data against the original articles for accuracy. Extracted information included: study characteristics (authors, country, publication year, sample size, attrition rates); participant characteristics (age, sex, ethnicity, socioeconomic status); aims of the study; methods used, including: measurement tool examined (e.g. sleep diary), comparison measurement tool(s), study setting (free

living/laboratory), number of nights of data collected, number of nights of data used in analysis, whether week nights and weekend nights were calculated separately, 24 hour sleep time or night time sleep only, whether data were systematically collected on day time naps, and for device based measurement tools: type, placement, epoch, cut point and algorithms used; outcome measures explored (e.g. sleep duration); measurement property examined (e.g. convergent validity); statistical method used; results of the study (including information on missing data) and sources of funding.

4.4.5 Risk of bias

As outlined in chapter three (section 3.4.5, p.149-151), I conducted comprehensive searches to determine the most appropriate risk of bias assessment to conduct on studies of measurement properties. In this review, I made the decision to use the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) risk of bias checklist on all included studies (Mokkink et al., 2018b; Prinsen et al., 2018; Terwee et al., 2018a). Although the COSMIN risk of bias checklist was devised for studies on patient reported outcome measures, the majority of the studies included in this review were evaluating parental reported tools, with a very limited number of studies examining accelerometers. Therefore, to ensure consistency in the risk of bias assessment of all included studies I felt it was appropriate to use the COSMIN checklist, rather than to complete a separate risk of bias assessment on just the small sub-sample of studies that were evaluating accelerometers, as this could subsequently bias the results due to inconsistencies in interpretation.

I conducted the risk of bias assessment on all included studies, to ensure accuracy in interpretation and to follow PRISMA guidelines as closely as possible, the assessment was

double checked against the original articles by Professor Carolyn Summerbell. Based on the studies included in the review, I conducted the assessment using the sub-sections relating to structural (construct) validity, criterion validity, construct (convergent) validity, test-retest reliability, and internal consistency. Only those sections relevant to the particular study were conducted. Each item was scored using the 4 point scale outlined by COSMIN (very good, adequate, doubtful, inadequate) (Mokkink et al., 2018b). The overall quality assessment of a study was determined using ‘the worst score counts’ principle (e.g. if one item was scored as ‘inadequate’, the overall score of the measurement property in that study would be ‘inadequate’)(Terwee et al., 2012). Risk of bias may be present if the overall quality assessment is doubtful or inadequate, or if there is only one study where the risk of bias is quality assessed as adequate (Mokkink et al., 2018a).

4.4.6 Definitions of validity, reliability and feasibility

For clarity with definitions used to determine key constructs outlined in this review, Table 4.2 provides an overview of the definitions of measurement properties that were examined in this review (Bowen et al., 2009; Evenson, Goto and Furberg, 2015; Mokkink et al., 2010b; Terwee et al., 2018a).

Table 4.2. Definition of each of the measurement properties included in this review

Measurement property	Definition
Validity	<i>Ability for a measure to accurately reflect the construct it is designed to measure.</i>
Construct (or structural) validity	The extent to which the measurement tool actually tests the hypothesis or theory they are measuring. This is usually examined by just the one measurement tool being explored, using statistical methods such as confirmatory or exploratory factor analysis.
Content validity	Extent to which a measure covers all aspects of the intended domains or dimensions that it claims to measure. This is usually examined through qualitative means. Often includes assessment of face validity , which is the appearance of a measure, in that it appears to measure what it claims to measure.
Criterion validity	Output of a measure produces similar results to a 'gold standard'. This includes studies that have examined a tool against polysomnography (which may be alongside videosomnography or direct observation).
Convergent (concurrent) validity	The extent of the agreement between measures. This includes studies that have examined comparisons between multiple measurement tools (such as accelerometers, diaries and questionnaires) but do not include a criterion method of polysomnography or videosomnography.
Reliability	<i>Extent to which a tool gives measurements that are consistent, stable and repeatable.</i>
Test-retest reliability	The extent to which a measure can obtain similar results in repeated trials, keeping as many conditions stable as possible.
Internal consistency	The extent to which items among a measurement tool that propose to measure the same construct are interrelated.
Feasibility	The extent to which a measurement tool: is suitable for the target population; can be successfully delivered in the target population/context; shows promise of being successful within the intended population. Includes participant and researcher acceptability, and cost, which can be assessed for all measurement tools through qualitative feedback of participants and through missing or lost data occurred from the measurement tool.

Definitions guided by: Bowen et al., 2009 Evenson, Goto and Furberg, 2015; Mokkink et al., 2010b; Terwee et al., 2018a

4.4.7 Interpretation and synthesis of results

I used the same approach as outlined in *Chapter three* to ensure consistent interpretation of the statistical results of individual studies (see section 3.4.7, p156-158). An overview of what constitutes a 'weak', 'moderate', or 'good' statistical result for the measurement properties of validity or reliability has been provided in Table 4.3 (De Vet et al., 2011; Evenson, Goto and Furberg, 2015; Hidding et al., 2017; Hidding et al., 2018; Tavakol and Dennick, 2011). Individual studies have been ranked based on the interpretation outlined. Feasibility assessment included: how much missing data or non-compliance occurred with any measurement tool (measurement tool under study or comparison tool), qualitative feedback, or information on the cost or acceptability of the tools. In addition to measurement property analysis, I also extracted information relating to SES and ethnicity of participants, so that it would be possible to analyse results from these studies to identify how well tools worked across or within different sociodemographic groups. This review presents all results using a narrative synthesis approach; a meta-analysis was not possible due to a lack of multiple studies examining the same measurement tools and the use of different comparison measurement tools.

Table 4.3. Main statistical analyses and interpretation of statistics

Statistical Analysis		Weak	Moderate	Good
Correlations (r)	Pearson's	<0.60	0.60-0.79	≥0.80
	Spearman's	<0.60	0.60-0.79	≥0.80
	Polyserial correlation coefficient	<0.60	0.60-0.79	≥0.80
	Intraclass correlation coefficient	<0.60	0.60-0.69	≥0.70
Cronbach's alpha		<0.60	0.60-0.69	≥0.70
Bland Altman ^a		>31 minutes deemed not satisfactory agreement		≤30 minutes
T-tests		^b	^b	^b
Feasibility		Narrative description		

Interpretation in line with: De Vet et al., 2011; Evenson, Goto and Furberg, 2015; Hidding et al., 2017; Hidding et al., 2018; Tavakol and Dennick, 2011 ^a

Extracted based on the interpretation of the included studies using this analysis method, a priori defined satisfactory agreement of ≤30 minutes only. ^b

Depends on the units of measure.

4.5 Results

4.5.1 Study selection

Initial database searches identified 4,298 articles. From this, 73 full text articles were screened and, of these, 13 articles met the eligibility criteria and were included in the review. A further 3 articles were identified through searching of reference lists of included studies and were included in the review (Figure 4.1). Studies were mainly excluded due to including children outside of the specified age range. Excluded studies with reasons are outlined in appendix 4.3.

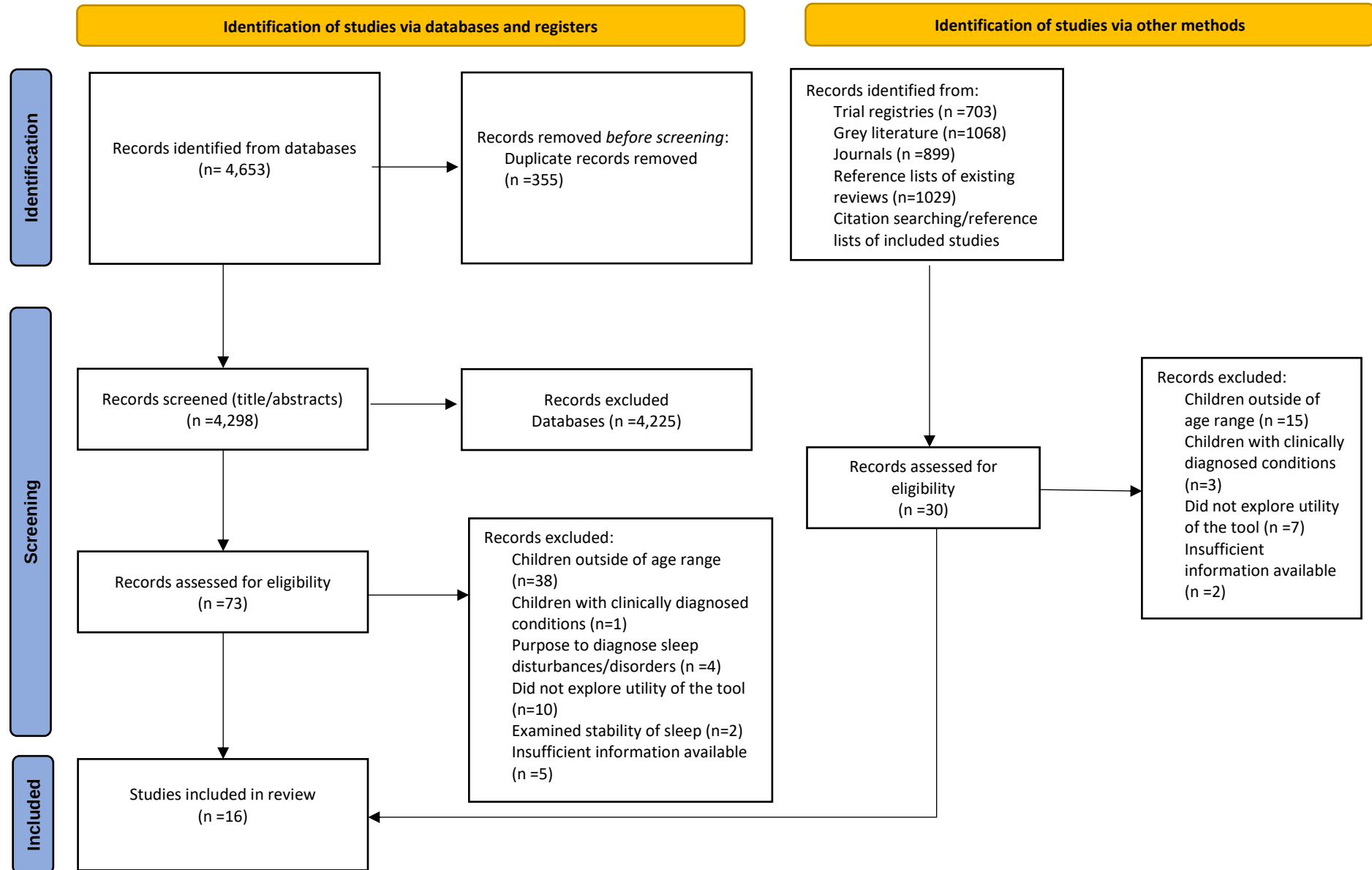


Figure 4.1. PRISMA flow chart of included studies

4.5.2 Description of studies

Included articles were published between 2001 and 2021 and were all conducted in high income countries (World Bank Data, 2021): USA (n=5), Japan (n=3), Israel (n=2), China (n=1), New Zealand (n=1), Portugal (n=1), Spain (n=1), Switzerland (n=1) and UK (n=1). Sample sizes ranged from 14 (Meltzer et al., 2015) to 346 children (Ishihara, Doi and Uchiyama, 2014). All studies that reported the sex of the children included both male and female children. The median age of the included children was 4.9 years. Most studies examined the measurement properties of the tools in free-living conditions to determine habitual sleep behaviours (n=15). One study was laboratory based and used PSG as the criterion method to measure sleep (Meltzer et al., 2015).

Thirteen studies examined parental reported measurement tools (questionnaires and diaries), and three remaining studies examined accelerometers. Fourteen studies examined the measurement properties of one measurement tool only and two studies examined the measurement properties of two measurement tools (Iwasaki et al., 2010; Werner et al., 2008). Nine studies examined questionnaires; two studies generated a new questionnaire (Sekine et al., 2002; Werner et al., 2008), three adapted questionnaires developed for other age groups (Ishihara, Doi and Uchiyama, 2014; Iwasaki et al., 2010; Kushnir and Sadeh, 2013), and four studies used the *Children's Sleep Habits Questionnaire* (CSHQ) (two in English, one variation translated into Spanish and one in Chinese), a measure originally developed for children aged 4-10 years old (Duraccio et al., 2018; Lucas-de la Cruz et al., 2016; Perpétuo, Fernandes, and Veríssimo, 2020; Tan et al., 2018). Six studies examined the measurement properties of different diaries. Diaries were specific to the study, with two studies utilising parental diaries frequently used within a clinical care setting

(Lam et al., 2011; Werner et al., 2008). Three studies examined the measurement properties of three types of accelerometer: Actigraph GT3X+, Fitbit Ultra and MicroMini Motionlogger.

The most frequently reported measurement property was convergent validity (n=13). Three studies examined the construct (structural) validity of the tools, one examined criterion validity, two studies the test-retest reliability and four the internal consistency. No studies reported the qualitative feasibility of the tools, though twelve studies reported missing data or non-completion that demonstrated an element of feasibility. Table 4.4 provides an overview of each study and the measurement properties that the study examined.

Table 4.4. Overview of measurement properties examined in each study

Study	Country	Measurement instrument under study	Construct Validity	Criterion Validity	Convergent Validity	Test-retest reliability	Internal consistency	Feasibility
Accelerometers								
Meltzer et al., 2015	USA	Fitbit Ultra		●				●
Tracy et al., 2021	USA	Actigraph GT3X+			●			
Staples et al., 2019	USA	MicroMini Motionlogger	●				●	●
Questionnaires								
Duraccio et al., 2018	USA	Children's Sleep Habits Questionnaire (CSHQ)			●			● ^a
Perpétuo, Fernandes, and Veríssimo, 2020	Portugal	Children's Sleep Habits Questionnaire (CSHQ)			●			●
Lucas-de la Cruz et al., 2016	Spain	Child Sleep Habits Questionnaire (CSHQ-Spanish)	●		●	●	●	●
Tan et al., 2018	China	Children's Sleep Habits Questionnaire (CSHQ- Chinese)	●				●	
Ishihara, Doi and Uchiyama, 2014	Japan	Children's ChronoType Questionnaire (CCTQ-Japanese)			●	●	●	●

Iwasaki et al., 2010	Japan	Brief 12 item questionnaire			●			● ^a
Kushnir and Sadeh, 2013	Israel	Brief Child Sleep Questionnaire (BCSQ)			●			● ^a
Sekine et al., 2002	Japan	Sleep habits questionnaire			●			● ^a
Werner et al., 2008	Switzerland	Sleep schedule time questionnaire (SSTQ)			●			● ^a
Diaries								
Galland et al., 2016	New Zealand	Sleep diary			●			
Iwasaki et al., 2010	Japan	Sleep diary			●			● ^a
Jones and Ball, 2014	UK	Sleep diary			●			● ^a
Lam et al., 2011	USA	Sleep diary			●			● ^a
Tikotzky and Sadeh, 2001	Israel	Sleep diary			●			
Werner et al., 2008	Switzerland	Sleep diary			●			● ^a

^a=Feasibility assessment for comparison tool only

4.5.3 Risk of bias assessment

Three studies assessing construct (structural) validity were quality assessed as adequate (Staples et al., 2019; Tan et al., 2018) or doubtful (Lucas-de la Cruz et al., 2016). One study assessing criterion validity was quality assessed as inadequate (Meltzer et al., 2015) due to the type of analysis conducted. The majority of studies assessing the convergent validity of the tools were quality assessed as high risk of bias; eight inadequate (Duraccio et al., 2018; Ishihara, Doi and Uchiyama, 2014; Iwasaki et al., 2010; Kushnir and Sadeh, 2013; Lam et al., 2011; Lucas-de la Cruz et al., 2016; Perpétuo, Fernandes, and Veríssimo, 2020; Werner et al., 2008) and four doubtful (Galland et al., 2016; Jones and Ball, 2014; Sekine et al., 2002; Tracy et al., 2021), with the exception of one study quality assessed as very good (Tikotzky and Sadeh, 2001). Main reasons for poor methodological quality of the studies was due to unknown measurement properties of the comparator tools. Additionally, in some studies the measurement tools were assessing different time periods (e.g. questionnaire measuring the week before the period of accelerometer wearing (Duraccio et al., 2018; Kushnir and Sadeh, 2013). Two studies examining test-retest reliability were quality assessed as inadequate (Ishihara, Doi and Uchiyama, 2014; Lucas-de la Cruz et al., 2016) due to the choice of statistical method used. Four studies examining internal consistency were quality assessed as very good (Ishihara, Doi and Uchiyama, 2014; Lucas-de la Cruz et al., 2016; Staples et al., 2019; Tan et al., 2018). The full risk of bias assessment can be found in appendix 4.4. No studies were removed from the overall analysis based on the risk of bias assessment; however, the quality of studies is acknowledged throughout the results and discussion.

4.5.4 Summary of measurement properties of tools

Detailed information of studies for each measurement property can be found in Tables 4.5 (construct validity), 4.6 (criterion validity), 4.7 (convergent validity), 4.8 (test-retest reliability), 4.9 (internal consistency) and 4.10 (feasibility). Table 4.11 provides a summary table of the evidence for validity (criterion and convergent) and reliability of the measurement tools. Construct validity and feasibility are not included in this table due to not being able to present the results of these studies in this way.

4.5.5 Validity

4.5.5.1 Construct (structural) Validity

Two studies examined the construct validity of the CSHQ: one a Spanish translation (Lucas-de la Cruz et al., 2016) and a Chinese translation (Tan et al., 2018). Lucas-de la Cruz and colleagues conducted an exploratory factor analysis and deemed keeping the same factor structure as the original questionnaire to be appropriate (Owens, Spirito and McGuinn, 2000). Tan and colleagues conducted a confirmatory factor analysis and determined that no existing factor structures were suitable for the data with pre-school aged children. Following this, they performed an exploratory and a further confirmatory factor analysis to determine a new four factor structure, including: bedtime behaviours, sleep behaviours, morning waking, and daytime sleepiness (Tan et al., 2018).

One study examined the construct validity of the outcome measures of the MicroMini Motionlogger accelerometer using principal component analysis (Staples et al., 2019). This study revealed a four component structure: sleep activity, sleep variability, sleep timing and sleep duration, which could categorise accelerometer outcome variables.

Daytime sleep and sleep latency represented exclusive elements of sleep that could not be categorised into factor structures.

Table 4.5. Detailed information of studies examining construct validity

Study	Study sample	Measurement tool under study	Methods	Structural Validity Results	Risk of bias
Lucas-de la Cruz et al., 2016	Spain; N=286; 4-7 years;	Child Sleep Habits Questionnaire (CSHQ)- Spanish version	Parents completed the Child Sleep Habits Questionnaire (CSHQ)- Spanish version	Exploratory factor analysis conducted – 10 factors with eigenvalues >1. When analysed items' content in each subscale they did not have logical connection. Eight factor model could be appropriated but many items not allocated in correct subscales. Kept the same factor structure as original questionnaire.	Doubtful
Tan et al., 2018	N=171; 4-5 years; 53% female, 47% male; China Chinese kindergarteners. Most parents worked full-time. Over 80% of mothers and 85% fathers had at least a college level education. Over half of the mothers and 80% fathers made ¥5000 or more monthly.	Children's Sleep Habits Questionnaire (CSHQ)- Chinese version	Parents completed the Chinese version of the CSHQ-C	Confirmatory factor analysis –Concluded that none of the existing factor structures (5 factor, 8 factor, 4 factor) outlined in previous studies and the original questionnaire were suitable for the data with this age group. Exploratory factor analysis – 6 factor structure best comparative fit = 0.89, followed this up with item reduction and confirmatory factor analysis with 4 factor structure: bedtime behaviours ($\alpha = .59$), sleep behaviours ($\alpha=0.62$), morning wakings ($\alpha=0.69$) and daytime sleepiness ($\alpha=0.67$).	Adequate

Staples et al., 2019	N=107 (at 3 years), n=104 (at 3.5 years); 3-3.5 years; Sex not reported for this age category; USA Information on parental education, ethnicity and socioeconomic status was available but for the whole sample, not separated for age groups of interest. Predominantly college educated parents, European American, high socioeconomic status.	MicroMini Motionlogger Accelerometer; <i>Wrist worn;</i> <i>Scored using Action W-2 software and Sadeh algorithm⁶⁴</i> <i>Sleep diary also completed to record bedtime, night awakenings, rise times and other issues relating to child's sleep.</i>	Accelerometer worn for 7 days at three time points in longitudinal study (2.5 years, 3 years, 3.5 years).	Principal component analysis (with oblique rotation): Strong loadings for a four factor structure at both time points, explaining 82% of the variance (when child aged 3) and 81% (when child age 3.5 years): 1. Sleep activity -Average time awake after sleep onset -Average minute to minute activity level -Average n of night awakenings lasting 5 minutes or more -Average duration of longest wake episode after sleep onset -Average % of active epochs after sleep onset 2. Sleep variability -Time of sleep onset -Duration of time in bed -Duration of sleep period -Time of midsleep -Bedtime -Minutes asleep whilst in bed 3. Sleep timing -Average time of midsleep -Average time of sleep onset -Average bedtime 4. Sleep duration -Average sleep period -Average duration of time in bed -Average minutes asleep while in bed	Adequate

				The daytime sleep variables (duration and variability) and night time latency to sleep onset were indicated to be independent from the 4 component structure when assessed with children aged 2.5 years, so were not included in these analyses.	

4.5.5.2 Criterion Validity

One study examined the criterion validity of the Fitbit Ultra compared with PSG (Meltzer et al., 2015). The Fitbit Ultra (sensitive mode) underestimated sleep time and sleep efficiency, and overestimated wake after sleep onset. The Fitbit Ultra (normal mode) overestimated sleep time and sleep efficiency, but there was no significant difference for wake after sleep onset.

Table 4.6. Detailed information of studies examining criterion validity

Study	Study sample	Measurement tool under study	Comparison measurement tool(s)	Methods	Outcome measures	Validity Results	Risk of bias
Meltzer et al., 2015	N=14 3-5 years; Sex not reported for this age category; USA Ethnicity reported for whole sample, but not separately by age category.	Fitbit Ultra <i>Non-dominant wrist; 1 minute epoch; Proprietary software algorithm for both normal and sensitive sleep recording modes.</i>	Polysomnography (PSG): Sandman 10.0 platform and video and audio recordings.	Laboratory based (day of week unknown) Overnight polysomnography in the pediatric sleep laboratory at Children's of Alabama. Children wore a Fitbit Ultra for the duration of the night. Some children randomly assigned to also wear two accelerometers, but results not available as to whether this included pre-school aged children.	Both measurement tools: - Total sleep time - Wake after sleep onset - Sleep efficiency	Fitbit Ultra (Normal setting) significantly overestimated total sleep time and sleep efficiency ($p<0.05$). No significant difference for wake after sleep onset, but was lower from the Fitbit Ultra (26.6 minutes) than PSG (45.5 minutes). Fitbit Ultra (sensitive mode) significantly underestimated total sleep time, overestimated wake after sleep onset and underestimated sleep efficiency ($p<0.001$).	Inadequate

4.5.5.3 *Convergent Validity*

Accelerometers

One study examined the ability of the Actigraph GT3X+ accelerometer to determine bedrest and wake time, using a newly developed decision tree algorithm specific to pre-school aged children (Tracy et al., 2021). The developed algorithm was able to detect bedrest and wake time similarly to visual identification of the data and was better at identifying bedrest than algorithms developed in other age groups that are often applied to young children (Sadeh, Sharkey and Carskadon, 1994; Tracy et al., 2014). There were significant differences between the outcomes of the algorithm and parental report ($p < 0.001$).

Questionnaires

Eight studies examined the convergent validity of six different questionnaires (all parental reported). The most frequently assessed questionnaire was the CSHQ, in English (Duraccio et al., 2018; Perpétuo, Fernandes, and Veríssimo, 2020) and Spanish (Lucas-de la Cruz et al., 2016). The majority of questionnaires assessed 'typical' sleep (Duraccio et al., 2018; Lucas-de la Cruz et al., 2016; Perpétuo, Fernandes, and Veríssimo, 2020; Tan et al., 2018; Werner et al., 2008), others varied including: daily reporting (Sekine et al., 2002), past week (Kushnir and Sadeh, 2013) and the past month (Iwasaki et al., 2010). Three questionnaires systematically collected data on naps (Duraccio et al., 2018; Perpétuo, Fernandes, and Veríssimo, 2020; Werner et al., 2008), however none reported on this outcome measure. Four studies examined weekday and weekend days separately (Ishihara, Doi and Uchiyama, 2014; Iwasaki et al., 2010; Perpétuo, Fernandes, and Veríssimo, 2020;

Werner et al., 2008), one study stated that data were collected on weekdays only (Duraccio et al., 2018).

All studies used an accelerometer as the comparison tool, with only two studies using the same accelerometer (*Actiwatch 2*) (Ishihara, Doi and Uchiyama, 2014; Perpétuo, Fernandes, and Veríssimo, 2020). Two studies also used sleep diaries alongside accelerometers, to determine sleep onset and offset (Ishihara, Doi and Uchiyama, 2014; Lucas-de la Cruz et al., 2016) and average bedtime and wake time (Duraccio et al., 2018).

All questionnaires assessed sleep duration. The *Brief Child Sleep Questionnaire* (BCSQ) (Kushnir and Sadeh, 2013) and *sleep habits questionnaire* (Sekine et al., 2002) showed high convergence with accelerometers for determining sleep duration ($r=0.85$ to $r=0.90$). The *Children's ChronoType Questionnaire* (CCTQ) showed a moderate correlation for weekdays but weak for weekend days (Ishihara, Doi and Uchiyama, 2014). The remaining questionnaires generally showed weak associations with accelerometry for determining sleep duration (Duraccio et al., 2018; Iwasaki et al., 2010; Lucas-de la Cruz et al., 2016; Perpétuo, Fernandes, and Veríssimo, 2020; Werner et al., 2008). No questionnaires were deemed satisfactory at determining sleep latency (Duraccio et al., 2018; Ishihara, Doi and Uchiyama, 2014; Iwasaki, 2010; Lucas-de la Cruz et al., 2016) or night awakenings (Duraccio et al., 2018; Iwasaki et al., 2010; Kushnir and Sadeh, 2013; Lucas-de la Cruz et al., 2016; Perpétuo, Fernandes, and Veríssimo, 2020).

Results for wake up time were variable; the CSHQ was highly convergent with accelerometry, more so for weekdays (7 min difference) than weekend days (28 min difference) (Perpétuo, Fernandes, and Veríssimo, 2020), with similar patterns observed for the CCTQ (Ishihara, Doi and Uchiyama, 2014), The *Sleep Schedule Time Questionnaire* (SSTQ)

and *brief questionnaire* were deemed to be weak at determining wake up time (Iwasaki et al., 2010; Werner et al., 2008).

Bedtime reports were moderate for the CCTQ (Ishihara, Doi and Uchiyama, 2014) and weak for the *brief questionnaire*, for weekday and weekend days (Iwasaki et al., 2010). Sleep onset time reports were moderate for the CCTQ and BCSQ (Ishihara, Doi and Uchiyama, 2014; Kushnir and Sadeh, 2013) and weak for the *brief questionnaire* (Iwasaki et al., 2010) and SSTQ (Werner et al., 2008).

Diaries

Six studies examined the convergent validity of sleep diaries. The format of the diaries varied, some used time intervals to determine sleep and wake status (Galland et al., 2016; Lam et al., 2011), whilst others requested certain information, such as sleep onset and end time (Iwasaki et al., 2010; Jones and Ball, 2014; Lam et al., 2011; Tikotzky and Sadeh, 2001). All diaries were parental reported, however, three studies explicitly stated that day time sleep records were based on information from nursery teachers (Galland et al., 2016; Iwasaki et al., 2010; Lam et al., 2011).

Diaries were completed daily in all studies (Galland et al., 2016; Iwasaki et al., 2010; Jones and Ball, 2014; Lam et al., 2011; Tikotzky and Sadeh, 2001; Werner et al., 2008), for a duration of 2 (Galland et al., 2016) to 6-8 days (Werner et al., 2008). The majority of diaries collected data on nap times (Galland et al., 2016; Jones and Ball, 2014; Lam et al., 2011; Tikotzky and Sadeh, 2001; Werner et al., 2008), however only three reported results on this outcome measure (Galland et al., 2016; Jones and Ball, 2014; Lam et al., 2011). Three studies reported results for weekday and weekend days separately (Iwasaki et al., 2010; Lam et al., 2011; Werner et al., 2008).

All studies used an accelerometer as the comparison tool; different types of accelerometers were used, although, it was uncertain whether two studies used the same accelerometer as reporting was unclear (Iwasaki et al., 2010; Tikotzky and Sadeh, 2001). Diary and accelerometer data were collected simultaneously in all studies.

Three studies demonstrated that the sleep diaries were highly convergent for determining nap times in comparison with accelerometry (Galland et al., 2016; Jones and Ball, 2014; Lam et al., 2011). Diaries were generally similar to accelerometry for determining sleep onset (Iwasaki et al., 2010; Jones and Ball, 2014; Lam et al., 2011; Tikotzky and Sadeh, 2001; Werner et al., 2008) and sleep end/wake time (Iwasaki et al., 2010; Jones and Ball, 2014; Tikotzky and Sadeh, 2001; Werner et al., 2008), with the exception of one study where parents reported that the child rose later than that detected by the accelerometer (Lam et al., 2011). Mixed results were reported for the diaries association with accelerometry for sleep duration, including: good (Tikotzky and Sadeh, 2001), moderate (Werner et al., 2008), and weak (Lam et al., 2011). One diary showed weak correlations for weekday but moderate for weekend days (Iwasaki et al., 2010). Similar to questionnaires, diaries were all rated as weak for determining night awakenings (Iwasaki et al., 2010; Lam et al., 2011; Tikotzky and Sadeh, 2001; Werner et al., 2008). One study reported weak associations between the diary and accelerometer for sleep latency and sleep quality (based on sleep efficiency, true sleep time and night awakenings) (Tikotzky and Sadeh, 2001).

Table 4.7. Detailed information of studies examining convergent validity

Study	Study sample	Measurement tool under study	Comparison measurement tool(s)	Methods	Outcome measures	Validity Results	Risk of bias rating
Accelerometers							
Tracy et al. 2021	N=200; 3-6 years; 49% male, 51% female; USA Ethnicity and parental education reported for whole sample, but not separately for development and validation groups. However, overall sample was largely (90%) Hispanic/Latino. Majority of parents (62%) had less than a high school degree or equivalent.	Actigraph GT3X+ <i>Waist; 60 sec epoch; Non wear periods identified and excluded using Choi et al., 2011 algorithm</i>	Visual identification of bedrest and wake periods by two trained raters. Parental reported sleep survey Pre-existing algorithms: Sadeh's automated sleep algorithm and decision tree (DT) youth algorithm by Tracy et al., 2014	Free living (weekday and weekend day required) Children wore the accelerometer on their waist for 7 days (including day time), a recording of ≥ 3 valid nights (of ≥ 6 hours) including ≥ 2 weeknights and ≥ 1 weeknight for inclusion.	For all tools: Bedrest (sleep time) and wake time periods. Decision tree algorithm and parental survey: Children taking naps, nap frequency and nap duration	Independent samples t-test: Decision tree performed better at identifying bedrest than decision tree - youth and Sadeh's algorithm ($p < 0.001$). Paired t-tests: Bedrest determined by decision tree was not significantly different from visual identification of the data ($p = 0.859$) but was significantly shorter than Sadeh's algorithm and parental survey ($p < 0.001$). Wake time determined by decision tree was not significantly longer from visual identification of the data ($p = 0.859$) but was significantly different to Sadeh's algorithm ($p < 0.001$).	Doubtful

						Bedrest determined by Sadeh's algorithm was significantly different to parental survey and visual identification ($p < 0.001$). Wake time determined by Sadeh's algorithm was significantly shorter than visual identification ($p < 0.001$). Bedrest determined by parental survey was significantly longer than from visual identification ($p < 0.001$). The number of children taking naps was significantly lower when determined by parental report than decision tree identified ($p < 0.001$). Nap frequency was higher and duration longer when reported by parental report than decision tree identified ($p < 0.001$). Overlap coefficients:	

						Total sleep significantly different between parental report and decision tree, parental report and Sadeh's algorithm and decision tree and Sadeh's algorithm (all $p < 0.001$). (Inter-rater agreement for visual identification of data was very good $k = 0.99$)	
Questionnaires							
Duraccio et al., 2017	N=108; 4-6 years; 42% female; 53% male; 5% missing USA 86% Caucasian, 2% Hispanic, >1% Native American, 2% Asian, 5% Other, 5% Missing Mean monthly income = \$5,693.73 (Likely high socioeconomic status as	Children's Sleep Habits Questionnaire (CSHQ) <i>Parental reported typical week of sleep</i>	Actigraph GT3X+ <i>Right hip 15 sec epoch Sadeh algorithm, sleep time established by parental report in a demographics questionnaire of average bed and wake up time</i>	Free living (week days only) CSHQ completed electronically, 3 days of continuous Actigraph GT3X+ wearing one week later	CSHQ: - Sleep duration - Sleep onset delay - Night awakenings - Daytime sleepiness Actigraph: - Minutes asleep - Sleep latency - Minutes awake after sleep onset - Number of awakenings and average	Pearson's Correlations: Minutes asleep (actigraph) and sleep duration (CSHQ) $r = -0.13$ Sleep latency (actigraph) and sleep onset delay (CHSQ) $r = -0.16$ Minutes awake after sleep onset (actigraph) and night awakenings (CHSQ) $r = -0.06$, Number of awakenings (actigraph) and night awakenings (CSHQ) $r = -0.05$. Average awakening length (actigraph) and night awakenings (CSHQ) $r = 0.16$.	Inadequate

	recruited through universities)				awakening length • Sleep efficiency	Sleep latency (Actigraph) and night awakenings (CSHQ)($r=0.30$, $p=0.002$) Sleep efficiency (actigraph) and daytime sleepiness (CSHQ) $r=-0.01$. Actigraph minutes asleep were significantly correlated with CSHQ daytime sleepiness ($r=-0.25$, $p=0.008$). Minutes asleep (Actigraph) and total score (CSHQ) ($r=-0.25$, $p=0.009$). No significant differences for age or sex.	
Ishihara et al., 2014	N=72; 3-6 years; 37 female, 35 male; Japan	Children's ChronoType Questionnaire (CCTQ)-Japanese version	Actiwatch 2 (Philips Respironics) <i>Non dominant wrist</i> <i>60 sec epoch</i> <i>Actiware software and sleep diary to determine sleep onset/offset</i>	Free living (week day and weekend day- night time sleep only) Actiwatch worn for 7 days, concurrent parental reported sleep diary, and CCTQ completed at one time point during the week	CCTQ and Actiwatch 2: • Bedtime • Time of lights off • Sleep onset time • Sleep offset time • Wake up time • Get up time	Spearman's Correlation: Sleep wake parameters measured with CCTQ were highly correlated to those reported or measured using Actiwatch ($p<0.001$). Correlation coefficients slightly smaller on free days than on scheduled days. Sleep latency on scheduled days and free days, and	Inadequate

			<i>and to check artifacts appearing on accelerometer recordings.</i>		• Midsleep point • Duration (time in bed, sleep period, sleep latency).	sleep period on free days, the correlation coefficients were smaller than those of the remaining parameters. Scheduled/weekday results: Bedtime $r=0.774$ Sleep onset time $r=0.775$ Wake up time/sleep offset time $r=0.824$ Get up time $r=0.830$ Midsleep point $r=0.836$ Time in bed $r=0.740$ Sleep period $r=0.743$ Sleep latency $r=0.512$. Free day/weekend results: Bedtime $r=0.641$ Sleep onset time $r=0.606$ Wake up time/sleep offset time $r=0.707$ Get up time $r=0.714$ Midsleep point $r=0.786$ Time in bed $r=0.556$ Sleep period $r=0.490$ Sleep latency $r=0.380$.	
Iwasaki et al., 2010 *also examined a sleep diary	N=47; 5 years old; 21 female, 27 male; Japan	Brief questionnaire <i>12 major questions regarding child's sleep</i>	Ambulatory Monitoring <i>Wrist (dominant or non-dominant)</i>	Free living (week day and weekend day- night time sleep only)	Brief questionnaire: • Usual bedtime • Sleep onset time	Pearson's Correlations: Brief questionnaire and ambulatory monitoring: <i>Bedtime</i>	Inadequate

		<i>patterns and sleep quality based on past 4 weeks</i>	<i>60 sec epoch Data analysed using Motion Logger ActFAST Analysis software, and sleep end time using Cole algorithm</i>	Ambulatory monitoring and sleep diary completed daily for 7 days, brief questionnaire completed at the end of the study period	• Sleep end time • Sleep period • Number of night wakings (>5 minutes) • Sleep latency Ambulatory Monitoring: • Bedtime • Sleep onset time • Sleep latency • Sleep end time • Sleep period • Number of night wakings	Overall $r=0.17$, $p<0.01$ Weekday $r=0.25$, $p<0.01$ Weekend $r=0.02$ Sleep onset time Overall $r=0.39$, $p<0.001$ Weekday $r=0.49$, $p<0.01$ Weekend $r=0.07$ Sleep latency Overall $r=0.04$ Weekday $r=0.07$ Weekend $r=0.02$ Sleep end time Overall $r=0.51$, $p<0.001$ Weekday $r=0.59$, $p<0.001$ Weekend $r=0.08$ Sleep period Overall $r=0.35$, $p<0.001$ Weekday: $r=0.38$, $p<0.001$ Weekend $r=0.01$ Night wakings Overall, weekday and weekend $r=0.00$ Comparison between questionnaire and diary: Sleep onset time: Overall $r=0.38$, $p<0.001$	

						Weekday $r=0.45$, $p<0.001$ Weekend $r=0.07$ Sleep end time: Overall $r=0.53$, $p<0.001$ Weekday $r=0.64$, $p<0.001$ Weekend $r=0.18$, $p<0.01$ Sleep period: Overall $r=0.29$, $p<0.001$ Weekday $r=0.27$, $p<0.001$ Weekend $r=0.09$ Night wakings: Overall $r=0.03$ Weekday $r=0.06$ Weekend $r=0.00$	
Kushnir and Sadeh, 2013	N=30; 4-6 years; 14 female, 16 male; Israel	Brief Child Sleep Questionnaire (BCSQ) <i>Parental reported, referring to their child's sleep during past week</i>	Mini Motionlogger <i>Non-dominant wrist; 60 sec epoch; Sadeh algorithm</i>	Free living (no separate report of weekday/weekend day sleep- night time sleep only) Parents completed BSCQ based on their child's sleep during past week, then child wore Mini Motionlogger, every night for 7 days.	BCSQ: - Sleep time - Total sleep duration during the night - Number of night wakings - Total time awake during night Mini Motionlogger:	Spearman's Correlation: Sleep onset time = $r=0.76$, $p<0.0001$. Sleep period $r=0.85$, $p<0.0001$ Night wakings $r=0.13$ Wake after sleep onset $r=-0.11$ Parents underestimated sleep onset time and overestimated sleep period. Parents significantly	Inadequate

					• Sleep onset time • Sleep period • Number of night wakings (lasting ≥ 5 minutes) • Wake time after sleep onset	underestimated the number of night wakings and wake after sleep onset.	
Lucas-de la Cruz et al., 2016	N=51; 4-7 years; 48% female, 52% male; Spain	Child Sleep Habits Questionnaire (CSHQ)-Spanish version	Actisleep monitor <i>Non-dominant wrist; 60 sec epoch; Nocturnal sleep/wake measures were estimated from actigraphic data using Sadeh algorithm. Actigraph recordings were checked against the parent's diary. The analysis</i>	Free living (weekday/weekend day differences unknown) Child wore monitor on non-dominant wrist for 7 consecutive days (≥ 3 more nights to be included). Parents instructed to keep a log for time spent in bed, out of bed for each measured sleep episode.	CSHQ: • Sleep latency • Total sleep time • Awakenings • Sleep efficiency Actisleep Monitor: • Sleep onset • Sleep latency • Total sleep time • Number and duration of awakenings Sleep efficiency	Polyserial correlation coefficient: Sleep latency (actisleep) with measures on CSHQ $r=-0.318$ to $r=0.355$ Awakenings (actisleep) with measures on CSHQ $r=-0.026$ to $r=0.334$ Total sleep time (actisleep) with measures on CSHQ $r=-0.154$ to $r=0.092$ Efficiency (actisleep) with measures on CSHQ $r=0.105$ to 0.133 . *ICC for 'time in bed' reported by CSHQ-S and sleep diary = 0.54 , $p<0.001$	Inadequate

			<i>period for the sleep data presented here (the scoring interval) included the period from the bedtime to the wake time as given in the diary.</i> Sleep diary completed alongside accelerometer				
Perpétuo et al., 2020	N=34; 3-6 years; 50% female, 50% male; Portugal Mothers' education level ranged from 6 to 21 years, and fathers from 6 to 17 years. Mothers worked between 26 and 56 weekly hours	Children's Sleep Habits Questionnaire (CSHQ) <i>Parental report on typical sleep</i>	Actiwatch 2 <i>Non-dominant wrist; 60 sec epoch; Actiware 6.0.9 software was used to classify different sleep states.</i>	Free living (week day and weekend day) Children asked to wear Actiwatch 2 for 7 days, questionnaire completed at one time point by parents. Data on naps systematically collected.	CSHQ: • Bed time • Wake up time • Daily sleep time • Number and duration of night wakings • Bedtime resistance • Sleep onset delay • Sleep duration	Correlations: Total sleep time from actigraphy was correlated with CSHQ bedtime resistance ($r = -0.43$, $p < 0.05$) and sleep problem global score ($r = -0.38$, $p < 0.05$). WASO (actigraphy) was significantly and positively correlated with CSHQ night-wakings ($r = 0.44$, $p < 0.05$). More night-wakings registered by actigraphy	Inadequate

	and fathers worked between 35 and 60 hours per week.				• Night wakings • Parasomnia • Sleep apnea • Daytime sleepiness • Total sleep score Actiwatch 2 • Bedtime • Wake up time • Time in bed • Total sleep time • Onset latency • Sleep efficiency • Wake after sleep onset • Number of night wakings	were associated with parents' report of greater instability on sleep duration ($r = 0.36$, $p < 0.05$) and with more parasomnias ($r = 0.35$, $p < 0.05$). Higher actigraphy sleep efficiency was significantly and negatively correlated with parent reports on bedtime resistance ($r = -0.42$, $p < 0.05$), night-wakings ($r = -0.45$, $p < 0.01$), and global scores of sleep problems ($r = -0.45$, $p < 0.01$). Bland Altman: With the exception of wake up time, there was no agreement between parental report and actigraphy measures During the week, parents reported a mean of 7 min earlier (SD = 0:19 min) wake-up time compared to time registered by actigraphy. Difference between methods is higher for weekends, where	

						parents reported wake-up time 28 min (SD = 0:47) later than the actigraph's. For total sleep time and wake after sleep onset, the longer the child slept (and the longer the night-wakings lasted) the greater the difference between methods was ($\beta = 0.49$, $p < 0.01$ and $\beta = -0.88$, $p < 0.001$ respectively). Results showed that, for total sleep time, only 6.5% of the pairs of observations showed a difference of 30 min or less, and for 25.8% of the pairs of observations, the difference was of 1 h 30 or less. For WASO, only for 3% of the paired observations was the difference < 30 min; for 27%, the difference was < 1 hour. For all other variables, the difference was > 1 hour.	
Sekine et al., 2002	N=21; 3-4 years old; 9 girls, 12 boys; Japan	Sleep habits questionnaire <i>3 consecutive nights of parental reported sleep</i>	Actiwatch (MiniMitter) <i>Ankle of non-dominant leg; 60 sec epoch;</i>	Free living (weekday/weekend day differences unknown- night time sleep only)	Sleep habits questionnaire: • Get up time • Bedtime • Sleeping hours	Pearson's Correlation Coefficient and Paired t-tests: The correlation and difference	Doubtful

			<i>Device automatically estimates various sleep parameters according to the algorithm (after researcher entered bedtime and rising time)</i>	Questionnaire completed and Actiwatch worn for 3 consecutive days	Actiwatch • Sleep start • Time of sleep onset • Sleep end • The time of sleep termination • Assumed sleeping hours • Actual sleeping hours • Sleep latency Sleep efficiency	between the reported sleeping hours during the measurement days and assumed or actual sleeping hours were 0.90 ($p<0.001$) and 0.79 hours (95%confidence interval: 0.59-0.99) and 0.90 ($p<0.001$) and 0.92 hours (0.73-1.10), respectively. The correlation and difference between the assumed and actual sleeping hours were 0.99 and 0.12 hours (0.06-0.19), respectively.	
Werner et al., 2008 *also examined a sleep diary	N=50; 4.5 to 7.3 years; 44% female, 56% male; Switzerland Predominantly white. 50% middle class, 50% upper class. Lower socioeconomic	Sleep schedule time questionnaire (SSTQ)	Actiwatch Plus AW4 <i>Non-dominant wrist; 60 sec epoch; Sleep measures using Actiware 5, using scoring procedure described by Acebo et al. 2005</i>	Free living (week day and weekend day) SSTQ answered before an interview about child's sleep that took place in parents' homes, followed by child wearing Actiwatch and parental reported sleep diary data collected for 6-8 consecutive days	Sleep schedule time questionnaire (SSTQ) • Normal wake up time • Get up time • Bedtime • Time of lights off • Sleep latency	Bland Altman and T-tests: Actiwatch and SSTQ: A prior defined satisfactory agreement of 30 minutes was not reached for any parameters tested: sleep start, sleep end or assumed sleep. Sleep diary was also used: Sleep diary and SSTQ: A prior defined satisfactory agreement of 30 minutes was not satisfied for any of	Inadequate

	class not represented.			(although 24 hour data collected, and SSTQ asked about nap times, there was no information available regarding day time sleep)	- Potential nap duration Actiwatch Plus AW4 - Sleep start time - Sleep end time - Assumed sleep - Actual wake time - True sleep time	the 3 parameters tested: sleep start, sleep end and assumed sleep, for weekdays or weekend days No significant, sex, or socioeconomic status effects on the differences.	
Diaries							
Galland et al., 2016	N=74; 3.5 and 5 years; Male to female ratio: 0:9 (3.5 yrs) and 1:3 (5 yrs); New Zealand Mostly European (98 % of 3.5 year olds; 91% 5 year olds) or Maori (2% 3.5 year olds ;	Sleep diary <i>Parental reported (diary to be filled out by the person who is caring for the child during the measurement period- incl nursery teachers for day time naps) 2 days</i>	Actical (MiniMitter) <i>Right hip 15 sec epoch Data scored using automated script developed in MATLAB</i>	Free living (24 hour sleep – unknown if weekday / weekend day) Children wore Actical for 5-7 days, sleep diary completed for 2 of these days.	Sleep diary and Actical: - Sleep onset time - Sleep offset time 24 hr sleep duration - Overnight sleep duration - Sleep efficiency	Percent agreement: Agreement in nap time between the diary and Actical: 93% Kappa: Kappa for agreement in nap time between diary and Actical= 0.58 Bland Altman: There were no significant systematic differences between any diary and Actical sleep–wake	Doubtful

	9% 5 of year olds) Most children resided in neighbourhoods within mid-range of deprivation index (age 3.5 yrs = 4.9; 5 yrs= 4.4				• Night wakings (number and duration) • Naps (number and duration)	variables at ages 3½ or 5 years. No significant differences in number of day time naps at any age. Nap duration was significantly longer in data derived from diary compared to Actical in children 3½ years old. No naps were identified in any child at age 5.	
Iwasaki et al., 2010 *also examine a sleep questionnaire	N=47; 5 years old; 21 female, 27 male; Japan	Sleep diary <i>Parental reported daily sleep diary, when children stayed at nursery the sleep diary was recorded based on the information provided by their caretakers.</i>	Ambulatory Monitoring <i>Wrist (dominant or non-dominant); 60 sec epoch; Data analysed using Motion Logger ActFAST Analysis software, and sleep end time using Cole algorithm⁷⁰</i>	Free living (night time sleep only-weekend and weekend day reports separately) Ambulatory monitoring and sleep diary completed daily for 7 days, brief questionnaire completed at the end of the study period	Sleep diary: • Sleep onset • Sleep end time • Sleep period • Number of night wakings (>5 minutes) Ambulatory Monitoring: • Bedtime • Sleep onset time • Sleep latency	Pearson's Correlations: Sleep diary and ambulatory monitoring: <i>Sleep onset time</i> Overall r=0.85, p<0.001 Weekday r=0.79, p<0.001 Weekend r=0.94, p<0.01 <i>Sleep end time</i> Overall r=0.83, p<0.001 Weekday r=0.81, p<0.001 Weekend r=0.70, p<0.001 <i>Sleep period</i> Overall r=0.57, p<0.001	Inadequate

					• Sleep end time • Sleep period • Number of night wakings	Weekday $r=0.43$, $p<0.001$ Weekend $r=0.73$, $p<0.001$ <i>Night wakings</i> Overall $r=0.02$ Weekday $r=0.01$ Weekend $r=0.03$ Comparison between questionnaire and diary: <i>Sleep onset time:</i> Overall $r=0.38$, $p<0.001$ Weekday $r=0.45$, $p<0.001$ Weekend $r=0.07$ <i>Sleep end time:</i> Overall $r=0.53$, $p<0.001$ Weekday $r=0.64$, $p<0.001$ Weekend $r=0.18$, $p<0.01$ <i>Sleep period:</i> Overall $r=0.29$, $p<0.001$ Weekday $r=0.27$, $p<0.001$ Weekend $r=0.09$ <i>Night wakings:</i> Overall $r=0.03$ Weekday $r=0.06$ Weekend $r=0.00$	

Jones and Ball, 2014	N=18; 3 years old; Sex not reported; England, UK 12 children from high socioeconomic status group and 6 from low socioeconomic status group	Sleep diary <i>Parental reported during a typical week (when children were not ill or on holiday).</i>	Actiwatch Mini <i>Placement and epoch not reported; Proprietary software using medium sensitivity setting as recommended by user manual.</i>	Free living (24 hour sleep- weekday and weekend days) Children wore Actiwatch and parent completed sleep diary for 5 consecutive days. Data on naps systematically collected.	Sleep diary: • Bedtime • Sleep onset time • Morning wake time • Daytime nap duration • Sleep onset location Actiwatch: • Sleep onset time • Wake time • Daytime nap duration	Paired sample t-tests and Pearson's Correlations: Sleep onset time $r=0.98$, $p<0.001$ Wake time $r=0.99$, $p<0.001$ Nap duration $r=0.81$, $p<0.001$. No systematic differences between families in low and high socioeconomic status groups.	Doubtful
Lam et al., 2011	N=59; 3-5 years; 30 female, 29 male; USA 58% African American; 22% Caucasian; 5% Asian; 15% multiracial	Parental sleep diary <i>Parents (with help of day care workers for nap times) instructed to maintain sleep diary for 7 days.</i>	Motionlogger Actiwatch (MiniMitter) <i>Non-dominant wrist; 60 sec epoch; Action W software program for sleep-wake cycles, Sadeh algorithm for sleep epochs</i>	Free living (24 hour sleep – weekend and weekend day reports separately) Children wore Motionlogger Actiwatch and parents completed sleep diary for duration of 7 days.	Parental sleep diary • Nap time • Bed time • Sleep latency • Night time awakenings • Sleep end time • Sleep onset time	Paired t-tests: Parents significantly overestimated children's night time sleep in comparison with Actiwatch, on both weekend and weekday nights. Weekday and weekend nap duration did not differ significantly between diary and Actiwatch. ANOVA:	Inadequate

					Motionlogger Actiwatch • Napping sleep onset and offset • Napping duration • Night time sleep onset and offset • Duration of night time sleep • Total weekday daytime napping • Total daytime weekend napping • Total weekday night time sleep • Total weekend night time sleep	Sleep onset times similar for parental report and Actiwatch ($p=0.48$) Parents reported later sleep offset time than Actiwatch ($p=0.34$) (more so for the weekend than weekdays). Parents reported fewer night time awakenings on all nights compared to Actiwatch ($p<0.0001$).	
Tikotzky and Sadeh, 2001	N=59; 3.8 to 6.1 years; 30 female, 29 male; Israel	Sleep diary <i>Parental reported daily sleep diary</i>	AMA 32 Ambulatory Monitoring <i>Non-dominant wrist;</i>	Free living (night time sleep only-weekdays only) Child wore AMA 32 Ambulatory	Sleep diary: • Lights off time • Morning rise time	Pearson's Correlations and t-tests: Lights off time (parental report) and sleep onset time (accel) $r=0.87$, $p<.0005$	Very good

	Most children belonged to two parent families of a middle-upper socioeconomic status, highly educated and employed parents.		<i>Epoch not reported; Actigraphic Scoring Analysis program</i>	Monitoring device and, parent report child's sleep in sleep diary for 4-5 consecutive nights	• Sleep duration • Number of night wakings • Sleep quality • Duration to fall asleep • Evening sleepiness • Morning drowsiness • Sleep duration during the day AMA 32 Ambulatory Monitoring Sleep schedule measures: • Sleep onset time • Morning rising time • Total sleep duration Sleep quality measures:	Sleep latency (parental report) and sleep onset (accel) $r=0.02$ Morning rise time from both measures $r=0.87$, $p<0.0005$ Sleep duration from both measures $r=0.81$, $p<0.0005$ Night wakings from both measures $r=0.42$, $p<0.005$. Significant difference between average n of night wakings according to parental (0.45) and accelerometer (2.66), $p<0.001$. Number of night wakings explained 86% of variance in the difference between parental and accelerometer reports on night wakings, age of child explain an additional 2.6% of the variance. Sleep quality (parental report) and sleep efficiency (accel) $r=-0.28$, $p<0.05$	

					• Sleep percent (sleep efficiency) • True sleep time • Number of night wakings	Sleep quality (parental report) and true sleep time $r=-0.12$ Sleep quality (parental report) and night time wakings $r=0.35$, $p<0.01$	
Werner et al., 2008 * also examine a sleep questionnaire	N=50; 4.5 to 7.3 years; 44% female, 56% male; Switzerland Predominantly white; 50% middle class, 50% upper class. Lower class not represented.	Sleep diary <i>Parental reported sleep diary</i>	Actiwatch Plus AW4 <i>Non-dominant wrist; 60 sec epoch; Sleep measures using Actiware 5, using scoring procedure described by Acebo et al. 2005</i>	Free living (week day and weekend day) SSTQ answered before an interview about child's sleep that took place in parents' homes, followed by child wearing Actiwatch and parental reported sleep diary data collected for 6-8 consecutive days (although 24 hour data collected, and SSTQ asked about nap times, there was no information available regarding day time sleep)	Sleep diary • Sleep start • Sleep end • Assumed sleep • Nocturnal wake time • Actual sleep time Actiwatch Plus AW4 • Sleep start time • Sleep end time • Assumed sleep • Actual wake time • True sleep time	Bland Altman and T-tests: Sleep diary and Actiwatch A priori, satisfactory agreement= differences > 30 minutes. This requirement was satisfied for 3 of our 5 parameters: sleep start (± 28), sleep end (± 24) and assumed sleep (± 32). This requirement was not satisfied for actual sleep time (± 72) and nocturnal wake time (± 55). For weekdays and weekend days, the differences were larger.	Inadequate

						Questionnaire was also used: Sleep diary and SSTQ: A prior defined satisfactory agreement of 30 minutes was not satisfied for any of the 3 parameters tested: sleep start, sleep end and assumed sleep, for weekdays or weekend days. No significant age, sex, or socioeconomic effects on any of these differences.	

4.5.6 Reliability

4.5.6.1 Test-retest

Two studies assessed the test-retest reliability of two different questionnaires (Ishihara, Doi and Uchiyama, 2014; Lucas-de la Cruz et al., 2016). The CCTQ (Japanese version) completed twice two weeks apart and the CSHQ (Spanish version) completed twice within three weeks. The questionnaires showed high correlations between administrations; CCTQ ($r=0.90$) (Ishihara, Doi and Uchiyama, 2014) and the CSHQ, particularly for sleep duration, ($r=0.81$). Night awakenings showed weak test-retest ($r=0.56$) (Lucas-de la Cruz et al., 2016).

Table 4.8. Detailed information of studies examining test-retest reliability

Study	Study Sample	Measurement tool under study (for reliability)	Methods	Test-retest results	Risk of bias rating
Ishihara et al., 2014	Japan; N=346; 3-6 years; 158 female, 188 male	Children's ChronoType Questionnaire (CCTQ)- Japanese version	Fill out the questionnaire two weeks apart.	Pearson's Correlation coefficient: $r=0.90$ ($p<0.001$)	Inadequate
Lucas-de la Cruz et al., 2016	Spain; N=52; 4-7 years; Sex of participants unknown	Child Sleep Habits Questionnaire (CSHQ)- Spanish version	Questionnaire completed twice in a three week interval.	Intraclass correlation coefficient Bedtime resistance: 0.75, $p<0.001$; sleep duration 0.78, $p<0.001$; sleep anxiety 0.79, $p<0.001$; night wakings 0.56, $p<0.001$; parasomnias 0.73, $p<0.001$; sleep disordered breathing 0.76, $p<0.001$; daytime sleepiness 0.54, $p<0.001$ Pearson's correlation coefficient Bedtime resistance: 0.75, $p<0.001$; sleep onset delay 0.69; sleep duration 0.81, $p<0.001$; sleep anxiety 0.79, $p<0.001$; night wakings 0.56, $p<0.001$; parasomnias 0.74, $p<0.001$; sleep disordered breathing 0.77, $p<0.001$; daytime sleepiness 0.60, $p<0.001$	Inadequate

4.5.6.2 Internal Consistency

Three studies examined the internal consistency of the questionnaires, results showed: moderate for the CSHQ (Chinese version) ($\alpha=0.67$) (Tan et al., 2018) and high for the CSHQ (Spanish version) ($\alpha=0.81$) (Lucas-de la Cruz et al., 2016) and CCTQ (Japanese version) ($\alpha=0.77$) (Ishihara, Doi and Uchiyama, 2014).

The newly devised four factor structure for the MicroMini Motionlogger accelerometer (sleep activity, sleep variability, sleep timing and sleep duration, which could categorise accelerometer outcome variables) were determined to have high internal consistency (ranging from $\alpha = 0.89$ to 0.95) (Staples et al., 2019).

Table 4.9. Detailed information of studies examining internal consistency

Study	Study Sample	Measurement tool under study (for reliability)	Methods	Internal consistency results	Risk of bias rating
Ishihara et al., 2014	Japan; N=346; 3-6 years; 158 female, 188 male	Children's ChronoType Questionnaire (CCTQ)- Japanese version	Parents completed the Japanese version of the CCTQ	Cronbach's alpha (α) Overall scale = 0.77, individual items ranged from 0.72 to 0.76.	Very good
Lucas-de la Cruz et al., 2016	Spain; N=286; 4-7 years;	Child Sleep Habits Questionnaire (CSHQ)- Spanish version	Parents completed the Child Sleep Habits Questionnaire (CSHQ)- Spanish version	Cronbach's alpha (α) Overall scale = 0.81. Bedtime resistance = 0.65 Sleep duration = 0.81 Sleep anxiety = 0.74 Night wakings = 0.69	Very good

				Parasomnias = 0.62 Sleep disordered breathing = 0.66 Daytime sleepiness = 0.60	
Tan et al., 2018	N=171; 4-5 years; 53% female, 47% male; China Chinese kindergarteners Most parents worked full-time. Over 80% of mothers and 85% fathers had at least a college level education. Over half of the mothers and 80% fathers made 5000 or more monthly.	Children's Sleep Habits Questionnaire (CSHQ)-Chinese version	Parents completed the Chinese version of the CSHQ-C	Cronbach's alpha (α) Overall scale = 0.67 Bedtime resistance = 0.59 Daytime sleepiness = 0.62 Night wakings = 0.11 Parasomnias = 0.58 Sleep anxiety = 0.54 Sleep disordered breathing = 0.36 Sleep duration = 0.65	Very good
Staples et al., 2019	N=107 (at 3 years), n=104 (at 3.5 years); 3-3.5 years; Sex not reported for this age category;	MicroMini Motionlogger Accelerometer; Wrist worn; Scored using Action W-2 software and Sadeh algorithm	Accelerometer worn for 7 days at three time points in longitudinal study (2.5 years, 3 years, 3.5 years). Principal component analysis to	Cronbach's alpha (α) Sleep activity Aged 3 years: 0.93 Aged 3.5 years: 0.93 Sleep variability Aged 3 years: 0.89 Aged 3.5 years: 0.90	Very good

	USA Information on parental education, ethnicity and socioeconomic status was available but for the whole sample, not separated for age groups of interest. Predominantly college educated parents, European American, high SES.	<i>Sleep diary also completed to record bedtime, night awakenings, rise times and other issues relating to child's sleep.</i>	determine composites.	Sleep timing Aged 3 years: 0.95 Aged 3.5 years: 0.95 Sleep duration Aged 3 years: 0.90 Aged 3.5 years: 0.92	

4.5.7 Feasibility

No studies explicitly examined the feasibility of the measurement tools through qualitative research. However, as per previous reviews (Evenson, Goto and Furberg, 2015), we included missing data, non-completion and other indicators of feasibility of the measure (such as completion time).

Twelve of the studies reported information that provided an indication on the feasibility of the measurement tools. This primarily consisted of missing data from the accelerometer, either when used as the comparison tool (Duraccio et al., 2018; Ishihara, Doi and Uchiyama, 2014; Jones and Ball, 2014; Kushnir and Sadeh, 2013; Lam et al., 2011; Lucas-de la Cruz et al., 2016; Perpétuo, Fernandes, and Veríssimo, 2020; Sekine et al., 2002; Werner et al., 2008) or tool under study (Meltzer et al., 2015; Staples et al., 2019) for reasons including: technical problems, refusal to wear the device, or a lack of available valid data. One study reported that although the accelerometers were tolerated by the children, the cost of such devices (and associated licensed software required) was expensive (Iwasaki et al., 2010).

One study reported that completion of the CSHQ (Spanish version) took 4-6 minutes (Lucas-de la Cruz et al., 2016). Several studies reported missing data from questionnaires either through non-completion or the questionnaire not being completed correctly (Ishihara, Doi and Uchiyama, 2014; Lucas-de la Cruz et al., 2016; Perpétuo, Fernandes, and Veríssimo, 2020).

Table 4.10. Detailed information on studies assessing feasibility of tools

Study	Study Sample	Measurement tool (<i>Measurement instrument under study or comparison tool</i>)	Feasibility information
Duraccio et al., 2017	N=108; 4-6 years; 42% female; 53% male; 5% missing USA 86% Caucasian, 2% Hispanic, >1% Native American, 2% Asian, 5% Other, 5% Missing Mean monthly income = \$5,693.73 (Likely high socioeconomic status as recruited through universities)	Accelerometer- Actigraph GT3X+ (<i>comparison tool</i>)	3% missing data due to participants missing 2 (out of 3) nights of data.
Ishihara et al., 2014	N=72 (validity), n=346 (reliability); 3-6 years; 37 female, 35 male; (validity), 158 female, 188 male (reliability); Japan	Children's ChronoType Questionnaire (CCTQ)- Japanese version (<i>measurement instrument under study</i>)	3% missing data on test-retest due to missing at least one question on the questionnaire
		Accelerometer- Actiwatch 2 (<i>comparison tool</i>)	4% missing from validation study due to missing data on accelerometer from participant non-adherence
Iwasaki et al., 2010	N=47; 5 years old; 21 female, 27 male; Japan	Accelerometer -Ambulatory monitoring (<i>comparison tool</i>)	1 child (out of 48 recruited) excluded prior to the study due to still taking day time naps.

			Actigraphy ultimately tolerated by all participants- data successfully collected for all children included in study. Cost of devices were expensive: Accelerometers costs more than US\$25,000 to introduce a system with licensed software and 15 devices. High cost resulting in more likely use of questionnaire/diary based sleep assessment tools
Jones and Ball, 2014	N=18; 3 years old; Sex not reported; England, UK 12 children from high socioeconomic status group and 6 from low socioeconomic status group	Accelerometer -Actiwatch (<i>comparison tool</i>)	49% missing data due to (17 out of 35) children refusing to keep Actiwatch on.
Kushnir and Sadeh, 2013	N=30; 4-6 years; 14 female, 16 male; Israel	Accelerometer -Mini Motionlogger (<i>comparison tool</i>)	Compliance problems and technical issues reported- actigraphy data were available for ≥ 4 nights in 88% percent of the children. For the rest of the sample, actigraphy data were available in 4.5%, 3%, and 4.5% for 1, 2, or 3 nights, respectively. However, it's not possible to distinguish between whether these reports were related to clinical or healthy sample of children.

Lam et al., 2011	N=59; 3-5 years; 30 female, 29 male; USA 58% African American; 22% Caucasian; 5% Asian; 15% multiracial	Accelerometer -Motionlogger Actiwatch (comparison tool)	8% missing data due to noncompliance with accelerometer (=5 children dropped out of the study due to noncompliance with accelerometer).
Lucas-de la Cruz et al., 2016	N=51 (validity); 4-7 years; 48% female, 52% male; Spain	Children's Sleep Habits Questionnaire (CSHQ)-Spanish version (Measurement instrument under study)	40% did not complete CSHQ correctly on first administration Completion of questionnaire takes 4-6 minutes
		Accelerometer -ActiSleep (comparison tool)	33% did not wear accelerometer for required time
		Both tools (Measurement instrument under study and comparison tool)	16% excluded from validity analysis due to not completing all measures (CSHQ, ActiSleep, sleep diary) - 61 parents completed second administration, of which 51 had appropriate data for all three measures.
Meltzer et al., 2015	N=14 3-5 years; Sex not reported for this age category; USA	Fitbit Ultra (measurement instrument under study)	19% missing data due to technical issues with Fitbit Ultra. However, it's not possible to distinguish between the different age

	Ethnicity reported for whole sample, but not separately by age category.		groups for the missing data (full sample = children aged 3-17 years)
		Polysomnography (<i>comparison tool</i>)	3% missing data due to corrupt polysomnography files However, it's not possible to distinguish between the different age groups for the missing data (full sample = children aged 3-17 years)
Perpétuo et al., 2020	N=34; 3-6 years; 50% female, 50% male; Portugal Mothers' education level ranged from 6 to 21 years, and fathers from 6 to 17 years. Mothers worked between 26 and 56 weekly hours and fathers worked between 35 and 60 hours per week.	Children's Sleep Habits Questionnaire (<i>comparison tool</i>)	20% missing data due to non-completion
		Accelerometer -Actiwatch 2 (<i>comparison tool</i>)	11% missing data due to not having valid data
		Both tools (<i>Measurement instrument under study and comparison tool</i>)	26% participants had missing data for both measurement tools Prior to measurement- 17% families dropped out for various reasons (changed residence or withdrew consent).
Sekine et al., 2002	N=21; 3-4 years old; 9 girls, 12 boys; Japan	Accelerometer -Actiwatch (<i>comparison tool</i>)	12% missing data from accelerometer (1 child reluctant to wear instrument, data from 2 other children were not successfully measured).

Staples et al., 2019	N=107 (at 3 years), n=104 (at 4.5 years); 3-3.5 years; Sex not reported for this age category; USA Information on parental education, ethnicity and SES was available but for the whole sample, not separated for age groups of interest. Predominantly college educated parents, European American, high SES.	Accelerometer -MicroMini Motionlogger (<i>measurement instrument under study</i>)	At age 3 years: 26% missing data (planned missingness) 5% accelerometer failure 6% not available to participate At age 3.5 years: 18% missing data (planned missingness) 7% accelerometer failure 4% not available to participate
Werner et al., 2008	N=50; 4.5 to 7.3 years; 44% female, 56% male; Switzerland Predominantly white. 50% middle class, 50% upper class SES. Lower SES class not represented.	Accelerometer -Actiwatch Plus AW4 (<i>comparison tool</i>)	26% total missing data: 13% insufficient language skills/deregistration, remaining 13% accelerometer problems, including: 4% refusal to wear accelerometer; 6% technical problems; 3% accelerometer loss

4.5.8 Generalisability of results and health inequality implications

4.5.8.1.1 Ethnicity

Studies that reported the ethnicity of the included children had samples who were predominantly Caucasian (Duraccio et al., 2018; Galland et al., 2016; Jones and Ball, 2014; Werner et al., 2008), Hispanic/Latino (Tracy et al., 2021), Chinese (Tan et al., 2018) and African American children (Lam et al., 2011).

4.5.8.1.2 Socioeconomic Status

Studies that reported the SES of participants described the families of the children as mainly middle-upper class (Duraccio et al., 2018; Jones and Ball, 2014; Tikotzky and Sadeh, 2001; Werner et al., 2008), with the exception of one study that reported that children lived in areas within 'mid-range' of deprivation (Galland et al., 2016). Two studies reported the educational level and working hours of parents of the included children (Perpétuo, Fernandes, and Veríssimo, 2020; Tan et al., 2018). One study reported no systematic differences in the outcomes of the diary and accelerometry between families in low and high SES groups (Jones and Ball, 2014).

Table 4.11. Summary table of the validity and reliability of measurement tools for assessing sleep of pre-school aged children

Measurement instrument under study	Validity							Reliability		Risk of bias assessment (Inadequate, Doubtful, Adequate, or Very good)	Reference to the included studies
	Sleep duration	Sleep onset time	Bedtime	Sleep end/ wake up time	Night wakings	Sleep latency	Nap time	Test-retest reliability	Internal consistency		
Criterion validity											
Accelerometers											
Fitbit Ultra (normal mode)										Inadequate	Meltzer et al., 2015
Convergent validity											
Accelerometers											
Actigraph GT3X+										Doubtful	Tracy et al. 2021
MicroMini Motionlogger										Very good	Staples et al., 2019
Questionnaires											
Children's sleep habits questionnaire (English)										Inadequate	Duraccio et al., 2017
Children's sleep habits questionnaire (English)										Inadequate	Perpétuo, Fernande and Veríssimo, 2020

Children's sleep habits questionnaire (Spanish)										Inadequate (convergent and test-retest)	Lucas-de la Cruz et al., 2016
										Very good (internal consistency)	
Children's Sleep Habits Questionnaire (Chinese)										Very good	Tan et al., 2018
Children's ChronoType Questionnaire (Japanese)	 			 						Inadequate (convergent and test-retest)	Ishihara et al., 2014
										Very good (internal consistency)	
Brief 12 item questionnaire										Inadequate	Iwasaki et al., 2010
Brief Child Sleep Questionnaire										Inadequate	Kushnir and Sadeh, 2013
Sleep habits questionnaire										Doubtful	Sekine et al., 2002
Sleep schedule time questionnaire										Inadequate	Werner et al., 2008
Diaries											
Sleep diary (5 min intervals)										Doubtful	Galland et al., 2016

Sleep diary (daily information)	 ²  ³	 ²  ³		 ²  ³						Inadequate	Iwasaki et al., 2010
Sleep diary (daily information)										Doubtful	Jones and Ball, 2014
Sleep diary (daily information)										Inadequate	Lam et al., 2011
Sleep diary (daily information)										Very good	Tikotzky and Sadeh, 2001
Sleep diary (15 min intervals)										Inadequate	Werner et al., 2008

*The methods used to assess the **rating of the measurement tool** are based on the interpretation of statistics detailed in Table 4.3 and are indicated in this summary table as: Good = ● Moderate=● Weak=●. These ratings do not include the risk of bias assessment; this is included in the final column to demonstrate the quality of the evidence of each study. Ratings of measurement tools assessed as low risk of bias are highlighted in grey cells.

** **Risk of bias** may be present if the overall quality assessment is doubtful or inadequate, or if there is only one study where the risk of bias is quality assessed as adequate.

¹ Representative of all subscales apart from night awakenings and daytime sleepiness which had weak ratings; ² Weekdays; ³ Weekend days

This table shows a summary of the results of studies where they assessed the criterion validity (in comparison with polysomnography), convergent validity (in comparison with accelerometry), test-retest reliability and internal consistency of the tools. This table does not include results on the construct (structural) validity or feasibility of the tools, due to the nature of these studies resulting in the information not being able to be presented in such a way. Please refer to the text for a narrative overview of studies that assessed these measurement properties.

4.6 Discussion

This review is the first to examine the measurement properties of tools used to assess sleep of pre-school aged children. The majority of studies (13/16) that met the inclusion criteria examined questionnaires and diaries, with minimal studies (3/16) on the validity of accelerometry to measure sleep in pre-school aged children. The limited evidence base is congruent with previous research reporting that the measurement properties of sleep measurement tools are often not assessed (Erwin and Bashore, 2017; Lewandowski, Toliver-Sokol and Palermo, 2011; McDowall et al, 2017; Spruyt and Gozal, 2011).

Measurement tools should have appropriate measurement properties, including being valid and reliable, alongside being feasible in the target population, however this is often not the case in studies assessing sleep of young children (Chaput et al., 2017b; Janssen et al., 2020). Previous development and evaluation of measurement tools used to assess sleep of children has predominantly focused on children aged 6 years and older (Sen and Spruyt, 2020) and, therefore, studies assessing sleep in pre-school aged children, rarely use valid and reliable methods (Chaput et al., 2017b; Janssen et al., 2020). The conclusions made from research studies implementing tools with either poor or unknown measurement properties may therefore be biased and invalid (Spruyt and Gozal, 2011).

Most studies included in the review assessed the measurement properties of questionnaires and diaries, primarily through convergent validity, using accelerometers as the comparison measure. Only one questionnaire (CSHQ) was assessed in multiple studies. There were very few studies that assessed the ability of the accelerometers to detect sleep specifically in the pre-school age range, with only one study examining the criterion validity of the Fitbit Ultra (Meltzer et al., 2015). The Fitbit Ultra was not used as a comparison tool in

any other study. This raises concerns about the results of studies when the measurement properties and accuracy of the tool being used as the comparison are unknown. It is critical that accelerometers are validated against a criterion method of polysomnography and/or videosomnography in the pre-school aged population prior to accelerometers being used as a comparison tool for validating further tools. This would require expensive and intensive research methodology, which may not always be feasible.

Overall, based on the current very limited evidence base, the '*subjective daily information*' reported in the sleep diary proposed by Tikotzky and Sadeh (2001) appeared most accurate for assessing sleep duration, and was based on a study with low risk of bias. However, this conclusion is from the results of one study only. The *sleep habits questionnaire*, BCSQ and CCTQ were concordant with accelerometry for assessing sleep duration, with the CCTQ also demonstrating good reliability (Ishihara, Doi and Uchiyama, 2014). However, these studies reported night time sleep only, were based on the results of one study each, and were determined to have high risk of bias. There were three diaries that directly assessed the outcome of day time sleep, all of which showed good accuracy (Galland et al., 2016; Jones and Ball, 2014; Lam et al., 2011).

All measurement tools assessed for the outcomes of sleep latency, night awakenings or wake after sleep onset were shown to be poor at determining these factors, with the exception of the Fitbit Ultra for measuring wake after sleep onset, using the normal mode only (Meltzer et al., 2015). This study was of inadequate quality, which suggests that at present, there is insufficient evidence to provide a conclusion on which measurement tools would be applicable to determine the sleep quality of pre-school aged children. Parental reported tools showing poor accuracy for determining night awakenings of young children

has been highlighted previously (Bauer and Blunden, 2008). Suggested reasons for this include that children of this age may stop signalling their parents if they wake during the night (Kushnir and Sadeh, 2013) and also that accelerometers (used as comparison methods) overestimate night awakening (Meltzer et al., 2012a; Sitnick, Goodlin-Jones and Anders, 2008). This potential bias is important when inferring the accuracy of parental reported tools.

Sleep onset and end time consistently showed higher convergence when reported from the diaries (Iwasaki et al., 2010; Jones and Ball, 2014; Tikotzky and Sadeh, 2001; Werner et al., 2008) than questionnaires (Iwasaki et al., 2010; Kushnir and Sadeh, 2013; Perpétuo, Fernandes, and Veríssimo, 2020; Werner et al., 2008). This is unsurprising given that in most instances diaries and accelerometry were being compared simultaneously, whilst questionnaires and accelerometers were not reflective of the same time frame. For example, questionnaires were measuring the week before the period of accelerometer wearing (Duraccio et al., 2018; Kushnir and Sadeh, 2013), or the questionnaire was measuring ‘typical’ sleep, whilst the accelerometer was measuring ‘in the moment’ sleep (Lucas-de la Cruz et al., 2016; Perpétuo, Fernandes, and Veríssimo, 2020; Werner et al., 2008). Research has demonstrated that the sleep patterns of pre-school age children vary even within a single week (Acebo et al., 1999; Taylor et al., 2015). As such, when comparing measurement tools it would be important to ensure they are representative of the same time frame – as the fluctuations and variations in sleep may be incorrectly attributed to the measurement tool being less accurate.

Previous research has suggested that reporting of sleep duration may be more accurate for weekday nights than weekend nights (Nascimento-Ferreira et al., 2016). There

were differences in accuracy of reporting between weekday and weekend days in this review (Ishihara, Doi and Uchiyama, 2014; Iwasaki et al., 2010), but there were no consistent patterns in terms of which days were more accurately reported. Although the reasons for this are unknown, the differences highlight the importance of assessing both weekday and weekend day sleep in research.

Accelerometers were used as the comparison tool in the majority of the studies. However, the types of accelerometer, placement, epochs, algorithms and procedures to detect sleep and wake varied between studies. The algorithms used to assess sleep parameters when using accelerometers (Duraccio et al., 2018; Iwasaki et al., 2010; Kushnir and Sadeh, 2013; Lam et al., 2011; Lucas-de la Cruz et al., 2016) were based on algorithms devised for adolescent and adult samples (Cole et al., 1992; Sadeh, Sharkey and Carskadon, 1994), despite known differences in the sleep of individuals of different ages (Grigg-Damberger et al., 2007). This review found that a pre-school specific algorithm was more accurate at detecting bedrest and wake time in comparison with visual identification of the data, and outcomes differed when compared with existing algorithms frequently used (Tracy et al., 2021). This highlights the importance of the data processing decisions when using accelerometers to detect sleep of pre-school aged children, and the need for device and age specific algorithms to improve accuracy.

Diaries were often used alongside accelerometers to indicate at least bed and wake time. This is common practice, as accelerometers cannot distinguish sleep from other low energy behaviours such as sedentary behaviour (Meltzer et al., 2012b; Sadeh, 2011). The use of a diary alongside an accelerometer has been shown to increase accuracy for measuring sleep (Tétreault et al., 2018). Accelerometers infer sleep based on the absence of

movement, rather than being a direct measure of sleep (Sadeh, 2015). Additionally, accelerometer output data is heavily reliant on subjective data interpretation choices (Migueles et al., 2017). As such, arguably, accelerometry should not be defined as an 'objective' measure of sleep, as is often the case within the literature.

There is limited research on the feasibility of measurement tools used to assess sleep of pre-school aged children. There were few reports on the feasibility of the measurement tools and no studies directly assessed feasibility qualitatively. Measurement tools are only applicable for use when feasible in the population in which they are to be used, as such, feasibility should be given as much attention as validity and reliability during tool development and evaluation. The majority of included studies reported missing data for the accelerometer, either when this was the tool under study or comparison tool. This is important feasibility information to note, as device malfunction can result in whole datasets being disregarded (Fisher, Krietsch and Janicke, 2019). Future research should ensure that feasibility is assessed. In particular, more qualitative work exploring the acceptability and feasibility of measurement tools is warranted to truly understand perceptions of the tools.

Further, the content validity of the included tools was unknown as no studies commented on the development process of the proxy reported tools (Terwee et al., 2018a). Further qualitative research with parents and carers of pre-school aged children during development of proxy report based measurement tools is recommended to ensure the items of the tool, and the tools themselves, are relevant and comprehensive to the construct, population and context of use (Terwee et al., 2018a).

Disparities in sleep behaviours of young children based on ethnicity, income, and mother's level of education have been noted (Graham et al., 2020). Recent research has

shown that parents of Hispanic children were more accurate at estimating their child's sleep duration and wake time than parents of White children (Li et al., 2021). This may be explained by different sleep habits (e.g. higher rates of bed sharing in the Hispanic families) (Li et al., 2021). This demonstrates the importance of ensuring tools are developed and evaluated with the population of interest. However, the studies included in this review showed that the literature is skewed towards middle-upper class, Caucasian populations. There was very limited evidence to suggest that measurement properties of the tools do not differ between low and high SES groups, with only one study explicitly examining and reporting on this. Additionally, studies included in this review were all conducted in high income countries, it is important for sleep measurement research with pre-school aged children in lower and middle income countries (Nascimento-Ferreira et al., 2016).

4.6.1 Strengths and limitations of review, including limitations on the methodology used

A strength of this review is that it is the first of its kind, to collate and summarise the evidence base on the measurement properties of the full range of tools used to assess sleep of pre-school children. Although termed 'rapid' in line with definitions of reviews, extensive searching was conducted on multiple platforms for international evidence on the topic.

Children's sleeping patterns change rapidly in the early years of life (Hirshkowitz et al., 2015a; Hirshkowitz 2015b). Although we included a broad age range in the current review to ensure inclusion of children who have not yet reached the age of formal schooling internationally (World Bank Data, 2020), only two studies included children aged 7 in their samples (Lucas-de la Cruz et al., 2016; Werner et al., 2008). The findings of this review are only applicable to children between 3-7 years, and do not provide information on how to capture sleep of infants or toddlers. In line with this, there were some studies that fell just

outside of the age range for inclusion in this review. The inclusion criteria were strict meaning that studies that included children outside of this age range were excluded, in line with previous reviews (Carson et al., 2015). However, the studies that fell just outside of this age bracket (e.g. LeBourgeois and Harsh, 2016; Bélanger et al., 2013) may have provided some additional interesting insights.

The main study limitation is the potential bias introduced by having only one reviewer to screen the studies and conduct the data extraction and risk of bias of included studies (Ganann, Ciliska and Thomas, 2010). Rapid reviews, even when conducted using systematic procedures such as the review outlined in this chapter, contain more biases than traditional systematic reviews. These include that there is a higher likelihood that potentially relevant studies are missed when only one person is conducting screening of studies. Additionally, a rapid review is more prone to errors in interpretation of studies and reporting due to only one person conducting the data extraction and risk of bias assessment (Ganann, Ciliska and Thomas, 2010). As a team, we tried to overcome these barriers by any uncertainty on study eligibility resolved through consultation with a second reviewer. Further, a second reviewer checked the outcomes of the data extraction and risk of bias assessment against the original studies.

Similar to the review outlined in *Chapter 3*, it was not possible to conduct a meta-analysis on the included studies due to the heterogeneity of studies including the measurement tools being examined and comparison measurement tools used, and the lack of multiple studies assessing the same measurement tool. This means that only a narrative synthesis approach could be used and it was not possible to provide an overall outcome based on combining statistics from multiple studies (Deeks, Higgins, and Altman, 2022).

4.6.2 Implications and recommendations for future research

This review has important implications for the measurement of sleep moving forward, particularly due to the presence of sleep in public health discourse and in global recommendations for pre-school aged children (WHO, 2019). Firstly, this is the first available review that has collated the evidence on the full range of measurement tools used to assess sleep specifically in the pre-school population. This offers a useful resource both in the conduct and interpretation of studies that assess sleep. However, the review demonstrates some stark findings on the lack of evaluation of tools used to assess sleep within this population; despite multiple tools frequently being used within the literature to draw conclusions on associations between sleep and a range of health and developmental outcomes. This has wider implications for policy makers when developing policy in this area, as it is important to ensure that policy is being developed based on high quality evidence (and the measurement tool of choice can impact on the outcomes of research, in which policy is then developed). This highlights that there needs to be more in-depth evaluation work on measurement tools in this field.

Further to this, accelerometers are increasingly used to assess sleep of pre-school children, however in this review only one eligible study was identified that assessed the use of an accelerometer (Fitbit Ultra) alongside polysomnography (Meltzer et al., 2015). The lack of evidence on accelerometers for assessment of sleep of pre-schoolers has wider implications for research for several reasons. Firstly, the outcomes of studies using accelerometers to detect sleep must be viewed with caution if it is unknown if the tool is measuring sleep accurately. Second, it is difficult to truly know the validity of other

measurement tools (primarily questionnaires and diaries) that are validated against accelerometry, when the validity of the accelerometer is not known within this age group.

Quality tools with known measurement properties are needed both to develop an appropriate evidence base and to effectively monitor and evaluate sleep at population level. This review highlights clear gaps that must be addressed including:

1. Evaluation of the validity, reliability, and feasibility of accelerometry for the use of sleep measurement in pre-school aged children, including the data processing decisions, age and device specific algorithms, and placement. There is a particular need for validation against criterion methods including polysomnography and videosomnography.
2. Qualitative evaluation of the feasibility and acceptability of measurement tools used to assess sleep of pre-school aged children.
3. Representation of lower SES and ethnically diverse groups in research exploring measurement properties of sleep measurement tools.

4.7 Conclusion

This review highlights the scarcity of evidence exploring the measurement properties of tools used to examine the sleep of pre-school aged children and some clear gaps in knowledge. At present, one sleep diary was shown to be valid for assessing sleep duration (based on one study only). Most studies had fundamental flaws in quality impacting on the ability to make confident conclusions. The review demonstrated that the literature is heavily skewed towards inclusion of middle-upper class, Caucasian samples, with limited and insufficient evidence to conclude which tools are most suitable for low SES and ethnic minority groups. There is a need for further evaluation of measurement tools used to assess

the sleep of pre-school aged children, including: 1) Evaluation of the validity and reliability of accelerometers, to improve the quality of studies assessing measurement properties of questionnaires and diaries; 2) Assessing the qualitative feasibility of all measurement tools; 3) Developing tools, or examining the content validity of the tools, with the target population prior to use.

4.8 [Summary and link to next chapter](#)

The findings of this review provide information on which measurement tools have reasonable measurement properties (namely validity) for examining the sleep of pre-school children. The review also demonstrated the lack of evidence on the acceptability and feasibility of measurement tools, which it is important to examine and understand to ensure that a tool can be successfully used. As such, this is the focus of the next part of this PhD, which was to assess the acceptability and feasibility of a range of measurement tools from the perspective of parents and carers, described in *Chapter 5*. In line with *Chapter 3*, this review also highlighted that there is a need for more in-depth development and evaluation work of proxy reported tools, with the target population, which will be addressed in *Chapter 6* of this thesis.

5 Chapter Five: Qualitative acceptability and feasibility of tools used to measure movement behaviours of pre-school aged children

The study in this chapter has been published in the International Journal of Environmental Research and Public Health:

Phillips, SM., Summerbell, C., Hesketh, K., Saxena S., and Hillier-Brown, FC. (2022). Parental views on the acceptability and feasibility of measurement tools used to assess movement behaviours of pre-school children: a qualitative study. *International Journal of Environmental Research and Public Health*, 19, 3733.

<https://doi.org/10.3390/ijerph19063733>

5.1 Prelude

In the endeavour to assess the best way of measuring movement behaviours of pre-school children at scale, knowing about the measurement properties of validity and reliability is important, but not sufficient. It is also important to know about the acceptability and feasibility of tools, to ensure that even if a tool is accurate, it is still something that people would like and agree to use. As highlighted in the previous two chapters (*Chapters 3 and 4*), there is minimal evidence on the acceptability and feasibility of measurement tools in pre-school children, with a particular lack of evidence assessing tools in families from lower SES groups or across different ethnic groups. On this basis, we conducted a study to fill this knowledge gap, to determine the acceptability and feasibility of measurement tools that are used to assess movement behaviours of pre-school children, deliberately recruiting families from more deprived areas in the UK, including from one of the most ethnically diverse areas in the UK.

5.2 Introduction

This chapter outlines a qualitative study that aimed to examine the acceptability and feasibility of measurement tools used to assess movement behaviours of pre-school aged children, from the perspective of parents and carers. The two reviews conducted to assess the evidence base on the validity, reliability and feasibility of tools (*Chapter 3 and Chapter 4*) demonstrated that there is limited research examining the acceptability and feasibility of measurement tools used to examine movement behaviours of pre-school aged children. Although the key measurement properties (validity, reliability) are fundamental when selecting a measurement tool, the acceptability (whether individuals are willing to do something) and feasibility (whether individuals are able to do something) of a tool may

determine success of the study (De Vet et al., 2011; Dowd et al., 2018). The feasibility and acceptability of tools can affect retention in research studies and the amount of usable data, including through low compliance with protocols (Barber et al., 2016; Robinson et al., 2019; Nyström et al., 2020). Further, in some instances, the use of tools has been suspended during data collection for these reasons (Hossain et al., 2021; Okely et al., 2021b).

The reviews highlighted that studies examining the feasibility or acceptability used survey based research, whereby parents or researchers would rate sentences such as '*the device was uncomfortable to wear*' (Bar-Or et al., 1996; Cardon and De Bourdeaudhuij, 2007; Davies et al., 2012; De Craemer et al., 2021; Van Cauwenberghe et al., 2012) or through researcher observation (Shin, 2015). Whilst these studies are informative in inferring the practical acceptability of devices, there is a need for more qualitative research to understand how different tools are perceived and accepted in the target population to ensure they are appropriate and accessible prior to use (Kowalski et al., 2012; Lancaster, Dodd and Williamson, 2004; Summerbell et al., 2012; Thabane et al., 2010). This is particularly key to ensure that tools are equally useful across all levels of deprivation, and that research will not increase inequalities in evidence through participation in research being with more advantaged population samples only (Hillier, Pedley and Summerbell, 2011; Love, Adams and van Sluijs, 2017). In particular, this is important among low SES and ethnic minority groups, for whom a high proportion of public health interventions are targeted (Hillier-Brown et al., 2014).

There are specific challenges associated with measuring movement behaviours of pre-school children, including the size and weight of device-based tools and increased likelihood of measuring equipment being played with (Cliff, Reilly and Okely, 2009). Cognitive

capabilities prevent young children from performing detailed recall of their own movement behaviours, resulting in the need for proxy (parental/carer) reports (Atkin et al., 2012; Sallis, 1991). A greater understanding of the acceptability and feasibility of possible measurement tools is needed to further develop the evidence base and understanding of the movement behaviours of pre-school children (Aubert et al., 2021). As the importance of concurrently assessing these behaviours grows, there is a need to address measurement issues and practical considerations by looking at how various tools will be applicable for 24-hour measurement of movement (Kuzik et al., 2017; Rosenberger et al., 2019; Quante et al., 2015).

To my knowledge, only one previous study of this kind in young children exists, whereby focus groups were conducted with mothers of toddlers (aged 2-3 years), to determine the qualitative feasibility of accelerometers used to assess PA of toddlers and their parents. This study identified issues that had potential to impact recruitment and compliance within their subsequent studies, and helped to select the accelerometer of choice in their future study (Costa et al., 2013). However, at present, no study had qualitatively explored the acceptability and feasibility of a range of measurement tools used to examine the movement behaviours (PA, SB, and sleep) of pre-school children (aged 3-4 years).

5.3 Aims and objectives

The aim of this study was to:

- Determine the feasibility and acceptability of a range of measurement tools used to examine the movement behaviours of pre-school children, from the perspective of parents and carers

The objectives of this study were to:

- Qualitatively assess, through workshops and focus group discussions, parent and carer views on a range of tools (accelerometers, diaries, questionnaires) used to assess movement behaviours of pre-school aged children
- Determine the acceptability and feasibility of tools used to measure the full spectrum of movement behaviours of pre-school aged children

5.4 Methods

5.4.1 Sampling and recruitment

Early years settings (children's centres, nurseries, and schools with pre-school provision) in the North of England were initially contacted using purposive sampling. In England, there is a health divide between the North and South, with the North of England experiencing notoriously worse health outcomes than the South of England (Bambra, Barr and Milne, 2014). A recent (2021) report titled '*The Child of the North Report*' highlighted the stark inequalities that children in the North of England experience (Pickett et al., 2021) emphasising the importance of research in this region. All early years settings (hereon termed '*settings*') contacted were in the highest quartile of index multiple deprivation (IMD), indicating they were in the most deprived areas in the United Kingdom (UK) (Ministry of Housing, Communities and Local Government, 2015). I contacted the settings via telephone, e-mail, and via personal networks. If a setting agreed to participate, an opportunity sampling method was used, whereby posters (appendix 5.1), letters (appendix 5.2), and information sheets (appendix 5.3) were provided to parents and carers of children at the setting. In one instance, I attended a children's centre several times to chat informally with parents and carers about the research and to build rapport with the centre. Participants were eligible to participate if they were a parent or carer of a pre-school child

(aged 3-4 years old). Children were able to be present during the whole session (demonstrations and focus groups) to ensure inclusivity.

5.4.2 Procedure

The study consisted of a demonstration of the different tools to give context and provide a focus to stimulate discussion, directly followed by a focus group discussion. Focus groups were the methodology of choice, as a qualitative methodology should be employed when aiming to understand views and perceptions on a given topic (Silverman, 2017). Focus groups are believed to be an appropriate method of data collection when identifying beliefs, likes, dislikes, opinions and attitudes around a certain product or topic (Fern and Fern, 2001). Group discussions via focus groups are thought to produce insights that would be less accessible without the group interaction, as the focus groups provide an opportunity for thoughts around a topic to be shared and compared (Morgan, 2012). Additionally, a similar previous piece of research that aimed to determine the qualitative feasibility of accelerometers used to assess PA of toddlers (aged 2-3 years) and their parents used focus groups, and identified issues that had potential to impact recruitment and compliance with subsequent studies (Costa et al., 2013).

As outlined by Silverman (2017), there are some key characteristics of a focus group, including:

- Recruitment of small groups of people who share particular characteristics (being a parent or carer of a pre-school child).
- Encourage an informal discussion around a particular topic (the measurement of movement using the outlined tools)

- Schedule of questions around some kind of stimulus material (demonstration of tools)
- Facilitating of discussion by the moderator rather than asking each individual participant all questions (I moderated the discussion between the group)
- A recorded discussion that is transcribed and analysed using typical techniques for qualitative data, usually thematic analysis

All sessions took place between December 2019 and January 2020 in an available room within the participating setting. Settings were located in three deprived areas in England, UK (West Yorkshire, County Durham and Northumberland). The setting managers conversed with parents and carers to arrange an appropriate date and time for the focus group, based on when a group of interested potential participants were available. Participants were fully informed verbally and in written format the purpose and procedures of the research, as well as being provided with a privacy notice (see appendix 5.4). Participants were reminded that participation is completely voluntary and that if they decided to take part they were able to withdraw at any point throughout the focus group and up to 7 days after, by getting in touch and providing their participant number. Participating parents and carers provided written informed consent at the beginning of the group session (see appendix 5.5). Participants then completed a brief demographic information questionnaire which included questions on: age, sex, ethnicity, age of child, education level, employment status, household income and home postcode (see appendix 5.6). Following this, the main session took place. First, I delivered a demonstration and explanation of the measurement tools used to examine movement behaviours of pre-school

children, where parents, carers and their children were able to trial the methods out. This included:

1. Accelerometers (see Figure 5.1 for images and appendix 5.7 for full information) with accompanying instructions (appendix 5.8) and logs (appendix 5.9):
 - a) Actigraph GT3X+ (ActiGraph, Pensacola, USA), a triaxial accelerometer that measures 4.6 cm × 3.3 cm × 1.5 cm and weighs 19g
 - b) activPAL4 micro (PAL Technologies Ltd., Glasgow, UK), an uniaxial accelerometer and inclinometer that measures 2.35cm x 4.3cm x 0.5 cm and weighs 9.5g
 - c) Actical (Philips Respironics Inc., Murrysville PA, USA), an omnidirectional accelerometer that measures 2.8 cm x 2.7cm 27 x 1 cm and weighs 17 g
2. Proxy reported tools including: a sleep diary (appendix 5.10) and a questionnaire (appendix 5.11).



Figure 5.1. Images of accelerometers (in order: Actigraph GT3X+, ActivPAL4 micro, Actical)

The accelerometers were selected based on the first step of this work (*Chapter 3*) determining these to be most valid for this age group. The sleep diary devised for the present study was a modified version of the National Sleep Foundation (NSF) consensus diary (Carney et al., 2012). Sleep diaries were generally shown to be most accurate in the review as part of this work (*Chapter 4*). However, no diaries were identified that examined both sleep duration and quality, so a modification of the NSF consensus diary (developed for adults, not young children) was used to determine the acceptability of such a tool. The questionnaire was a provisional questionnaire that was being developed as part of this PhD research (full development process outlined in *Chapter 6*). Parents, carers and children had the opportunity to trial the measurement tools during the demonstration session.

Following the demonstrations, I conducted the focus group discussions using a semi-structured discussion guide (appendix 5.12). The nature of the questions were based on previous studies that have explored the feasibility and acceptability of measurement tools used to assess physical activity (Costa et al., 2013; Huberty et al., 2015), as well as being based on the aspects that constitute the definition of feasibility (Bowen et al., 2009). The discussions started by asking for initial thoughts and opinions on the tools, which stimulated discussion on the practicality, acceptability, suitability, and any envisaged problems for using the methods to assess the behaviours of their pre-school child. Specific questions were asked regarding whether the tools and accompanying instructions were understandable, comprehensive, and relevant, if discussion had not already covered this. Parents and carers were asked for their opinions on the use of each of the measurement tools for their child aged 3 or 4 years old, for a period of 7 days, as a 24-hour wear protocol if they were to take part in a future research study. The choice of 7 days was based on this being the optimal and usual length of measurement (Bingham et al., 2016b; Cliff, Reilly and

Okely, 2009). The 24 hour wear period was in order to assess the whole 24 hours of the day (Cliff et al., 2017), as well as higher compliance reported for 24 hour wear protocols (Migueles et al., 2017). On completion of the focus group, participants were fully debriefed (in written (see appendix 5.13) and verbal formats), including information on how the responses will be used, alongside being thanked for their time. My contact details were provided in case participants wanted to withdraw their responses, or if they had any questions or concerns regarding the focus group session.

5.4.3 Materials

- Voice recorder
- Information sheets, privacy notice, consent forms, debrief forms
- A list of participants
- Name tags (for pseudonyms)
- Semi-structured focus group topic guide
- Copies of 1) different accelerometers with instructions of use and logs; 2) sleep diary; 3) the provisional movement behaviour measurement tool.

5.4.4 Data management

The focus group discussions were recorded on a voice recorder and transferred onto a password protected research laptop and saved in an encrypted folder. Once the data had been securely saved on the laptop, it was instantly deleted from the voice recorder. The data from the demographic questionnaire was inputted into a file and stored electronically in a password protected file. Participants were assigned unique participant numbers and chose pseudonyms that were present on the consent forms, and from this point anonymity was employed throughout the whole of the study, transcription, and all written outputs of

the work. Participants were able to withdraw their participation/data, without reason, by providing their pseudonym or participant number. The paper copies of the consent forms and completed questionnaires were stored in a locked cabinet in my office. All data will be kept for 10 years, the electronic data will stay stored electronically and manual data will be kept within the university. After 10 years the data will be discarded appropriately (deletion of electronic files or shredding of manual files), unless used for a publishable journal, in which case, the data will be discarded 10 years from publishing.

5.4.5 Data analysis

All focus group discussions were voice recorded and transcribed verbatim. A thematic analysis approach was used to identify important themes and analyse patterns within the data (Braun and Clarke, 2006; Braun and Clarke, 2013; Braun and Clarke, 2014; Clarke and Braun, 2017). I used an inductive approach to the analysis, to ensure that all parent and carer views on the tools could be captured, rather than within prescribed a-priori defined categories used in previous work (Costa et al., 2013). I conducted the thematic analysis manually using a systematic step-by-step process proposed by Braun and Clarke (Braun and Clarke, 2006) (see table 5.1).

Table 5.1: Stages of thematic analysis

Stage of analysis	Analysis conducted
Familiarisation	I collected the data through interactive means through focus group discussions – I came to the analysis with prior knowledge of the data as I was collecting the data, with some initial ideas already present from this. Once the transcription was complete, I actively read through the transcripts multiple times over several months.
Initial coding generation	I conducted line by line coding of each transcript where I wrote notes at the end of each line to state what that line represented. From this, I created a document outlining the codes, notes, and relevant extracts of the data. This was done to show where extracts and notes from the data could be grouped under the coded subheadings.
Identifying key themes	This step included grouping together relevant codes in a meaningful way, to develop the themes of the data. This step was done on paper where I wrote down each of the codes on a separate bit of card, from within and across transcripts, and grouped these together when codes were around a similar topic into ‘themes’. This was done to consider how different codes could be combined and grouped together under overarching themes. This step ended by creating a further document that included a page for each theme heading, with all of the codes written underneath as sub-headings, with all data extracts underneath that had been coded and included in these themes.
Reviewing of the themes	This included going back to the original table of codes that was devised, as well as the overarching themes, to check whether the codes and the data extracts were representative of the overarching theme to ensure that the themes adequately captured the coded data. Following this, I went back to reading the original transcripts to ensure that the developed themes were reflective of the original data, and that each theme contained the relevant and appropriate extracts of data from the transcripts.
Theme definition and labelling	The final stage included defining and labelling the themes by revisiting all phases of analysis and ensuring that each of the key themes captured the essence of the data.

5.4.6 Insurance

The project was covered by Durham University's public liability insurance and professional indemnity insurance.

5.4.7 Ethics

The study was conducted according to the guidelines of the Declaration of Helsinki, and ethical approval was obtained from the Department of Sport and Exercise Sciences Ethics Committee at the University of Durham (*SPORT-2019-07-01T15:01:13-pdhd55* on 19/09/19).

5.4.8 Risk assessment

I completed a risk assessment for this work to determine any hazards and risks that may have arisen for participants and myself as the researcher in the conduct of this research, there were minimal risks and all were identified as 'low risk'. The completed risk assessment can be found in appendix 5.14.

5.5 Results

Three focus groups with a total of 11 parents and carers (including pre-school teachers) of pre-school children (aged 3-4 years) took place. Children were present in two of the three focus groups. Of the nine early years settings contacted, a total of five settings (three nurseries, one children's centre and one pre-school) were willing to participate, with three (one nursery, one children's centre, and one pre-school) remaining in the research due to closures of education settings as a result of Covid-19 pandemic nationwide lockdowns (Mahase, 2020). Table 5.2 provides information on the demographic

characteristics of the participants. A total of 23-25 parents were confirmed to be attending the focus group sessions by the three participating settings prior to the researcher visit (52-56% drop out, no reasons provided).

Table 5.2. Demographic characteristics of participants

Demographic characteristics		n=11
Sex (%)	Female	100
Ethnicity (%)	White British	100
Age of parent/carers (years)	Median	29
	Range	21 to 61
Age of pre-school child participant cared for (years)	Median	3.7
	Range	3.3-4.9
Education level (%)	Masters/PhD or equivalent	0
	Bachelor degree or equivalent	9
	A levels or equivalent	9
	Diploma in higher education/BTEC or equivalent	18
	GCSE's or equivalent	27
	Vocational qualifications (NVQ Level 2)	9
	National nursery examination board	9
	No formal qualifications	9
	Did not specify	9
Employment status (%)	Working full-time	36
	Working part-time	27
	Looking after the home	9
	Not working	27
Household income per year (%)	< £4,999	0
	£5,000 - £9,999	9
	£10,000 -£14,999	9
	£15,000 - £19,999	18
	£20,000 - £24,999	18
	£25,000 - £29,999	18
	£30,000 - £34,999	0
	£35,000 - £39,999	9
	>£40,000 - £44,999	0
	Don't know	18
Index Multiple Deprivation quintile (%) (1 = most deprived, 5 =least deprived)	1	81
	2	9
	3	0
	4	0
	5	9

Results are presented according to the four overarching themes (Figure 5.2). Verbatim quotes from the focus groups are used to support these themes.

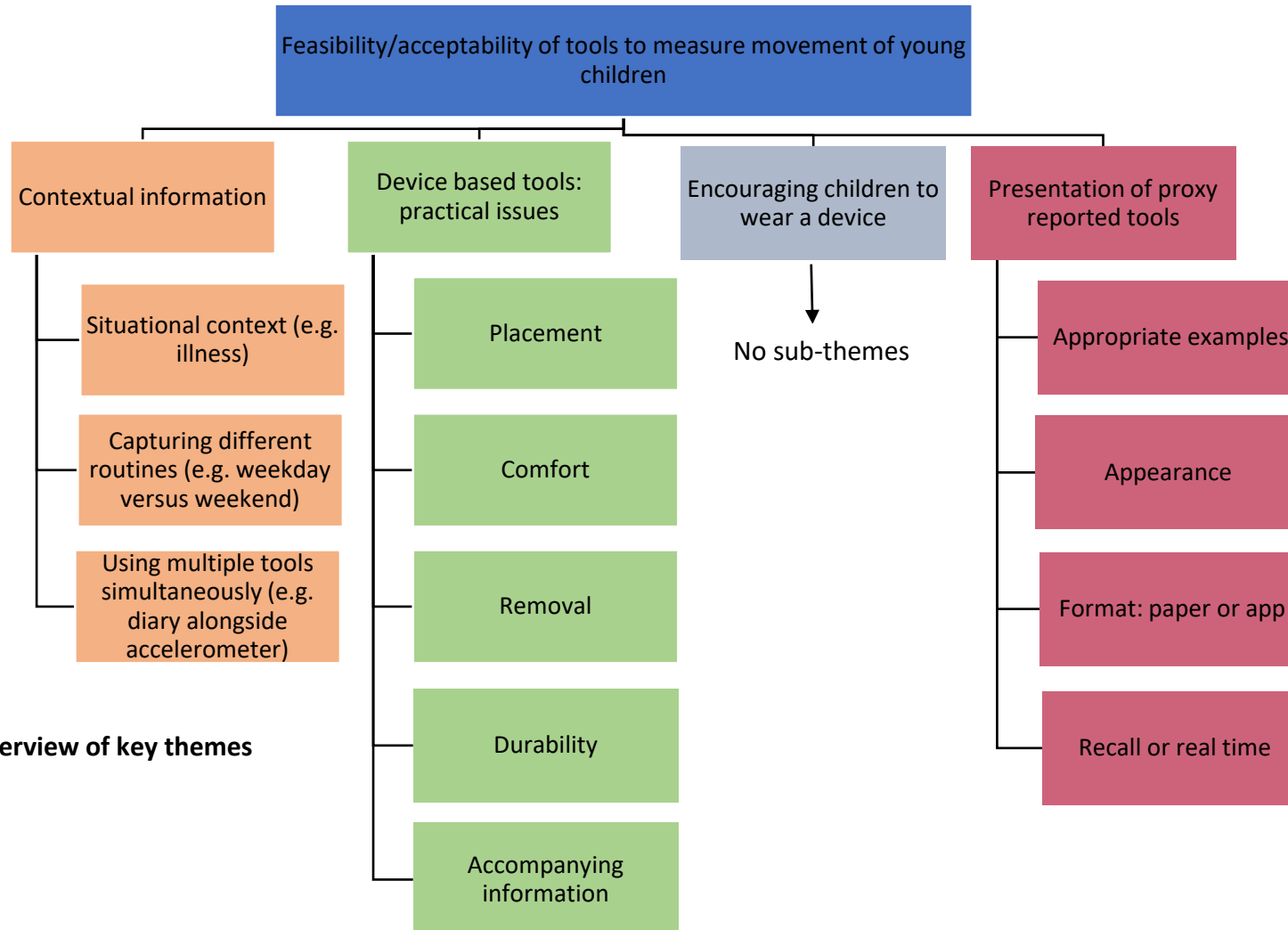


Figure 5.2. Overview of key themes

5.5.1 Theme 1: Reporting contextual information is important to parents and carers

The first theme relates to the importance of being able to report context when using any type of measurement tool, including situational context (e.g. if a child is unwell), the ability for tools to capture different routines, and some considerations when using accelerometers as lone tools for 24 hour movement measurement.

a) Situational context

All groups suggested that reporting key contextual information, regardless of measurement tool, was important for measuring the movement behaviours of their young children. Parents specified that it would be important to report on whether their child has a disability, if they suffer from any sleep disturbances, or whether they were suffering from an illness at the time of measurement. Parents stated their children's behaviours could be starkly different if they are unwell, in that they are more likely to be sedentary, nap during the day, and run around less when they are unwell. Reporting on illness was therefore important to provide an adequate representation of their child's usual 24-hour day.

'I think it all depends on whether they're poorly or not...[name] sleeps like a trooper when she's poorly during the day...But then they have their off-days, so it's like if they're poorly they're going to have a nap.' (P5).

b) Capturing different routines

Parents reported the importance of tools capturing different routines, for example, parents stated that activity would vary depending on if their child was at home or school, at their grandparents or relatives, or if it was a weekday or weekend. Parents were conscious

that they do not see their child's school activity and so they would need to ask the teacher to be able to report this. Parents felt that it was important this information was captured to present an adequate reflection of their child's activity. It was suggested that a format where there was space for both home and school activities would make it slightly easier to obtain information about school activities: *'And I think when it comes to nursery and stuff, if it's a daily thing they don't have to fill it in, you can just say to them, have they played out.'* (P5). This was also discussed by a nursery teacher in the focus group, who suggested that if proxy reported tools were separated by both school and home activities, then nursery could help to complete the day time sections when children are in school: *'...so that time would just be like the time they spent outside, so that would be easier for us, as a nursery, to do just a quick tick box'* (P6). This emphasised the importance of researchers needing to engage with early years settings to have teacher involvement.

Additionally, parents reported that activities and sleeping routines are often different on the weekend than the weekdays where the days are structured around school and work. *'Yeah. That's a weekday, but then during the weekend it's-'* (P4), *'It's a bit different like us.'* (P2) *'we're out or we're going for a walk or we're at home or -the weekend activity and the week, yeah.'* (P4) (Focus group 1 extract).

Parents suggested that they would like to be able to say why routines and behaviours may be different on certain days and that it would be important to explain any anomalies or provide an explanation. Parents said that having the space for *'additional comments'* or to provide more information where you cannot fit it in the generic tool would be helpful to explain anything in more detail. This was thought to be particularly important to explain their child's sleep characteristics, such as why sleep may be different on certain

nights or provide extra information about the context of sleeping arrangements (e.g. if their child moved into their parents bed during the night).

'I definitely think that a box or a question in relation to how often is this usually the case, like being allowed to give a reason as to why on this particular day your child didn't sleep very long or didn't actually go to bed because they might have been ill or it might be a weekend, so they might be allowed to have a later night because it's Friday night.' (P11)

c) Using multiple measurement tools simultaneously

Parents had concerns about the accelerometers being used as a lone tool, particularly for the measurement of sleep, as they thought in these instances it may look as though their child had not slept by examining accelerometer data alone. Parents discussed that their children move around during the night even when they are asleep, either through sleepwalking, having night terrors, or that they move themselves around the bed whilst asleep. Parents proposed that both diaries and accelerometers could be used simultaneously to ensure higher accuracy of their child's behaviour and so information obtained from accelerometers was not misunderstood: *'Then at least you know whether they sleepwalk or they climb mountains or do whatever they do.'* (P5).

5.5.2 Theme 2: Device based tools: practical issues

The second theme concerns practicalities of using devices to measure movement in pre-school age children, including the comfort, placement, removal, and durability of devices. A sub-theme also covers the acceptability of accompanying instructions and logs to report wear time of devices.

a) Placement of devices

When asked for initial thoughts on the various devices (Actigraph GT3X+, Actical and activPAL), participants in all groups showed a clear aversion to devices being attached directly to the skin and most frequently regarded the activPAL as the least accepted tool for these reasons. Participants felt that the device would cause irritation, be painful when removing, were concerned of allergic reactions, described that their child may try pull it off (and subsequently it may be easier to become lost/misplaced), and felt that it would be the most difficult of the devices to re-apply. This was the only device that some of the parents explicitly stated their child would not wear.

'Yeah, they wouldn't want something stuck to their body like that. I just don't think they'd like it at all. Plasters don't stay on very well, let alone...' (P2)

Parents felt that for children with skin conditions such as eczema these types of devices would not be appropriate or feasible: *'my son has eczema so he couldn't have the one that sticks'* (P11). This same concern was raised in relation to children with sensory issues, as exemplified in the extract below.

'I think they're great. The only issue I have with the actual monitors themselves is how children would react if they have sensory issues. So some children with sensory issues are fine with things touching their skin, but something that's actually attached to the skin might be kind of irritating for some children.' (P9)

Parents reported that devices that could be clipped on or worn on top of clothes rather than being directly attached to the skin were preferred; however, there was variation in preferences for device placement. Some parents preferred the hip placement for devices, reporting that devices would be better located out of sight of the children to reduce the

chances that they would play with the device: *'...I think if it's on a belt and you can put like a jumper over the top or something, they'll forget it's there'* (P7). Others stated a preference for the wrist placement as children may think it was a watch, which they reported that children begin to become interested in at this age.

b) Comfort/ease of wearing

There was a lot of emphasis placed on the size of devices, with parents and carers stating that smaller devices would be favoured for this age group. Despite the activPAL receiving negative reactions due to the attachment to the skin, the size of this device was seen favourably. Whereas larger devices may be too big for children of this age, with the Actigraph frequently referred to as *'bulky'* throughout the focus group discussions: *'No, I think it's too bulky for her for it to be on the wrist, she's only dinky, so I think it [Actigraph] would be too big for her.'* (P5). However, it was also suggested that there were positives to having larger devices as it would be easier to identify if it had fallen off: *'...they'll not know it's gone, whereas the ones with the straps you'll be able to see if it's going to fall off, it'll be dangling.'* (P9).

There were some evident issues raised around the applicability of a child wearing the same device for 24 hours per day. This was primarily when the size of devices became problematic, with parents concerned that a larger device may be uncomfortable when their child was sleeping which may result in the child removing and displacing the device, whilst smaller devices attached to the skin may stay on during the night but would be the least practical in terms of applying and for daytime use. The Actical was thought to be more appropriate for 24-hour wear time *'because it's not as bulky'* (P7), *'it's slimline they might not know it's there, they might forget'* (P6) and *'Even when they're asleep it probably won't*

dig in as well.' (P7). This raised the idea that device preference may also be partly dependent on the behaviour and outcome of interest; whereby larger devices not attached to the skin may be preferable for use during the day but may cause more difficulties during sleep. *'But I don't think this would be worn 24 hours. For sleeping, I don't think it'd [Actigraph] be comfortable, if you turn and it's on your hip, I don't think it would be'* (P2). Whilst smaller devices may be better when only measuring sleep: *'Oh, absolutely, because it's like thinner it would be good for bedtimes and they might not feel it as much, but the other one I don't think that would be very good for bed.'* (P7).

c) Removal of devices

Parents spoke about the likelihood that children will try and remove the device regardless of where it is placed, as the devices would initially be a *'novelty'* but then children may get bored of them after a while. Parents spoke of this in relation to watches that their children had previously worn: *'I just, they've worn like little kiddie watches and things before or bracelets, but they always take it, oh, I don't want to wear that anymore, and they'll take it off.'* (P2) *'On/off, on/off.'* (P4) *'Yeah, on and off.'* (P1) (Focus group 1 extract). Parents did state that regardless of where the device was, even if it was attached in the middle of their child's back that they would still find a way to take it off if they wanted: *'Between the shoulder blades... She can't reach that... She'll be like a bear and just scratch it off on the wall!'* (P2). The groups suggested that if the device was out of sight it may help to reduce the chances of children removing it, but it would be preferable to place the device in a location where children could remove it themselves if they were experiencing any discomfort, rather than something that was attached to the skin that could be painful to remove.

'...just for practicality of, just so they're not stressed as well, if they do want to take it off they can quite easily take it off themselves.' (P2).

Although parents thought it was highly likely that the child would want to remove the device at some point, they suggested that the Actigraph and Actical would be easier to reattach as opposed to the activPAL that would require wiping the area and re-applying with new adhesive. This was also a preference for when children are with grandparents or in the care of other adults.

'I think the strap one would probably be the easiest for like grandparents and that, because then they don't have to strip them to stick it back on their thigh' (P10).

d) Durability of devices

Parents and carers talked about the practicalities of the devices in relation to activities that children engage in stating that devices must be '*sand resistant*' and '*water resistant*'. Parents and carers spoke of the need for these considerations when thinking of devices that are wrist worn: '*they're into mud and dirt and soil and digging in the mud...And anything else, slime, if we have slime in the water tank or anything like that*' (P6).

Additionally, parents and carers stated the need for devices to be durable and to be child appropriate '*childproof, like hard, like not feeble, if you know what I mean-so they'll easily break.*' (P7) '*Absolutely- no small parts either we have to say these days.*' (P8) '*Yeah, very good. Childproof!*' (P7) (Focus group 2 extract).

e) Acceptability of accompanying information- instructions and accelerometer logs

Basic instructions usually administered with an accelerometer (see appendix 5.8) were well received across participants, stated to be '*straightforward*' enough to manage the

accelerometer wearing protocol. It was felt the instructions would be understandable even if children were in the care of family members who did not often use technology. *'My grandma's not good with tech and I think she could understand that, so!'* (P11).

Participants suggested that an accompanying visual method of instruction would also be useful, such as a link to a video demonstration on how to put the device on, particularly if devices are being used via remote administration, but also to act as a reminder if the device is removed during the week.

'Ideally probably a demonstration.' (P6) *'Yeah, that'd be good, like a video or something to go on, wouldn't it? That would be handy...Like just say oh if you go onto this YouTube website or whatever, it shows you how to do it, because all right it's reading something, but sometimes I need showing...something to show me what to do. Because then otherwise you might not put it on properly and then it might not get the right reading for what you need.'* (P7) (Extract from focus group 2).

Three types of accelerometer log that ranged in the amount of content captured were presented in the demonstration and focus group (see appendix 5.9). Participants across groups consistently reported that they preferred the detailed logs with space for commenting on whether the accelerometer had been worn, if it had been removed, time and reason for removal, and information on bed and wake up times. Parents reported that this format was *'easier to figure out'* what needed to be reported and how to keep track of when the child was wearing the accelerometer. Simple logs containing less information (e.g. just time monitor was removed, time monitor was put back on, and reason) were seen to be more complex in that parents would have to think more about what they had to write: *'It looks a little bit like...Like where to start, am I putting the right thing in.'* (P4).

Parents stated that plenty of space would be needed on the accompanying accelerometer logs to state when the device had been removed, so that even if children did remove the device frequently, they were able to report this: *'Yeah, I just think I need more room for how many times it would be removed!'* (P5).

5.5.3 Theme 3: Encouraging children to wear a device

The third theme concerns encouraging children to wear the devices, including compliance and ideas on how to make tools more appealing to children of this age.

It was apparent through the focus group discussions that parents and carers thought that compliance with the accelerometer protocol was child dependent. Some suggested that they would not have any real problems with their children wearing any of the devices: *'Yeah, that's the only thing. If I didn't have a choice I wouldn't be too bothered. So I would potentially use any and I think she'd do any...'* (P4). Whilst others felt they would really struggle to get their child to comply with the device wearing. *'I don't think mine would wear them. I don't think. This one I think would be a novelty to start with.'* (P2). Although parents offered some practical solutions to try and increase compliance, it may be that to a certain extent it is individualistic as to whether children would take part, demonstrating the importance of over-sampling in research studies with pre-school aged children: *'...every child's different. Some children might leave it alone, they might adapt to it brilliantly; whereas, other children might be like no they just don't like wearing it.'* (P8)

Participants confidently reported that the way in which devices are framed to children could impact on compliance *'Just because they look cool, they're kind of fun and if it's explained to them in a way they can understand they might think it's really cool'* (P11). An example that arose through all focus groups was framing the device to children as something

that could *'give you superpowers'* or that was a *'magic belt'*. During one focus group, a child showed interest in the devices, so their parent asked them if they would want to wear the device to get superpowers, in which the parent stated: *'[Child] just said he wants superpowers, so he'd have it on all day long.'* (P9). It was suggested that devices could be turned into a game, in particular, this was stated in relation to the Actigraph: *'So, I like the Actigraph one, although it's chunky and everything, I don't think that would be an issue with kids because it stands out for them and you can make it a game.'* (P9).

Similarly, suggestions were made on how to make the devices more appealing for children so that they would be easier to use, including the use of colourful and different patterned belts *'...if they were more appealing, more child friendly. Like colourful bands if you want them to wear it and like you know with like the characters and things.'* (P2). Further suggestions included being able to personalise the devices in some way, such as children designing stickers or pictures that they could put on their device: *'Create your own one. Give the kid some stickers to decorate their own'* (P7). Alongside making the device more appealing to children, these strategies would help practically, to determine who the device belongs to if it falls off whilst at nursery: *'And you would know who's was who, especially around nursery, because then you could just say... this child's has got the flower on, this one's got a car on. So then you'd be like right oh you've lost yours, there you go, rather than going like this is number 574, whose is this!'* (P8).

Parents stated that for children with disabilities, if some of the devices (that did not have a direct skin attachment) were framed appropriately or explained to the child then they could be applicable. With this, it was evident that there needed to be a degree of parental

buy-in to encourage children to wear the devices and to help facilitate compliance from their children.

‘I’ve got a son that’s got these special boots. Like my kids know that he’s got to wear them. So I reckon if I told them they had to wear that they’d just wear it.’ “Yeah, the four-year-old, he’s going through an autism diagnosis at the moment. So this one [Actigraph] we could say was like some kind of magic belt.’ (P2).’

5.5.4 Theme 4: Presentation of diaries and questionnaires (proxy-reported tools)

Theme four is specific to the presentation of diaries and questionnaires, including the format of these types of tools, timing of data collection, and ensuring age specific activities.

a) Examples of movement behaviours

Ideas about the presentation of the proxy-reported tools included that it was favourable if tools were clearly aimed at young children, with one parent reporting that having examples of activities that your child would *‘typically actually do’* made it more appealing, than something that appeared to have unrealistic expectations of young children and the fact that children *‘don’t sit still and play nicely’* (P11). Parents also suggested that as children engage in so many activities, and that these vary per day, it would be useful to have space to report the different activities that their child does, without always having to stay within prescribed activities listed. However, it was recommended that tools should contain some age-specific activities, so that parents can understand the aim of the question but with additional space for reporting further activities.

‘Yeah. So we could just have examples so people don’t think hmm, what have they done?- Yeah, but then have the choice to also write your own if you can do that.’ (P4).

b) Appearance of diaries and questionnaires

There was some contradiction between groups as to whether the proxy reported tools should be made visually appealing or made to look 'boring'. Parents reported that the tool could be quite intimidating and daunting, but if it was colourful or was made to look appealing in some way then it would be more likely that they would want to complete it: *'And maybe colour coordinate it... just something that makes it look a bit like that other thing that you've got... Yeah, just pleasing to the eye and like you, I don't know, it sounds silly, but you want to fill it out, because it's... it looks more pleasing, doesn't it...'* (P4). Others suggested that it would be preferable for the tool to be as dull as possible, as otherwise their children may think that it is for them: *'Exactly, then they'd be like oh I'm just going to colour that in for you mummy, that's all right. It's like, don't do that! But that's what kids do- any bit of paper they find they're like oh I'm just going to draw mummy a nice picture on that. They don't think it's important.'* (P8).

c) Format of tools: paper or app?

Whilst there was little discussion around the mode of administration, in one focus group, it was suggested that a phone application (app) rather than a paper diary may be helpful – particularly for sleep. It was suggested that normally parents will pick up their phone to see what time it is when their child wakes up during the night. *'Yeah, or even like you say yeah an app would be good... An app would be easier, absolutely... you could literally tick a box, bedtime... well, they go up at this time, and then it's so much easier keeping track rather than... having to write this down.'* (P8).

d) *Recall or real time*

Parents reported that questionnaires with previous week recall were confusing and too difficult to complete. Recall was a concern that continued to arise throughout the focus groups, both in terms of how parents felt they would not accurately retrospectively report past week activity, but also their concern that they do not always know if children have woken up during the night so this would not be correctly detected:

'I think it's a bit too hard for me this, because like she does so much each day, like all the different things she does through the day, and then you're having to record them and remember, oh, it's so confusing, I just don't understand it.' (P3)

Rather, parents suggested that having space to report every day would be much more beneficial: *'Most of this it would. I personally would find it easier having it daily instead of how many days this week'* (P5). Parents felt this would be easier as they would not have to think back to the past week and what their child spent the whole of each day doing, were less likely to get children mixed up, but would also be able to capture the intricacies and differences between each of the days. In this regard, the proposed sleep diary was generally well accepted and understandable, with participants suggesting that filling in the parental reported sleep diary for a seven day period was *'Oh, easy'* and would *'be straightforward'*. A diary based format was preferable over a questionnaire; parents stated that a questionnaire may take less time to complete, but they felt that having the space to report in more detail and have a longer tool would be beneficial in terms of the tool meaningfully detecting the activity of their child.

'It might mean it's more paperwork, but it makes it more- easier to look at. So it kind of looks like it's more but it's...' (P4) *'In the long run it'll be easier.'* (P2) *'I think it's going to*

be easier, especially if you've got more kids, because then you're not getting the kids mixed up either, because I know that's what I do. I'm like, [child's] done this today, no, he hasn't, it was [other child], no, it was [other child]-' (P5). (Extract focus group 1).

5.6 Discussion

5.6.1 Principal findings

The results of this study revealed four main themes on the feasibility and acceptability of measurement tools used to assess movement behaviours of pre-school children from the perspective of parents and carers. These include:

- 1) Importance of providing contextual information when using any measurement tool to report on children's movement behaviours (e.g. child illness, capturing different routines such as home and school activity, and potentially using multiple measurement tools to obtain context alongside device based tools)
- 2) Practical issues associated with devices (e.g. placement of devices and aversion to devices being attached directly to the skin of their child; concern of larger devices during sleep time; likelihood of removal of devices (and subsequently needing space to report this on an accompanying log), worry about durability of devices during child play, preference for written and visual instructions for device administration)
- 3) Encouraging children to wear a device (e.g. making devices attractive to children- 'superpowers', colourful belts, and personalised stickers).
- 4) Presentation of diaries and questionnaires (e.g. parents would prefer examples of age appropriate movement activities, much prefer real-time recording than recall).

Overall, the activPAL was the least preferred device and raised the highest amounts of concerns for children in this age group, namely due to this tool being attached to the skin. This finding is in line with previous research reporting that pre-school aged children 'strongly opposed' to having chest worn devices attached to their skin (Shin, 2015). During a structured PA session, children requested to stop wearing the device due to the device feeling uncomfortable and slipping from their skin (Shin, 2015). Devices were said to be more accepted when worn over the clothes, findings also highlighted by Costa and colleagues with younger children (Costa et al., 2013). Similar concerns were raised in other research with the use of the ActivPAL, whereby 38% of parents reported that their child had some skin irritation due to wearing the device, with a small proportion of parents indicating that their child found it unpleasant to wear the device (De Decker et al., 2013). However, further studies have reported that the ActivPAL was generally accepted (Davies et al., 2012; De Craemer et al., 2021) with minimal reports of removal of the device due to the dressing of the device being uncomfortable for the child (Davies et al., 2012). Despite this, in some research studies it may be that for accuracy reasons, devices must be attached to the skin. In these instances, researchers may want to address some of the perceived concerns including a demonstration of the device attachment and removal both to children and the caregivers, and clearly explain what the device is for in an interactive and fun way to children prior to administration. The Actical was the preferred device overall and raised fewest issues for use with young children, mainly due to the smaller size of the device and that attachment was not directly to the skin. In a previous study, researcher reported acceptability of the Actical and ActivPAL suggested that most children had high acceptability to wearing both devices. However, children only wore the devices for one day in this study and authors did note that a proportion of children did not assent to wearing the devices and

some were lost/removed during the day (Van Cauwenberghe et al., 2012). Where possible, devices similar in size to the Actical may be preferred for measuring movement behaviours in this age group.

The Actigraph was generally well accepted by parents and carers, however there were some concerns raised with the size of this device mainly for sleep, in parents stating that the device was 'bulky' and may be uncomfortable if a child rolls on their side whilst asleep. However, it was also suggested that the size of this device made it less likely to become misplaced. In previous research, parents have reported that for a wear time of five days, the majority of children found the Actigraph (GT1M) pleasant or very pleasant to wear, with only few (2%) reporting that their child found the device unpleasant (Cardon and De Bourdeaudhuij, 2007). Similar results were reported using the Actigraph (GT1M) in younger children (aged 2-3 years), whereby no parents reported that their child found the accelerometer unpleasant to wear but three parents reported that the accelerometer sometimes moved out of the correct position (Van Cauwenberghe et al., 2011a). It is important to note that these previous studies have focused on day time activity rather than sleep. Removal of device based measurement tools is a long standing challenge with measurement of PA and related behaviours of young children (Cliff, Reilly and Okely, 2009). Whilst it may be inevitable that devices are removed for at least part of the measurement period, ways to make the tools more attractive and fun may help to increase compliance with the accelerometer wearing protocol. With this, parents suggested that the Actigraph would be the easiest device to turn into a game and if framed appropriately in a fun way and with children having autonomy to make the device their own, then children may enjoy wearing the device. Previous research reported that some children enjoyed wearing the Actigraph accelerometer by thinking of themselves as some of their favourite characters by

wearing the device (Barber et al., 2015). In line with this, parents and carers suggested the use of stickers and colourful belts would make the devices more appealing to children, a strategy that has been suggested in a previous piece of research to try and increase willingness and compliance with pre-schoolers to wear devices (Van Cauwenberghe et al., 2012). This may be a helpful consideration in the design and development of device-based tools and their wear protocol.

Actigraph and Actical accelerometers have most frequently been validated and worn using a hip placement (Dobell et al., 2019; Janssen et al., 2015; Tracy et al., 2021), however, both hip and wrist placements have shown potential to classify intensity of movement in pre-school aged children (Migueles et al., 2017). The findings of the present study showed variation in terms of preference between wrist and hip; with wrist being said to be feasible if the device can be worn like a watch, and the hip placement if the device is under the clothes then children are more likely to forget about it. Additionally, if both tools are framed in an attractive way to children as recommended by parents, placement location may be less relevant or problematic. In light of this, the most accurate placement for the device may be chosen with likely minimal impact on the acceptability of the tools.

In previous PA intervention studies using the Actigraph accelerometer, it was reported that parents sometimes forgot how to apply the device during the seven day protocol (Barber et al., 2015). In the present study, participants reported that written instructions accompanied with video demonstrations would be helpful, this would also help with remote administration and reduce the need for research assistants to visit families during the study. Similarly, parents reported that having detailed accelerometer logs were easier to follow and less complex than simple logs for reporting device removal.

In relation to their own experiences, participants were very aware of the recall bias when responding to questionnaires. Similarly, given their busy schedules with having young children, parents felt that retrospective reporting may be too difficult and confusing. Parents stated that they felt diary based tools would be a much easier and more accurate method of measurement, as these could be completed daily. Where feasible (in terms of resources- researcher time, financial resources), it may be more appropriate to use diaries rather than questionnaires. These findings highlight the importance of involving the target population in the development of new tools to ensure that the tool will be meaningful to complete, that the items on the tool capture the appropriate and relevant behaviours, and that the questions are understandable (Terwee et al., 2018a). Additionally, parents suggested that having space to help put their responses or data from accelerometry into context, as well as being able to explain any anomalies in their child's behaviour was important. Not only would this help with contextualising the information, but may also help parents feel their children's behaviours are more accurately captured. These findings further demonstrate the importance of public involvement and engagement in the design and conduct of research, to ensure that subsequent participation in research is meaningful through appropriately co-designed study materials (Rolfe et al., 2018).

Further to this, it was suggested that the use of electronic formats of data collection, such as phone apps, may provide a welcome alternative to traditional paper formats of questionnaires and diaries (Quante et al., 2015). Such methods are already seen in national data collection schemes such as the Active Lives Survey, where both electronic and paper based formats of the questionnaire are available (Ipsos MORI, 2021); however this does not include proxy report for children as young as 3 and 4 years old. Compliance issues have more frequently been reported for electronic surveys compared with telephone surveys,

with particular high attrition rates for parents with young children (Bennett et al., 2012). As such, it is important that electronic-based methods of data collection for the measurement of movement behaviours of young children are thoroughly evaluated to ensure accuracy and feasibility prior to use (Bennett et al., 2012).

5.6.2 Strengths and limitations, including limitations of the methodology used

To my knowledge, based on the extensive reviewing conducted as part of this PhD (*Chapters 3 and 4*), this is the first study to assess the acceptability and feasibility of a range of measurement tools used to examine the movement behaviours of pre-school aged children, from the perspective of parents and carers. This study used novel methodology for the measurement field that is largely based on quantitative studies only. Through this, the study has highlighted practical issues and helpful insights that researchers may wish to consider when designing studies and selecting materials to use for measuring movement behaviours of pre-school aged children. Although a strength of qualitative research is the rich information and insights obtained from participants, it does mean that the sample sizes are smaller. Although this study included a small sample size, the study offers practical insights and strategies into the design of studies examining movement behaviours of pre-school aged children that may be generalisable through transferability, however; I will leave this at the discretion of the reader as to whether the findings apply to their group of interest (Smith, 2018). The insights obtained by the parents and carers involved in this research could be shared with populations of interest through public involvement and engagement prior to full scale studies. Moreover, a further strength of this study is that the recruitment strategy was used to deliberately draw from groups whose views are often underrepresented in measurement literature (as identified in *Chapter 3 and 4*).

This study also has some limitations that must be recognised. Firstly, I tried to recruit an ethnically diverse population by recruiting two of the five settings in one of the most ethnically diverse cities in the UK (Office for National Statistics, 2011). Despite this, the sample in the present study was not ethnically diverse, as such, generalisability of the findings across different ethnic groups is unknown. It is plausible that I did not achieve an ethnically diverse sample for several reasons, including: 1) The recruitment poster including the logos of my host institution, Durham University, and funders. This may have been disconcerting to potential participants including feeling as though the research was not for them and 2) The positionality of me as a White British researcher relative to the ethnic background of the communities in which I was trying to recruit. Similarly, although an advantage of this research was the recruitment of underrepresented groups, which is a key focus of this PhD, it would be useful to test these tools using similar methodology across a range of SES groups. Although participants had the opportunity to trial the tools during the demonstration conducted as the first part of the discussion, they did not use the measurement tools for extended periods of time prior to the focus groups. It is plausible that in practice more feasibility concerns could be raised than those discussed in this research. The aim of this research was to explore the acceptability and feasibility of the tools prior to use in future research projects. Therefore, although this study provides important insights into the feasibility, acceptability, and practicality of a range of measurement tools that may otherwise not have been considered, we cannot determine from this study whether or how this would impact compliance in research studies. Finally, this research was conducted prior to the coronavirus pandemic, it is therefore plausible that some of the issues raised may be exacerbated as a result of this. For example, participants may be more vigilant to the cleanliness of device-based tools.

In this study we presented certain models of devices, but due to rapid and continuous advances in accelerometry technology (as discussed in section 2.11.2, p.126 onwards), it is likely that newer versions of the devices will be available, now and in the future. The findings of this research are not intended to be a criticism of any particular devices, but rather to present practicalities of using the devices from the perspective of parents of young children, to help with the development and implementation of research studies with this age group. We hope that the findings of this study will also be useful to manufacturers in future development of revised versions of the tools.

Focus groups were seen to be the most appropriate method of choice for the work conducted in this chapter due to the group interactions that would enable insights that may not be accessed if examining the measurement tools using other methods such as interviews (Morgan, 2012). However, some limitations of this method must be acknowledged. In the present study it was a challenge to get a substantial group of parents together in one place at the same time which limited the size of the focus groups that took place. In addition, a well-known limitation of focus groups, is that some participants may restrict their opinions to support the dominant opinions or views seen as more socially acceptable (Smithson, 2000). A further limitation is that naturally in group interactions there are often more dominant voices that can overpower the group during focus group discussions (Smithson, 2000). In the present study, I tried to overcome this by involving quieter individuals through eye-contact and asking them specifically for their views on the different tools.

5.6.3 Implications of findings and recommendations

An implication of this work is that parents and carers raised multiple practical issues, alongside some useful strategies, for the use of a variety of measurement tools that can be used to assess movement behaviour of pre-schoolers. These findings have implications for the conduct of future research with pre-school children, as they provide an understanding on the feasibility of different tools for different purposes. For example, parents demonstrated preferences for the size of different devices depending on what movement behaviour is being monitored. Additionally, awareness of these issues and the strategies to overcome some of the potential concerns may help with recruitment and retention of participants to movement behaviour research (e.g. providing demonstrations of use of different tools may help to overcome some perceived barriers of the tools).

The findings of the research can also be used by practitioners and policy makers when considering ways of assessing movement behaviours in practice. This may be in the form of knowing important questions to ask to make the tools meaningful if using proxy reported tools, or when determining whether tools may be appropriate for large-scale surveillance based on perceived feasibility. Alongside this, these findings can be used by manufacturers and tool developers for future development of both proxy and device-based measurement tools. Manufacturers of device-based tools can use the findings of this research to ensure that future tools can be designed to be applicable for pre-school children. For example, ensuring that tools are robust enough to counteract the concern of durability of devices for when children are playing in sand/mud.

This research has enabled the development of a series of recommendations for the measurement of movement behaviours of pre-school aged children that may be used in the

design and development of research projects. These recommendations can be found in the Infographic displayed as figure 5.3.

10 recommendations for measuring movement behaviours of pre-school aged children



1

Context is important to parents and carers– having space to explain their child’s health status (e.g. illness) helps make measurement meaningful.

2

Ensure that measurement captures different routines to be reflective of children’s movement e.g. home vs school, weekday vs weekend.

3

Devices worn as watches or placed ‘out of sight’ preferred– devices stuck to the skin were less favourable.

4

Smaller devices preferred for 24-hour movement measurement –but larger devices may be preferred for day time measurement only and smaller devices for night time.

5

Removal of devices may be inevitable- include detailed device wear time logs with plenty of space to report when the device is worn/removed.

6

‘Child-friendly’ devices – suitable for playing in mud, slime, sand and water- and no small parts! Demonstrate safety of devices prior to studies.

7

Written and visual (video demonstrations) study instructions most helpful.

8

Frame research to young children so that taking part is ‘cool’ – devices can give ‘superpowers’ or ‘magic’

9

Modify tools to make them ‘childlike’ –stickers so children can decorate/personalise their device, colourful belts or ones with children’s favourite characters on.

10

Daily reporting easier for proxy reported tools – recall can be particularly challenging with young children. Ensure that age appropriate activities are included on the tool.



Source:

NIHR | School for Public Health Research

Figure 5.3. Infographic of recommendations for measuring movement behaviours of pre-school aged children

5.7 Conclusion

This novel study highlights the importance of being able to provide contextual information when reporting movement behaviours of young children (e.g. child illness, capturing different routines); some practical issues associated with devices (e.g. aversion to devices being attached directly to the skin of their child; concern of larger devices during sleep time); encouraging children to wear a device (e.g. making devices attractive to children- ‘superpowers’, colourful bands and stickers); and preferences on presentation of diaries and questionnaires (e.g. age appropriate activities, preference for real-time recording over recall). The findings of this research provide valuable insights from parents and carers on the use of various measurement tools with young children that can be used to help with the design, development and implementation of studies measuring movement behaviours of pre-school aged children.

5.8 Summary and link to next chapter

This chapter highlighted some important findings around preferences and views of parents and carers on various measurement tools used to assess movement behaviours of pre-school children, including specifically some insights on views on the presentation of proxy reported tools. As highlighted in the reviews presented in this thesis (*Chapters 3 and 4*), there is a need for more in-depth development and evaluation work of proxy reported tools, with the target population. Additionally, the reviews both demonstrated that there is insufficient evidence to conclude which tools are most suitable for use with lower SES communities or different ethnic groups. On this basis, the purpose of the next chapter was to bring findings from *Chapters 3, 4 and 5* together by co-developing a measurement tool that could be used to examine the movement behaviours of pre-school children, together

with parents and carers, followed by assessing the content validity of the tool. The work presented in this chapter (*Chapter 5*) was conducted in parallel to the work that will be discussed in the next chapter (*Chapter 6*), with part of the focus group sessions including development of the new tool.

- 6 [Chapter Six: Movement Measurement in the Early Years](#)
(MoveMEY): Development and content validity of a movement
behaviour measurement tool for pre-school aged children.

6.1 Prelude

In this chapter, I draw on the findings from the previous chapters that identified that there are currently no suitable measurement tools for determining the movement behaviour guidelines of pre-school children at scale (*Chapters 2, 3 and 4*), as well as drawing on parent and carer views on the acceptability and feasibility of tools (*Chapter 5*). As highlighted in *Chapters 3 and 4* of this thesis, there is a need for more in-depth development and evaluation work of proxy reported tools, with the target population. Additionally, the reviews presented in *Chapters 3 and 4* both demonstrated that there is insufficient evidence to conclude which tools are most suitable for use with lower SES communities. On this basis, the study outlined in this chapter was conducted to fill the gap, by co-developing (with parents, carers, and topic expert researchers) a measurement tool that could be used to examine the movement behaviours of pre-school children, followed by assessing the content validity of the tool. In this empirical work, I aimed to deliberately recruit families from deprived areas in the United Kingdom (UK). The work presented in this chapter was conducted in parallel to the work discussed in the previous chapter of this thesis (*Chapter 5*), with part of the focus group sessions including development of the new tool, which is presented in this chapter, *Chapter 6*.

6.2 Introduction

This chapter outlines the development and content validity assessment of a tool used to assess movement behaviours of pre-school aged children at scale. As identified through *Chapter 3* and *Chapter 4* of this thesis, there are no proxy reported tools available that have been shown to be suitable, including valid, reliable, and feasible, for examining the PA, SB (*Chapter 3*) and sleep (*Chapter 4*) of pre-school aged children. Further, no proxy

report tools encompassing all movement behaviours for this age group exist with acceptable validity and reliability (as described in *Chapter 3* and *Chapter 4*, see also Arts et al., 2022).

The lack of valid, reliable, and feasible proxy reported tools may be in part due to the limitations outlined in *Chapter 2* of this thesis (see section 2.7.10 (p. 92), 2.8.8 (p.112) and 2.10 (p.122)), such as recall bias or difficulty examining the sporadic and intermittent nature of young children's movement. An alternative explanation is the lack of involvement of the target population in the development of these tools, and the lack of focus on content validity assessment. Information on the development and content validity of tools is often not described or examined (Hidding et al., 2017; Hidding et al., 2018; Kelly, Fitzsimons and Baker, 2016), or is very minimally described (Bacardi-Gascón et al., 2012; Dwyer et al., 2011). However, it has been recommended that the first (and most important) step when designing a new measurement tool is comprehensive development (including the target population and topic experts) and subsequent pilot of the tool with the target population of interest (De Vet et al., 2011; Terwee et al., 2018b). This helps to ensure that the tool works as intended for measuring the construct it aims to measure, within the target population, and for the context of use; to improve the quality of tools prior to further evaluation of validity, reliability, and responsiveness (De Vet et al., 2011; Terwee et al., 2018b). Otherwise, for example, there may be a discrepancy between behaviours that pre-school aged children engage in and what the measurement tool assesses.

Content validity is defined as: '*the degree to which the content of an instrument is an adequate reflection of the construct to be measured*' (Mokkink et al., 2010b). Content validity relates to the relevance, comprehensiveness and comprehensibility of a measurement tool for the construct, target population and context of use. As content validity is difficult to measure, it is suggested that a well-designed development of tool

study should be conducted to ensure content validity (Olshansky et al., 2012; Terwee et al., 2018b).

At present, no study has comprehensively developed and assessed the content validity of a measurement tool used to examine the movement behaviours of PA, SB, and sleep of pre-school children (aged 3-4 years) for assessing movement behaviour guidelines at scale.

6.3 Aims and objectives

The aim of this study was to:

- Co-develop and assess the content validity (including comprehensibility, comprehensiveness and relevance) of a measurement tool for assessing the movement behaviours of pre-school aged children at scale, together with parents and carers of pre-school aged children and topic experts.

The objectives of this study were:

- To develop a measurement tool used to assess the movement behaviours of pre-school aged children, together with parents and carers of pre-school aged children, and topic experts.
- To pilot the devised measurement tool, and conduct interviews with parents and carers of pre-school aged children following the pilot, to determine the content validity of the devised measurement tool for assessing the movement behaviours of pre-school aged children.

6.4 Methods

6.4.1 Conceptual Framework

As this chapter of the thesis involved the development of a new tool, I examined measurement theory to ensure that the tool development had solid underpinnings and could be appropriately guided by a theoretical or conceptual framework where relevant. Measurement theory relates to how the output (e.g. scores generated on a questionnaire) from the items of a tool represents the construct intended to be measured (Edwards and Bagozzi, 2000). Theoretical frameworks regarding measurement are not required when observable characteristics are being measured, such as PA (De Vet et al., 2011). The reason for this is that with observable constructs, it is usually apparent how the items contribute to the construct being measured. Measurement theories predominantly focus on unobservable constructs (De Vet et al., 2011; Petrillo et al., 2015). An example of such a construct includes anxiety; although *symptoms* of anxiety may be observable, we cannot directly observe anxiety itself. This is demonstrated through the use of methods such as the ‘*Revised Child Anxiety and Depression Scale*’ to determine anxiety status, which uses a range of symptoms to determine whether a child is classified as having anxiety (Mathyssek et al., 2013). Therefore, the current project does not propose a theoretical framework, as the measure being developed and evaluated is to survey movement behaviours.

Despite this, a conceptual framework, defined as “*a model representing the relationships between the items and the construct to be measured*” (De Vet et al., 2011, p.8) is useful for the measurement of observable constructs. The conceptual framework can determine the underlying relationships between the items that make up the tool, and the construct that these items aim to measure. Although measurement tools developed and

evaluated to assess PA often exclude the use of a conceptual framework (Gimento-Santos et al., 2011), research has suggested that a lack of conceptual frameworks used to devise measurement instruments leads to inadequate development and validation of measures, as well as possible ambiguity with what the tool is measuring (Rothman et al., 2007).

Therefore, the research presented in this chapter used the World Health Organization movement behaviour guidelines for pre-school children as a conceptual framework (WHO, 2019), to ensure that the constructs being investigated (PA, SB, and sleep) and the relationships between these and the developed measurement tool are clearly distinguishable. The relationships between the constructs being measured and tool itself were further enhanced by the target population (parents and carers of pre-school children), informing the development of the measurement tool, using qualitative methods, to ensure that the tool was appropriate for measuring the behaviours (Dobbels et al., 2014).

6.4.2 General procedures

COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) is an initiative comprising of an international team of researchers aiming to improve measurement in research and practice. COSMIN provide guidelines for assessing the quality of development of tools and content validity studies, mainly for the assessment of quality of studies when conducting systematic reviews. However, these standards have an overall aim to improve health outcome measurement tools. As these set of standards demonstrate important design requirements for content validity studies, they provide a strong basis for a development of tool and content validity research study (Terwee et al., 2018b). Researchers have previously recommended following the COSMIN framework and using the checklist to report on the development and content validity of measurement tools

used to examine PA and SB (Degroote et al., 2020; Terwee et al., 2018a). As such, the development and content validity assessment of the measurement tool described in this work was conducted using the COSMIN guidelines as a framework (Terwee et al., 2018a; Terwee et al., 2018b) (see appendix 6.1). This research is also reported in line with the COSMIN general recommendations for the design of any study on measurement properties of a tool (Mokkink et al., 2019).

The first step when developing a new measurement tool is to conduct thorough searches of the measurement properties of existing measurement tools. This searching is considered to be important for several reasons:

1. To prevent the development of new tools that are a replication of pre-existing tools.
2. To develop ideas on what the tool should/should not look like. Pre-existing tools that are not sufficient for the purpose or have poor quality can still provide important information.
3. To potentially identify tools that can be translated/adapted for specific needs, including providing a set of potentially relevant questions.

Tool development should also involve experts who have engaged with the target population on a particular topic (e.g. researchers of PA in young children) and involvement of the target population themselves, as the key experts. In particular, through the use of qualitative methodologies. The current work included development of the measurement tool with the target population (parents and carers of pre-school aged children) through focus groups, and topic expert researchers through qualitative surveys. Both parents/carers and researchers were asked about the relevance of items, comprehensiveness, and

comprehensibility for the content validity assessment. This included interviews with parents following completion of Movement Measurement in the Early Years (MoveMEY). This study included four distinct steps (Figure 6.1), and four different methods were applied (Figure 6.2).

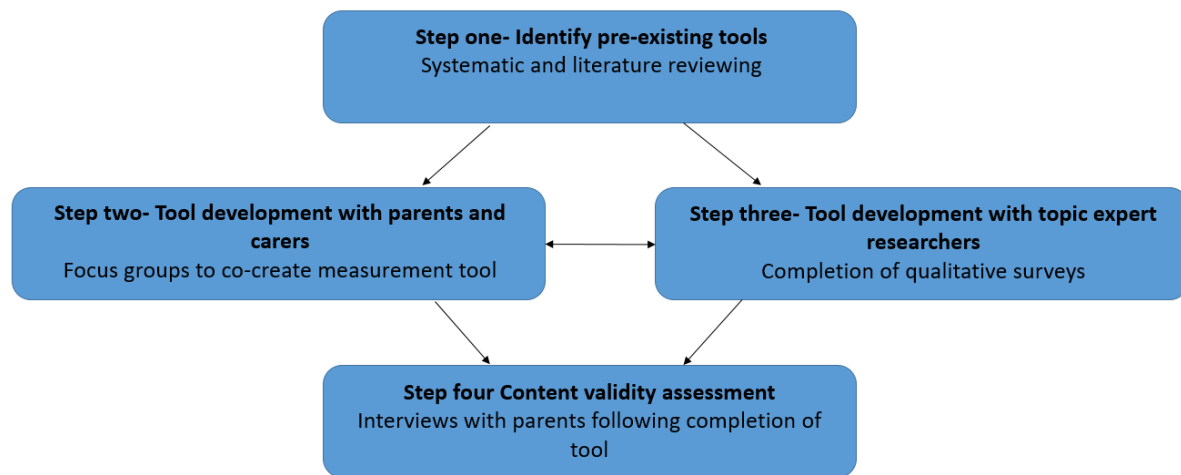


Figure 6.1. Development and content validity assessment

In summary: 1) Initial development of a tool based on extensive searching of the literature and mapping questions to global guidelines for movement behaviours in pre-school children (WHO, 2019), 2) Development of MoveMEY through focus groups with parents and carers of pre-school children using the provisional tool as a stimulus for discussion, 3) Qualitative survey consultation with topic relevant researchers, to aid in appropriate development of MoveMEY, 4) Assessment of the content validity (relevance of items, comprehensiveness, and comprehensibility of MoveMEY) through interviews with carers of pre-school children following completion of MoveMEY . The detailed procedures of each step are outlined in section 6.4.5 of this thesis and an overview can be found in Figure 6.2

Step 1	Step 2	Step 3	Step 4
Aim: Examine the pre-existing literature on measurement tools for assessing the movement behaviours of pre-school children	Aim: Developing the Movement Measurement in Early Years (MoveMEY) tool	Aim: Developing the MoveMEY tool	Aim: Examine the content validity of MoveMEY
Method: Literature reviewing and guideline mapping	Method: Focus groups with parents and carers of pre-school children	Method: Qualitative surveys completed by topic expert researchers	Method: Interviews with parents following completion of MoveMEY
Procedure: Provisional tool created based on scoping of pre-existing literature and mapping the questions to the World Health Organization movement guidelines for pre-school children	Procedure: Co-created MoveMEY through focus group discussions with parents and carers of pre-school aged children	Procedure: Feedback through qualitative surveys from topic expert researchers to aid development of the tool	Procedure: Discussions with parents of pre-schoolers following a week of completing MoveMEY, regarding comprehensiveness, understandability and relevance of items.

Figure 6.2. Overview of steps taken in developing and determining content validity of MoveMEY.

6.4.3 Sampling and recruitment

The focus for the measurement tool devised as part of this work was on pre-school children aged 3-4 years (for reasons outlined in section 2.2.2 p.32-33) of this thesis). For the focus group discussions (**Step 2**) and interviews (**Step 4**), participants were parents and carers of pre-school children (aged 3-4 years old) recruited through early years settings (children's centres, nurseries, and schools with pre-school provision) in the North of England, UK. Pre-school children themselves were not recruited due to evidence suggesting that children aged <9 years old do not yet have the cognitive capabilities to be able to fill out self-report methods or be able to participate in participatory research methods (Atkin et al., 2012; Hidding et al., 2020).

6.4.3.1 Step 1: Provisional development of MoveMEY

The methodologies of **Step 1** including systematic and literature reviewing meant that no participants were required for this step.

6.4.3.2 Step 2: Focus groups with parents and carers of pre-school children

The procedures for recruitment for this step are outlined in *Chapter 5* section 5.4.1 (p.387) of this thesis, but I will provide a summary here for ease of the reader. Parents and carers were recruited through early years settings using purposive sampling, to select schools in the highest quartile of most deprived areas in the United Kingdom (UK), obtained by the index multiple deprivation (IMD) score, indicating they were in the most deprived areas in the UK (Ministry of Housing, Communities and Local Government, 2015). Children were able to be present during the whole session (demonstrations and focus groups) to ensure inclusivity.

6.4.3.3 Step 3: Topic relevant researchers

Topic relevant researchers were recruited through my personal network (and personal networks of my supervisory team). Researchers were approached via email to identify if they would wish to comment on the measurement tool. On receipt of a confirmation email, the provisional MoveMEY measurement tool (see appendix 5.11) and scoring sheet (see appendix 6.4) and 2) a qualitative survey outlining face and content validity questions (see appendix 6.5) was sent and all contact proceeded via email.

6.4.3.4 Step 4: Interviews with parents of pre-school aged children

For the content validity study, all research took place remotely due to the ongoing COVID-19 pandemic. Recruitment was limited to English speaking parents, due to the study materials and the diary only being available in English at this time. A different group of parents and carers to those involved in the development of MoveMEY (**Step 2**) were recruited through early years settings using an opportunity sampling method, whereby information about the study was provided to parents and carers of pre-school children through several methods including: posters in the early years settings (see appendix 6.6), or their online platforms (e.g. social media), and through word of mouth. Parents who wanted to take part were asked to contact me using the contact details provided, at this point I provided a written overview of the work via email, including the information sheet (see appendix 6.7), privacy notice (see appendix 6.8), and a consent form (see appendix 6.9).

This involved several recruitment strategies:

1. I used purposive sampling to contact settings in Yorkshire and County Durham who were in the highest quartile of most deprived areas in the UK, as assessed by

the index multiple deprivation (IMD) score (Ministry of Housing, Communities and Local Government, 2019), through telephone and e-mail (some settings I had existing links with).

2. A local public health practitioner who focuses on early years distributed the study poster to nurseries and schools in the area (County Durham).
3. Snowballing recruitment whereby participants provided the study poster to other family and friend members to take part in the research.

6.4.4 Remuneration for participation

For **Step 4** only, participants received £20 in Love2Shop vouchers upon completion of the study. Children were also sent some 'MoveMEY' stickers in the study pack containing in the MoveMEY tool. The compensation for time structure was outlined via email at the beginning of the study and was also provided on the written documentation (poster and information sheet).

6.4.5 Procedure

6.4.5.1 Step 1: Provisional development of MoveMEY

The provisional version of the MoveMEY measurement tool was a paper-based questionnaire. The development process included:

1. Conducting a systematic review of the measurement properties (validity, reliability, feasibility) of measurement tools used to examine PA and SB of pre-school aged children (*Chapter 3*).
2. A rapid review of the measurement properties (validity, reliability, feasibility) of measurement tools used to examine sleep of pre-school aged children (*Chapter 4*).

3. Literature screening of published journal articles whereby the movement behaviours of pre-school aged children had been examined at scale (>100 participants), to identify specific questions or items used to measure all or part of the guidelines (see appendix 6.2). Although the aim of this research is to develop a tool that could be used to determine compliance with the guidelines in pre-school children at scale (using a definition of >1000 children), when scoping the literature, it was apparent that a very small proportion of studies had examined the three movement behaviours collectively in pre-school aged children (n=1). Therefore, it was decided that it would be more pragmatic to extend the searches to studies that included more than 100 participants in their final sample.
4. Screening of current worldwide activity surveillance systems (see appendix 6.3).
5. Mapping of questions to current movement behaviours guidelines for pre-school aged children (WHO, 2019).

The reviews were used to gain insight into useful aspects of existing measures used to examine PA, SB, and sleep of pre-school aged children. The literature screening of published journal articles and current surveillance systems were examined to identify what methods and questions were currently being used in large scale studies and at population level, including what aspects of the guidelines they targeted and how questions were framed. Following this, the current (and recent) 24 hour movement behaviour guidelines published by the WHO (WHO, 2019), and the Chief Medical Officers' UK PA guidelines (Chief Medical Officers, 2019) were examined. A tool was developed based on these concepts,

ensuring that guidelines were captured, and that applicable formats and questions from the extensive searching were included.

6.4.5.2 Step 2: Focus groups with parents and carers of pre-school aged children

The procedures for the focus group discussions are outlined in detail in *Chapter 5*, section 5.4.2 (p. 380-390). In brief, the study consisted of a demonstration of the different tools to give context and provide a focus to stimulate discussion, directly followed by a focus group discussion with parents and carers. In addition to the procedure outlined previously, for the section of the focus group focusing on the development of the tool (provisional tool shown to act as a stimulus and to help facilitate discussion, which can be found in appendix 5.11), participants were informed that the purpose was to develop the tool with them, rather than for them. This included discussions with parents and carers to identify what kind of tool would be best and what the tool should look like, as well as discussions around the types, frequency, and duration of activities that their pre-school aged child engages in to ensure that the tool would appropriately capture the behaviours of pre-school aged children. It was made clear that the measurement tool could be completely re-designed, re-formatted, could include different items, and that items could be added or removed. Explicit questions were also asked on the relevance, comprehensiveness, and understandability of the tool, and how it could be improved. All sessions took place between December 2019 and January 2020 in an available room within the participating setting. Settings were located in three deprived areas in England, UK (West Yorkshire, County Durham and Northumberland).

6.4.5.3 Step 3: Qualitative surveys with topic expert researchers

Qualitative surveys were completed by topic expert researchers using the same provisional tool (see appendix 5.11) presented to parents and carers in the focus groups. Once agreement to comment on MoveMEY was obtained, topic experts were sent a copy of the provisional MoveMEY questionnaire (appendix 5.11), the scoring sheet (appendix 6.4), along with a qualitative survey (appendix 6.5), which included items on the relevance and comprehensiveness of the tool, and their perceived ability of the tool to capture the movement behaviour guidelines. There was also space for any additional comments, feedback or suggestions in relation to the tool. The questions outlined on the qualitative survey were devised using a range of references and resources (Chief Medical Officers, 2019; Kelly, Fitzsimons and Baker, 2016; Mokkink et al., 2010b; Mokkink et al., 2018b; Terwee et al., 2010; WHO, 2019). Surveys were returned via email upon completion between December 2019 and January 2020.

6.4.5.4 Step 4: Interviews with parents of pre-school aged children

Interviews with parents of pre-school aged children were conducted following completion of MoveMEY to assess the content validity (comprehensiveness, understandability and relevance of the tool). Interviews are the most widely used qualitative method of data collection, which involve a conversation between the researcher and participant, to understand participants' experiences on a particular topic (Willig, 2013). As this piece of research was to determine parents' personal experiences of completing the diary, a one-to-one interview was deemed the most appropriate methodology. This is in line with interviews being proposed as an appropriate method when determining the content validity of a tool (Terwee et al., 2018a). The interviews followed a semi-structured guide

(see appendix 6.10), based on the quality criteria for good content validity outlined by COSMIN (Terwee et al., 2018a; Terwee et al., 2018b), to ensure that the basic questions needed to determine the content validity of the tool were asked, whilst allowing free flowing conversation in the direction guided by parents. The interviews included questions on the parents' opinions and experiences of completing the measurement tool, including questions on the relevance of specific items, the comprehensiveness and comprehensibility of the tool, instructions, response options and recall period, as well as asking if any behaviour is accounted for twice, or if anything was missing.

Participants were provided with all study information via email at the first point of contact, including: information sheets (appendix 6.7), privacy notices (appendix 6.8) and consent forms (appendix 6.9). Following informed consent (by way of a signed and digitally returned consent form) participants were asked to fill in a demographic questionnaire which included questions on: age, sex, ethnicity and religion of child, relationship to the child (e.g. mother, father), age, sex, ethnicity, religion, education level, employment status of parent, household income and home postcode (see appendix 6.11). After completion of the demographic information questionnaire, a 'research study pack' was posted to the participating family (using the address the parent provided); this included the MoveMEY measurement tool developed during **Step 2** and **Step 3** (see appendix 6.12), which parents were asked to complete for the period of 7 days (see appendix 6.13 for study pack letter).

Following the week of completing the MoveMEY measurement tool, I conducted telephone interviews with the participants. Participants were given the option of phone or video call for the interview. At the beginning of the interviews, I briefly restated the study aims, asked participants if they had any questions, and if they were still happy for the

discussion to be voice recorded. All interviews were voice recorded. On completion of the interview, I sent participants a debrief form (appendix 6.14) via the email address provided, restated for participants to get in touch if they had any questions, and posted the vouchers using the address provided. Participants returned their completed MoveMEY diary back to me via post (see Figure 6.3 for image of full study procedure).

Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
•Information provided to participant about the study. •Consent if participant wishes to take part.	•Demographic information completed and sent to researcher via email •Study pack including the MoveMEY diary posted to the participant.	•7 days completing the MoveMEY measurement tool	•Remote interview via video or phone call (at a time convenient to the participant).	•Debrief •Diary returned •Vouchers sent to participant

Figure 6.3. Procedure for content validity study

6.4.6 Materials

6.4.6.1 Step 1- Provisional development of MoveMEY

This step included thorough systematic and literature reviewing (see *Chapter 3*, *Chapter 4*, appendix 6.2 and appendix 6.3). Alongside this, the provisional tool was developed by mapping questions to the WHO movement behaviour recommendations for pre-school aged children (WHO, 2019).

6.4.6.2 Step 2- Focus groups with parents and carers of pre-school aged children

All materials used for **step 2** can be found in section 5.4.3 (p.3928) of this thesis (*Chapter 5*). The provisional tool developed that was used as a stimulus in the focus group discussions can be found in appendix 5.11.

6.4.6.3 *Step 3- Topic relevant researchers*

Provisional tool developed (appendix 5.11) and provisional scoring sheet (appendix 6.4). Qualitative surveys asking questions on face and content validity of the provisional tool (see appendix 6.5).

6.4.6.4 *Step 4- Interviews with parents of pre-school aged children*

‘Research study pack’ – including:

- 1) Study information
 - 2) Pens
 - 3) One copy of the MoveMEY diary
 - 4) MoveMEY Stickers
 - 5) Pre-paid envelope to return the diary
- Information sheet, privacy notice, consent form, debrief form
 - Demographic data questionnaire
 - Vouchers (as compensation for time)
 - Research laptop
 - Voice recorder
 - Interview guide

6.4.7 *Data analysis*

A thematic analysis approach was used to code and summarise data across all qualitative data collected (**Step 2, Step 3, and Step 4**). For **Step 2** (focus groups) and **Step 4** (interviews) all discussions were voice recorded, transcribed verbatim, and analysed using a thematic analysis approach (Braun and Clarke, 2006; Braun and Clarke, 2013). Thematic

analysis is a qualitative research analysis method used to identify and analyse patterns of meaning within a set of data, which provides the researcher with the ability to identify themes that are important in relation to the phenomena being explored (Braun and Clarke, 2006). Thematic analysis is a widely recognised approach for analysing qualitative data during tool development and when determining the content validity of measurement tools (Terwee et al., 2018b). In **Step 2**, thematic analysis was used to code the data, summarise the opinions and ideas of the parents and carers, and to subsequently create the first official version of the tool. In **Step 3**, responses from topic experts were coded and themes were compared with the responses from parents and carers to adapt the tool. In **Step 4**, a similar approach was used whereby the data was coded and summarised to make any necessary adaptations to the tool. Additionally, following the pilot of the tool, diaries were scored to ensure the scoring sheet was appropriate, and the diaries were screened to determine any additional concerns such as if any questions were consistently unanswered that had not been raised through the interviews. The themes for each section were summarised under the following categories: 1) Format of MoveMEY, 2) Content of MoveMEY, and 3) Scoring of MoveMEY (for topic expert researchers and following the pilot). However, any other relevant data that did not fit under these categories was still included as additional themes.

6.4.8 Data management

For this chapter, data management procedures are in place for **Step 2**, **Step 3** and **Step 4**. The data management for all data collected as part of **Step 2** is outlined in *Chapter 5*, section 5.4.4 (p.392) of this thesis. **For Step 3**, the survey responses from topic expert researchers were stored on a Durham University password protected research laptop in password protected files. Approval was obtained from all participating researchers as to

whether they would like to be acknowledged and named or kept anonymous. The data management procedure for the data collection in **Step 4** was as follows:

All participants were given a unique participant number, which was present on the consent forms and from this point was used throughout the whole research process on all study documentation (demographic questionnaire, MoveMEY diary and transcripts). The name of the participant was never used in order to keep the participants data completely anonymous. All interviews took place via phone (although the option for Zoom was also available) and were recorded on a voice recorder, and subsequently transcribed verbatim. The data (demographic questionnaire and interviews- raw data and transcripts) were saved on a Durham University password protected research laptop in password protected files. Once the data was securely saved on the laptop, it was deleted from the voice recorder. The returned MoveMEY diaries were stored in a locked cabinet; the data from the diaries was scanned into a password protected file on the Durham University password protected research laptop and at this point the original paper files were shredded. All data will be stored electronically for 10 years. After 10 years the data will be discarded appropriately (deletion of electronic files), unless used for a publishable journal, in which case, the data will be discarded 10 years from publishing.

6.4.9 Insurance

The research conducted in **Step 2** and **Step 4** was covered by Durham University's public liability insurance and professional indemnity insurance.

6.4.10 Ethics

Ethical approval for **Step 2**, the focus groups, is reported in *Chapter 5* section 5.4.7 (p.395) of this thesis. **Step 4**, the pilot and content validity study, was also conducted

according to the guidelines of the Declaration of Helsinki, and ethical approval was obtained from the Department of Sport and Exercise Science Ethics Committee at the University of Durham (*SPORT-2021-06-23T15_37_44-pdhd55* on 09/07/21).

6.5 Results

6.5.1 Participants

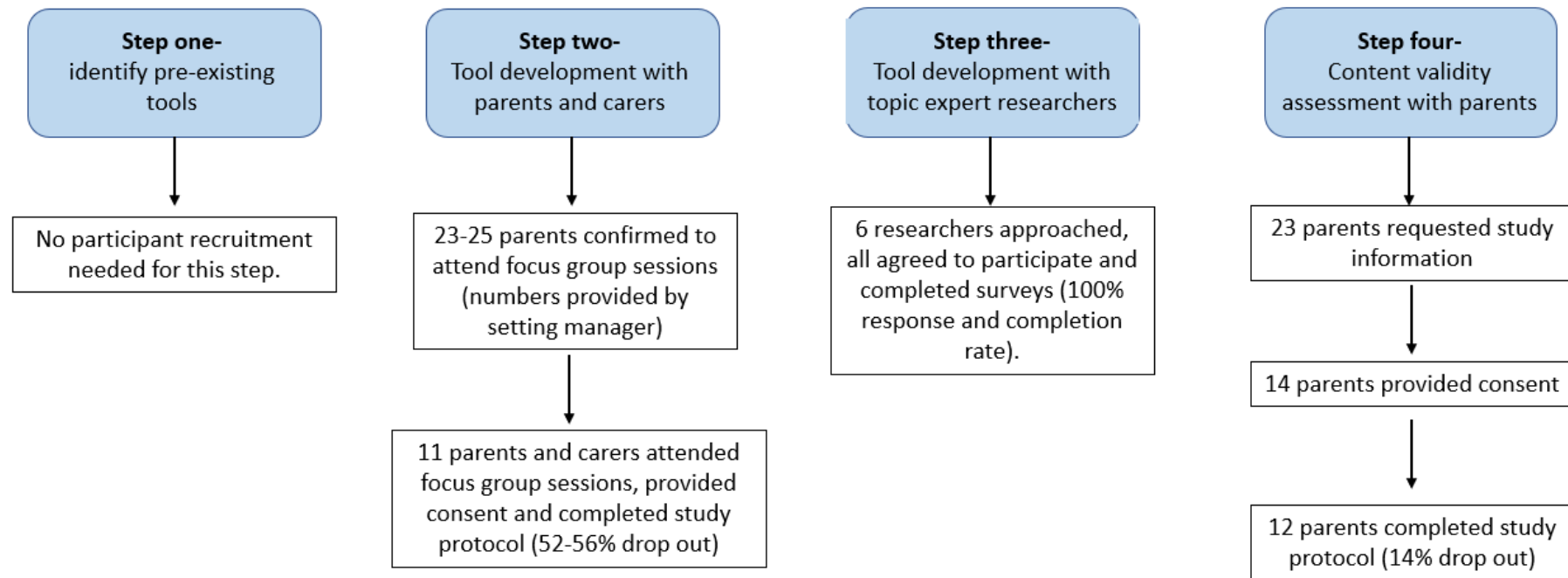


Figure 6.4. Participants for each step of the development and content validity assessment of MoveMEY

6.5.1.1 Step 1- Provisional development of MoveMEY

The methodologies of **Step 1** including systematic and literature reviewing, as such, no participants were required for this stage of the research.

6.5.1.2 Step 2- Focus groups with parents and carers of pre-school aged children

Three focus groups with a total of eleven parents and carers (including pre-school teachers) of pre-school children (aged 3-4 years) took place in this step. Participants are described in detail in *Chapter 5*, Table 5.2 (p.392).

6.5.1.3 Step 3- Topic expert researchers

Six topic expert researchers (response rate 100%) participated in the study. Researchers in the fields of movement behaviours, including being members of the early years expert committee for guidelines in the UK (Professor John Reilly, Dr Kathryn Hesketh), PA and SB measurement tool development (Dr Daniel Bingham, Dr Esther van Sluijs), sleep (Professor Helen Ball) and child public health (Professor Sonia Saxena). All researchers approached agreed to participate and are accredited in the declaration and acknowledgments for this chapter.

6.5.1.4 Step 4- Interviews with parents of pre-school aged children

Twelve interviews with parents of pre-school children (aged 3-4 years) took place. Table 6.1 provides information on the demographic characteristics of the participants. A total of 23 parents requested study information, of these, 14 provided consent, and 12 parents (reporting on 13 children) completed the study protocol. Of the two parents who provided consent but did not complete the study, one reported that both of her children

were unwell with Covid so could no longer participate, and the other was moving jobs and did not arrange the interview.

Table 6.1. Demographic characteristics of participants

Demographic characteristics		n=12 parents, 13 children
Sex of parent/carer (%)	Female	92
	Male	8
Ethnicity of parent (%)	White British	92
	Latvian	8
Age of parent (years)	Median	37
	Range	27-40
Parent's religion (%)	Christian	58
	No religion	42
Sex of child (%)	Female	31
	Male	69
Ethnicity of child (%)	White British	100
Age of pre-school child (years)	Median	3.8
	Range	3.1-4.11
Child's religion (%)	Christian	62
	No religion	38
Education level (%)	Masters/PhD or equivalent	17
	Bachelor degree or equivalent	75
	A levels or equivalent	8
	Diploma in higher education/BTEC or equivalent	0
	GCSE's or equivalent	0
	Vocational qualifications (NVQ Level 2)	0
	No formal qualifications	0
Employment status (%)	Working full-time	34
	Working part-time	50
	Full time stay at home parent	8
	Unable to work	8
Household income per year (%)	Under £2,500	0
	£2,500 - £4,999	8
	£5,000 - £9,999	0
	£10,000 -£14,999	8
	£15,000 - £19,999	8
	£20,000 - £24,999	0
	£25,000 - £29,999	17
	£30,000 - £34,999	0
	£35,000 - £39,999	8
	£40,000 - £44,999	8
	£45,000 - £49,999	8
	£50,000 - £74,999	8
	£75,000 - £99,999	17
	£100,000 or more	0
	Prefer not to say	8
Index Multiple Deprivation quintile (%) (1 = most deprived, 5 =least deprived)	1	42
	2	25
	3	17
	4	8
	5	8

6.5.2 Step 1: Provisional development of MoveMEY

During initial tool development, it was envisaged that the tool would take a proxy-report survey approach, as this is a more feasible method for large-scale surveillance (Bauman et al., 2006) due to often being most practical, low cost, low participant burden and generally accepted (Dishman, Washburn and Schoeller, 2001). The systematically conducted reviews for this research demonstrated that for PA and SB (*Chapter 3*), the most promising proxy reported tools for pre-school children were:

1) Pre-school age physical activity questionnaire (PRE-PAQ) (Dwyer et al., 2011).

However, this tool was designed to measure activity in the home environment only, and correlates of activity, meaning that the format and content would not provide overall activity levels, duration, or type of activity, which would be required to assess guidelines.

2) Netherland's PA questionnaire (Janz, Broffitt and Levy, 2005), which includes a Likert scale of activity preferences (e.g. prefers to play alone to prefers to play with other children), rather than recall of behaviours, and was not in a format that would produce results on overall activity, duration or type of activity, and would not be able to assess guidelines.

For assessing sleep in this age group, only one sleep diary (Tikotzky and Sadeh, 2001) was shown to be of reasonable quality and validity (see *Chapter 4*). Some of the relevant details from the diary were carried forward into the provisional tool (as discussed below), including a question on bed time, wake up time, and number of night wakings. The remaining questions used a Likert scale (e.g. sleep quality rated on 4 point scale from 0 (very good) to 3 (bad)), which would not be a suitable format for assessing the guidelines.

I also examined a plethora of literature that had assessed the movement behaviours of pre-school aged children at scale (>100 children in the sample) to determine questions and methods used (appendix 6.2). From this, I identified that data collection largely consisted of accelerometers being used to derive PA estimates, with questions asked for how many hours of screen time and sleep (e.g. Chaput et al., 2017a; Cliff et al., 2017; Berglind et al., 2018; De Craemer et al., 2018). Some studies used accelerometer derived estimates only (Meredith-Jones et al., 2019; Taylor et al., 2018); accelerometer derived PA and sleep, with questions on screen time (Carson et al., 2019; Chen et al., 2019) and accelerometer derived PA and SB, but parental reported sleeping hours (Carson, Tremblay and Chastin, 2017). For screen time, the questions most often included how many hours the child typically spent in various types of screen-based activities (e.g. watching TV, playing video games). Response options varied, including prescribed times (e.g. never, less than 30 min per day, <1 hour per day) or open responses. In one instance, questions were also asked to target the 'restrained' element of movement behaviour guidelines, this included asking the average time child was restrained in a motor vehicle e.g. in the car and a question on being safely restrained in a high chair/stroller (Kuzik and Carson, 2016). For sleep, the question was most often child's usual sleep duration, with questions such as *'How many hours does your child usually spend sleeping in a 24 hour period?'*.

Alongside this, I conducted searches of national surveillance systems for movement guidelines (as 24 hour movement and individual behaviours e.g. PA only) (see appendix 6.3). There was variation in the national surveillance systems, with some that did not capture the guidelines (NHS Digital, 2016), some focusing on screen time only (Ministry of Health New Zealand, 2017a), and average PA and sedentary screen based activities over the past 7 days (Australian Bureau of Statistics, 2013). The National Sleep Foundation (NSF) had an

extensive (>30 questions) questionnaire to assess sleep of pre-school children (National Sleep Foundation, 2002). Some questions for the provisional tool were adapted from this questionnaire; however, the NSF survey was too extensive to be used when assessing both PA and SB alongside sleep. The only surveillance system I encountered covering the whole 24 movement behaviours was the Canadian Health Measures Survey; this included accelerometer derived PA, screen time questions and average amount of sleep in a 24 hour period (Chaput et al., 2017a). Questions included how many hours the child spent engaged in screen time activities, and how many hours the child usually sleeps in a 24-hour period.

6.5.2.1 Format of MoveMEY

The results from the extensive searching as outlined above revealed that there was not a format that had previously been used, validated, and appropriate for the purpose of assessing the guidelines. As such, I expanded my searches to include relevant tools that measure movement behaviours of young children (outside of pre-school age range) for guidance on the preliminary formatting of the tool. This was only the preliminary development of the tool in order to present parents and carers with a stimulus for discussion, so whilst it needed to be formatted in some way, it was mainly to ensure that the questions covered what was needed to determine children meeting the guidelines. The questionnaire was separated into sections to capture the three main behaviours: PA, SB, and sleep. I used the format as outlined in the Early Years Physical Activity Questionnaire (EY-PAQ) for the PA and SB questions, which has shown reasonable validity and reliability (Bingham et al., 2016a). This tool has later been shown to be one of the highest quality pre-existing tools, in terms of reliability and construct validity, for measuring movement in young children (aged 0-5 years) (Arts et al., 2022). However, I changed the questioning from

‘last month’ to ‘last 7 days’ due to research suggesting this to be more accurate (e.g. Terwee et al., 2010). The format of the sleep section of the questionnaire was slightly different due to the complexity of sleep, and wanting to capture consistent wake and bed times, as well as sleep quality. This included modifications of various tools, including questions from the diary shown to be valid in *Chapter 4* of this thesis (Tikotzky and Sadeh, 2001), the most frequently assessed questionnaire, the Children’s Sleep Habits Questionnaire (Owens, Spirito and McGuinn, 2000) and the NSF survey (National Sleep Foundation, 2002). Modifications were made to ensure that the questions were able to detect the guidelines.

6.5.2.2 *Content of MoveMEY*

During the preliminary development of MoveMEY, the content of the tool was developed by ensuring that the questions were mapped on to the WHO movement behaviour guidelines (WHO, 2019). Figure 6.5 shows how each question mapped onto the WHO guidelines. There were very few questions that had been proposed previously to capture the guidelines (as outlined in section 6.5.2, p. 452-454). However, some questions used previously to assess the movement behaviour guidelines in large scale studies and as part of the Canadian surveillance system, in particular the questions on screen time (as listed above), were extracted across to form part of the tool.

For the PA section, there were three questions, broken down by type of activity: active play (outdoor and indoor separately) and active travel. During the pre-school years, much PA is achieved through active play. Guidelines now state specifically that children should engage in outdoor play. As such, in order to capture the context and type of activity and this intricacy of indoor/outdoor, it was important to separate the question on active play for both indoor and outdoor activity. This included stating that the responses should

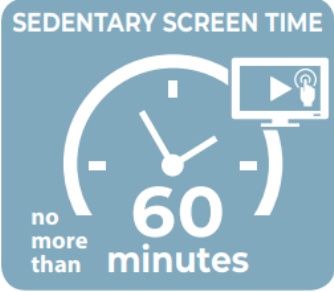
include PA through organised sessions (e.g. structured swimming sessions, football sessions). Other methods of PA include through active travel (e.g. walking/cycling to pre-school); therefore, there was a separate question to ensure that these activities were captured. All questions on PA also included sub-questions relating to whether the activity made the child sweat or breathe harder and the hours/minutes that this encompassed; the purpose of which was to determine the amount of MVPA the children engaged in.

For SB, three questions were proposed, that were separated to encapsulate the different elements of the guidelines, including: screen time, sedentary whilst engaging in additional pursuits (e.g. reading, storytelling), and whilst restrained (e.g. in a car seat / pram). Multiple examples of screen time were used to ensure that different screen exposures were captured (Byrne, Terranova and Trost, 2021). Questions included detecting the duration of time separately for each of the sedentary types of activity, due to guidelines noting differences between screen time and educational activities whilst sedentary.

For the sleep guidelines, five questions were proposed, regarding: bed time/wake time, consistency of bed/wake times, how long it takes child to fall asleep following bed time, night time awakenings and nap times. Bed time, wake time, how long it takes child to fall asleep, and nap times were to be used for sleep duration. How long it takes child to fall asleep and night-time awakenings were included to provide an indication of sleep quality, in line with the NSF sleep quality recommendations (Ohayon et al., 2017). Consistency of bed time and wake time was included to ensure the guideline on 'consistent bed and wake times' could be captured by the tool. The provisional tool can be found in appendix 5.11.

Figure 6.5. Provisional tool questions mapped to guidelines

Guidelines		Provisional questions		
Physical Activity 	Spend at least 180 minutes in a variety of types of physical activities at any intensity.	In the past week, how many days and for how long each day would you say your child has spent actively playing outdoors (e.g. playing in the garden/yard, playing at the park, riding a bike, jumping on trampoline, climbing, skipping, throw/catch, football sessions)?	In the past week, how many days and for how long each day would you say your child has spent actively playing indoors (e.g. messy play, dancing, running around, rough and tumble play, sit and ride/push toys, indoor play area, active computer games such as Wii Sports, gymnastics/swimming classes)?	In the past week, how many days and for how long each day would you say your child has spent travelling by walking or cycling (e.g. to and from school)?
	At least 60 minutes is moderate- to vigorous intensity PA,	Did any of this activity/active play make your child sweat or breathe harder? (Please circle)	Did any of this activity/active play make your child sweat or breathe harder? (Please circle)	Did any of this activity/active play make your child sweat or breathe harder? (Please circle)
	At least 60 minutes is moderate- to vigorous intensity PA,	If yes, how many hours/minutes of this activity/active play made your child sweat or breathe harder?	If yes, how many hours/minutes of this activity/active play made your child sweat or breathe harder?	If yes, how many hours/minutes of this time spent travelling made your child sweat or breathe harder?
	Not be restrained for more than 1 hour at a time (e.g. prams/strollers) or sit for extended periods of time	In the past week, how many days and for how long each day would you say your child has spent: in their pushchair, being carried, sat in the car/on public transport?		

Sedentary Behaviour 	Sedentary screen time should be no more than 1 hour; less is better	In the past week, how many days and for how long each day would you say your child has spent engaged in screen based activities (e.g watching TV, using a computer, tablet or mobile phone, playing computer games) whilst sitting, reclining or lying?		
	When sedentary, engaging in reading and storytelling with a caregiver is encouraged	In the past week, how many days and for how long each day would you say your child has spent engaged in other quiet? activities (e.g reading, storytelling, colouring, playing with toys), whilst sitting, reclining or lying?		
Sleep 	Have 10–13 hours of good quality sleep, which may include a nap	Please write your child's usual bed time and wake up time.	On average, from the time that your child goes to bed, how long does it take them to fall asleep?	On average, how many times and for how long each time does your child wake up during their night time sleep? On average, how many times and for how long each time does your child nap during the day?

	Regular sleep and wake-up times	On which days of the week is this usually the case? (Please circle)		

6.5.2.3 *Scoring of MoveMEY*

The PA and SB scoring of MoveMEY was based on the format of the EY-PAQ, as the format of this tool had been selected and therefore the accompanying scoring system should also be used (Bingham et al., 2016a). This would also provide appropriate response outputs to determine duration of the activities, which would enable guideline adherence to be determined. Each question asked ‘how often’ (number of days the week) and ‘for how long each day’ the child had engaged in the activity. The duration reporting options were 1) up to 15 minutes, 2) 16-30 minutes, 3) 31-45 minutes, 4) 46-60 minutes, or 5) if more than an hour per day, estimated time in hours and minutes. Unless free text responses were used, the strategy for scoring was to use the lowest threshold for PA (due to usual over-reporting) and highest threshold for sedentary behaviour (due to under-reporting). Therefore, daily PA for each question could be calculated by the duration of activity each day (e.g. 16 minutes) multiplied by the frequency that activity occurred (e.g. 5 days), divided by 7 to estimate daily minutes of PA. Estimated daily PA for each of the three questions asking about PA would be summed to provide total daily PA. For determining MVPA, free text responses of hours and minutes were used, so output could be calculated based on the total of the responses inputted by parents divided by 7 to determine daily MVPA. Daily sedentary time for each question could be calculated by the duration of activity (e.g. 45 minutes) multiplied by the frequency that activity occurred (e.g. 5 days), and then divided by 7 to estimate daily minutes of SB. Estimated daily sedentary time from each of the three questions asking about SB would then be summed to provide a total daily sedentary time score. To determine individual guidelines such as screen time, just the screen time question would be calculated. For the sleep section, total sleep duration was determined using total time from time in bed minus time to fall asleep minus total minutes awake during the night

plus total nap duration. Questions on bed time, wake up time and time to fall asleep were for what is 'usual', as such, these applied for every day of the week. Questions on night wakings and nap time were daily and the same scoring system as the questions on PA and SB was applied, whereby the duration of nap or night waking (e.g. 16 minutes) could be multiplied by the frequency that activity occurred (e.g. 2 days), and then divided by 7 to determine a daily average. The strategy for scoring was to use the highest threshold for night wakings (as these are usually underestimated by parents) and lowest threshold for nap duration (due to overestimation of sleep duration in this age group). 'Good quality' sleep was determined based on sleep efficiency, which included $\geq 85\%$ total sleep time from total time in bed to be classified as good sleep quality, based on NSF recommendations (Ohayon et al., 2017). 'Consistent' wake and bed times was determined to be if these were reported as consistent for ≥ 5 days, based on existing questionnaires using this definition (Owens, Spirito and McGuinn, 2000). Previously, researchers have used the following question to detect consistency of bedtime routine previously '*Does your child have a usual bedtime routine, that is, on most nights, do the same activities occur?*' (Mindell et al., 2009). Therefore, although there is currently no consensus in the literature on what is classified as 'consistent', this definition also made logical sense with being more consistent than not. The full scoring sheet for the provisional tool can be found in appendix 6.4.

6.5.3 Step 2: Focus groups with parents and carers of pre-school aged children

6.5.3.1 Format of MoveMEY

In this section, I will present the key findings from parents and carers of pre-school children on the format of MoveMEY. An overview of the key findings are displayed in table 6.2.

Table 6.2. Key findings on the format of MoveMEY from parents, carers and topic expert researchers

Parents/carers	Topic expert researchers
Daily reporting	Daily reporting
Diary format	Tool may have to be longer to be more meaningful
Distinguish between weekday/weekend	Distinguish between weekday/weekend
Enable space to report on home and school activity	Needing to capture home and school activity
Report nap time / night wakings daily as they differ	Report nap time / night wakings daily as they differ
Preferences differed in terms of entering own amount of time for each activity versus ticking boxes	Parents to enter own amount of time for each activity may make the tool less complex
Sleep section slightly different – sleep routines mean that bed/wake time question would not need to be daily	
	Some behaviours may be reported twice – make clear from the start what questions are ahead to avoid this happening

6.5.3.1.1 Capturing different routines

A main reoccurring theme through the focus group discussions in relation to reporting children's movement behaviours was differences in routine that result in different behaviours. The main concern was the difference and ability to detect school and home activity, and for parents to be able to capture this information when they are not with their child. *'I think the only problem with this one is that if they're at school or nursery, again, it's*

quite a long questionnaire for staff to fill in.' (P5). It was suggested that by having space to report for home/school activity separately, it would be easier to coordinate this with teachers. With this, a nursery teacher in the focus group said they could complete the tool during the day and give it back to the parent at handover. Secondly, the differences between weekday and weekend day activity were reported, stating that weekends are usually different from weekdays and activity can vary so it would be important to capture these differences: *'Yeah. That's a weekday, but then during the weekend it's-* (P4) *'It's a bit different like us.'* (P2) (Focus group 1 extract).

6.5.3.1.2 Daily reporting

In light of this, it was suggested that the preliminary draft of the MoveMEY tool was challenging as parents reported that their children participate in a large number and variation of activities, and it can change every day. Therefore, it would be very difficult to have to try recall and determine how much activity their child had done over the past week. *'I think it's a bit too hard for me this, because like she does so much each day, like all the different things she does through the day, and then you're having to record them and remember, oh, it's so confusing, I just don't understand it'* (P3). Others suggested that the tool was understandable: *'straightforward'* and *'easy to follow,'* but that it was quite long so may be a lot to fill in. Instead, it was proposed that daily reporting would be much easier and more convenient than retrospective reporting: *'I think maybe like this with different days of the week. So you're not having to sit and work out each day what they've done, at least you can have it as like a daily thing.'* (P5).

Parents suggested that they would rather write their own amount of time into boxes rather than tick how long their child had been doing the activity *'Like where you could write*

your own amount of time, like roughly' (P3), whereas other parents felt that ticking the amount of time spent engaging in the activity would be preferred *'Yeah, tick box. That'd be so much easier, yeah'* (P7). It was proposed that a diary based format, with activities down the side and space to report individually for each day, for both home and school activity would be most beneficial. With this, it was also reported that it would be easier to capture the nursery time activity as parents could ask the teacher at the end of each day when collecting their child. Parents did note that daily reporting for different activities would create *'more paperwork'* but that *'in the long run it'll be easier'*.

'Yeah, I don't know. It might look like a lot more, but it would be so much easier if you just did it every single day. If you just did it, if you had all this paperwork together and you just decided every day at this time you're going to fill it all out, it's a piece of - it's only one week, isn't it?' (P4).

Other formats considered by some parents were to separate the tool so that it was daily but have space to report activity for different segments of the day (morning, afternoon, evening) where parents could freely fill in what their child had engaged in during those times, as they felt this would help simplify the tool.

The structure for reporting the sleep section was regarded slightly differently, whilst in most instances parents still felt that daily reporting would be better. For some questions such as bed and wake times, parents suggested that a format of 'usual' may work due to routines that parents have with children of this age and these being fairly consistent, but felt it was important to keep the space to report differences and reasons for the difference. For specific behaviours, such as nap time and night wakings, it was suggested that reporting daily on these would help give a more accurate estimate as the format of the provisional

tool (how many days and how long each day) could inflate the estimate substantially.

Parents felt that daily for night waking was important not only because this can vary so much per night, but also because they may not be able to appropriately recall *'When you suffer from lack of sleep remembering what you did it's like, right, how long was she up for on Monday.'* (P5). Parents also suggested that even if their child does not usually nap, some circumstances may mean that children have a nap, for example: if on a long drive, if they are unwell, or if they have had a particularly busy day. *'And then it's like the number of times per day, well, if it's only once per week then if you put once, like say she's poorly and I put that she's napped once, you're just going to think that she's napped once every day for however long, whereas I think if it's daily.'* (P5).

6.5.3.1.3 Visual of the tool

In terms of the visual of the tool, some parents suggested that it would be better if the tool looked visually appealing: *'Maybe colour coordinate it... just pleasing to the eye and like you...you want to fill it out, because it's... looks more pleasing'* (P4). Whilst others stated it would be better to be *'as boring as possible so the kids don't get attracted to it.'* (P8).

6.5.3.1.4 App versus paper format

A suggestion in one focus group was that a phone application would be easier to use than a paper based diary, as they were concerned that they may end up losing the paper. It was suggested that an app or something similar, that could be used on an electronic device, would be more convenient if asking different caregivers to fill in the tool (e.g. parents, nursery teachers). This is illustrated in the extract below:

‘Yeah, so, say like on a tablet, if you could put it on a tablet you may be even able to put it on your child’s tablet, so then if they did go to nanny or grandma’s they’d take their tablet with them nine times out of ten and then just have a section that they can then log in and then just go yeah, yeah, yeah-’ (P8).

6.5.3.2 Content of MoveMEY

In this section I will present the key findings from parents and carers of pre-school children on the content of MoveMEY. An overview of key findings are displayed in table 6.3.

Table 6.3. Key findings on the content of MoveMEY from parents/carers and topic expert researchers

Parents/carers	Topic expert researchers
The provisional tool was relevant, included relevant activities, and covered most of the ‘basic’ activities that children of this age engage in.	The provisional tool was thought to be relevant and provided an adequate reflection of the movement behaviours of this age group, and appropriately targeted the guidelines.
Locations of PA /active play, rather than specific activities.	Locations of PA /active play, rather than specific activities.
Distinguish between different types of seated travel (e.g., separate spaces for travelling in car, on the bus).	Concern that seated travel question may not appropriately capture guidelines – separate the question to distinguish between different types of seated travel independently.
Space to add some additional activities of their own.	Additional activities to be added: ‘scooter’ (for leisure and active travel), soft play and puzzles. ‘Bath time’, ‘sitting whilst eating’ and ‘sitting whilst on the potty’ should be added, to ensure that these seated behaviours were captured.
Term ‘sweating’ not a word parents would use in relation to their pre-school child.	‘Out of breath’ or ‘breathe harder’ may be more relevant alternatives to the term ‘sweat’ for this age group.

Space to report 'general information' such as illness, sleep problems or disabilities.	Space to report 'general information' such as illness and sleep problems.
Children sometimes use screens before bed so unsure on sleep time.	Screen time before bed may be important additional factor to capture.
Question on when bed/wake time was consistent- space to explain why differences may be present on certain days.	Question on when bed/wake time was consistent- space to explain why differences may be present on certain days.
Allow space for active travel versus walking/cycling for leisure.	
Space for 'additional comments' to help put the responses into context.	
	Specific activities for sedentary behaviour and screen time
	Report '0' if activity didn't happen
	Add lay definitions of the behaviours to be captured

6.5.3.2.1 Relevance of items and comprehensiveness of the tool

The content of the provisional tool was generally well received, and parents stated that the activities outlined were relevant '*captured most of what they do.*' (P11) and comprehensive in terms of the activities that children engage in '*...I think you've covered the basic ones.*' (P2). It was suggested that in some instances including location of activity may help with recall, such as the question on '*outdoor activities*' including options such as '*in the park*'. It was also suggested that some example activities should be included (including those already outlined), to help parents understand the general idea of the question, alongside some blank spaces so that parents could fill in any additional activities not covered by the generic tool. '*So we could just have examples so people don't think hmm, what have they done?...*' Yeah, but then have the choice to also write your own if you can do that.' (P4).

Parents recommended having '*additional comments*' sections, to be able to explain any responses in more detail or say what happened on particular days '*...allowing people to actually write down the things as well*', including if a child was with another member of the family or childminder. Further, parents felt this would allow the opportunity to provide any context about sleeping arrangements, such as if the child shares their room with a sibling or changes bed during the night. Most parents said they had typical bedtimes for their children but suggested that being able to have space to say why certain bed/wake times might be different would be helpful to explain the differences.

Some specific issues raised by parents were that it may be important to distinguish between the different types of seated travel, as not all were relevant to them, and they were unsure whether it would be important to make these distinctions more clearly. '*...So we're in the car quite a lot when we go out because my partner drives so we go everywhere, because he's just pure lazy. He won't even walk to the corner shop! So I don't know if being in a car or on public transport would maybe be relevant.*' (P9). Additionally, one parent suggested ensuring that there was a specific space to report walking/cycling for leisure, rather than just as active travel, to make sure that all of these behaviours were detected. Further, parents reported that they do not always know when their child falls asleep, as even when they put them to bed they may watch TV or play on a tablet before sleeping. This highlighted that screen time before bed is something children of this age may engage in, and as such, may be an important factor to include in the tool.

The terminology of 'sweating' was something that some parents said they would not really say about their young children, and so alternative terminology may be better to capture higher intensity activity: '*I wouldn't ask them if she was sweating, I know my kids*

probably would anyway, but you can just say to her, has she played out and how long for, or something like that.' (P5).

6.5.3.2.2 General information

Parents suggested that it would be important to report on some 'general information' about their child that would help put their child's behaviours into context. Parents suggested that reporting if their child is suffering from illness would be important as this changes their behaviours, as demonstrated by this quote: *'...there should be a box here if they were ill, because [child] she never naps now, but if she's ill she would probably sleep nearly all day'* (P10). Secondly, reporting 'sleep problems' was suggested, so that parents are able to say if their child suffers from any night disturbances such as night terrors or to report the 'type of sleeper' that their child is. As night-time awakenings are common in pre-school children (National Sleep Foundation, 2004), the question in the tool stated 'medical' sleep problems, with space to state information on the sleep problem. A final point raised was that behaviours may be different if children have a disability that impacts on their movement; suggesting that inclusion of a question on whether their child has a disability that may impact on their movement behaviours may be beneficial.

6.5.4 Step 3: Qualitative surveys with topic expert researchers

6.5.4.1 Format of MoveMEY

In this section, I will present the key findings from topic expert researchers on the format of MoveMEY. An overview of key findings are displayed in table 6.2 (p.459).

6.5.4.1.1 Daily reporting

Topic expert researchers consistently suggested that questions should be separated to report for home versus preschool/nursery and to capture the weekday and weekend

differences in activity and sleeping behaviours. This was mainly due to activity varying substantially depending on the structure around the day of the week, but also because parents recalling movement behaviours during nursery time would be very difficult. Daily reporting was also stated specifically in relation to questions on nap time and night wakings, as it was suggested that this would vary by day so an average may not necessarily work as it could lead to an over or underestimation of time. It was suggested that night wakings may be difficult to report regardless of how this is asked; however, it was stated that it is plausible that this could be made more accurate by being able to distinguish between days, rather than having a weekly average.

The complexity of measuring the movement behaviours of pre-school children through a proxy reported tool was noted. The trade-off between brevity of a tool and the amount/type of responses was raised. It was also noted that this could potentially lead to a longer and more complicated tool, but this could result with the tool being more meaningful for parents to complete. Additionally, the concern of recall instruments and parental perceptions of children's PA were reported as problems with parental reported tools.

In line with this, it was thought that the format of reporting how many days and how long each day was too complex and breaking this up further would help with clarity and recall. Changing the format may help to simplify the tool, ensuring there is sufficient space for parents to report answers to help with ease of completion. Alternative solutions were to propose an 'at a glance' week of activity/table to reduce difficulty of recall. It was also suggested to have different sub-sections to the tool to break things down further including general health, PA at home, school, organised activity, and active travel, screen time activities at home, school, seated travel, other sedentary behaviour and sleep.

6.5.4.1.2 Ensuring behaviours are not reported twice

Expert researchers highlighted some instances where activities may be recorded twice. This was primarily an issue for distinguishing active travel from leisure-based activity (e.g. walking/cycling for leisure). It was suggested that clear sections that outlined the forthcoming questions in advance may help avoid activity being recorded in the incorrect place, for example stating that there would be questions on 1) outdoor active play, 2) indoor active play, 3) active travel. A further instance in which this was a concern was naps taken 'on the go', for example naps taken when travelling in the car, or in the pram, as this same period of time may be accounted for in two separate locations of the tool.

6.5.4.2 Content of MoveMEY

In this section I will present the key findings from topic expert researchers on the content of MoveMEY. An overview of key findings are displayed in table 6.3 (p.463).

6.5.4.2.1 Relevance of items

Topic expert researchers felt that the questions in MoveMEY were relevant and represented an adequate reflection of each of the movement behaviours of pre-schoolers, and appropriately targeted the guidelines. For the question targeting MVPA of pre-school children, it was suggested that 'out of breath' or 'breathe harder' may be better (and more relevant) alternatives to the term 'sweat'.

It was suggested that asking questions on specific types of PA, rather than having lots of different activities may be beneficial in aiding recall. In particular, location specific questions were recommended to help with recall for outdoor PA, rather than specific activities (e.g. time spent climbing, time spent running). In line with this, consistency is

important, to ensure that the tool does not have some locations e.g. 'play in the park' as well as activities e.g. 'climbing', as the activity may occur in the location and this could lead to replication of responses. On the other hand, outlining the specific activities that children engage in for sedentary behaviour and types of screen time may help with recall and help to *'focus the mind'*, rather than grouping 'screen time' together as one overarching question.

Suggestions were made to add some additional activities as examples of what the questions were asking, including: 'scooter' (for leisure and active travel), soft play and puzzles. It was recommended that activities such as 'bath time', 'sitting whilst eating' and 'sitting whilst on the potty' should be added, to ensure that these seated behaviours were captured.

6.5.4.2.2 [Comprehensiveness of the tool](#)

Generally, it was reported that the items on the tool were representative of the guidelines. A concern raised was the difficulty with making the distinction between standing and sitting behaviours for children of this age, as they may quickly go from one behaviour to the next. Other instances of concern that topic experts felt needed addressing included that the item on seated travel may not necessarily capture the restrained guideline, as for example, children sitting on public transport would not necessarily mean they were restrained during this time. It was rather suggested that the use of the word 'restrained' with specific examples may help to more accurately detect the guideline. Alternatively, it was suggested that this question should be separated by the different modes of sedentary behaviour so that parents would recall for each behaviour and distinctions could be made between seated activity and restrained activity (e.g. seated whilst on the bus, seated in the car).

An additional recommendation was that it may be beneficial to capture screen time before bed. Although this was not put forward as a UK or WHO recommendation, accumulating evidence in this area (see Janssen et al., 2020) meant that it was recommended by some expert committees, and some experts felt it would be useful information to collect.

Regarding sleep specifically, topic expert researchers stated that it was important to distinguish between in/out of bed and sleep versus wake. Further, a question to achieve consistency of bed time and wake time was presented in the provisional version of the tool, with parents to circle on which days this was the case. Topic expert researchers suggested that space should be provided here so that parents could supplement their responses with why things were different on certain days.

6.5.4.2.3 Understandability of the tool

Generally it was suggested that the tool, and the individual items of the tool were comprehensible. There were a few suggested additions to help with clarity of the tool including adding definitions in a lay format on the key behaviours (PA, SB, and sleep) on the first page. Ensuring that there was sufficient explanation on what the tool was aiming to capture (particularly if a researcher would not be present for data collection), including making sure that technical language was not used. It was also suggested that it should be clear to participants to mark '0' when the activity did not happen, so that it was clear between missing data/activity not happening.

6.5.4.2.4 Reporting contextual information

It was suggested that there should be a question on 'sleep problems' such as excessive night wakings and night terrors, as this could help explain the final results of someone reporting on child's sleep. Although no definition on what was considered 'excessive' was provided, the question was made explicit to be 'medical' sleep problems and space was provided in the question for parents to provide information on what the sleep problem was. Similarly, it was suggested that a question on whether the child has an illness at the time of reporting would be important to include. As well as knowing if the child is attending nursery or preschool, and on which days this is the case.

6.5.4.3 Scoring of MoveMEY

Concern was raised that using the current format of the provisional tool, scoring would result in a weekly average of activity over 7 days, providing a mean 'usual' time of each activity rather than a specific amount of daily activity. Although this would help determine usual activity, it would not be able to accurately detect whether children meet the guidelines in a 24 hour day, and as such would not appropriately capture the guidelines. It would be preferable to know how many days children are active/sedentary and the average PA and SB on those days, rather than average across the week; to extract daily and weekly activity. There were few comments on the sleep scoring, however, it was suggested that a clear definition on what constitutes 'consistent' bed and wake times would be important to rationalise.

Alternative suggestions were that parents could put in their own amount of time for each activity rather than keeping within prescribed times, as prescribed times may add to the complexity of completing the tool. A main concern for scoring was what would happen if

values equated to more than 24 hours for any one day. Additionally, researcher burden of calculating this for large samples of participants was raised as a problem.

The revised MoveMEY measurement tool following this co-development process outlined in **Step 2** and **Step 3**, which was then piloted in the next section can be found in full in appendix 6.12. The accompanying scoring sheet for the tool following co-development can be found in appendix 6.15.

6.5.5 Step 4: Interviews with parents of pre-school aged children

Following the pilot study, there were no major concerns with the tool that had been revised following **Step 2** and **Step 3** (appendix 6.12). The tool was stated to contain relevant items, was comprehensive to detect the range of children's movement behaviours, and there were no concerns with understandability of the tool and instructions. Despite this, the pilot study was a useful exercise to improve the tool and make some amendments in line with parental feedback. In the sections below, I discuss in detail the feedback on the format (including response options and recall period), content, practicalities, and scoring of the tool. I also discuss some overall feedback from participants on their experience of taking part in the research, specifically related to their views on the tool. I clearly state where any alterations to the tool were made in text, but these can also be found in Table 6.4. The tool piloted can be found in appendix 6.12 (and scoring sheet in appendix 6.15), and the revised and finalised tool can be found in appendix 6.16 and the scoring sheet in appendix 6.17.

Table 6.4. Amendments to the tool

Question in MoveMEY	Concerns	Revised version
Part one: General Questions		
1. Does your child have any physical, neurodevelopmental, or medical condition or disability that affects their ability to play and be physically active?	No concerns	No amendments made
2. Does your child have any medical sleep problems, such as night terrors?	No concerns	No amendments made
3. Is your child currently suffering from an illness that may affect their normal behaviours, including being active, movement, sitting or sleep?	Question appearing as though it is only aimed at long term illness, rather than just being unwell during the week of measurement	<i>Changed to include more example so that even short-term illnesses are captured. 'Is your child currently suffering from an illness, unwell or poorly (short or long term) that may affect their normal behaviours, including being active, movement, sitting or sleep?'</i>
Part Two: Physical Activity		
4a. Please state how many hours and minutes your child spends actively playing outdoors in each of the following (activities may include: running around, jumping on a trampoline, climbing, skipping, throw/catch).	Columns for home and nursery activity, but specific row was also available for 'indoors at nursery'.	<i>Crossed out 'home' boxes for the 'outdoor play at nursery' row</i>
4b. Did any of these activities make your child 'huff and puff' or breathe harder? (Please circle)	Parents did not always feel their child would be out of breath or 'huffing and puffing', but that a range of activities would suggest this intensity of activity.	<i>Added more examples of higher intensity activity, including makes child breathless, hot and sweaty, or need a drink or rest</i>
4c. If yes, please state how many hours/minutes of this activity made your child 'huff and puff' or breathe harder.		

5a. Please state how many hours and minutes your child spends actively travelling , which could include travelling for leisure (e.g. to/from school, the shops, the park) each day.	No concerns	No amendments made
5b. Did any of these activities make your child 'huff and puff' or breathe harder? (Please circle)	Parents did not always feel their child would be out of breath or 'huffing and puffing', but that a range of activities would suggest this intensity of activity.	Added more examples of higher intensity activity, including makes child breathless, hot and sweaty, or need a drink or rest
5c. If yes, please state how many hours/minutes of this activity made your child 'huff and puff' or breathe harder.		
6a. Please state how many hours and minutes your child spends actively playing indoors (activities may include: dancing, running around, rough and tumble play, sit and ride push toys).	Columns for home and nursery activity, but specific rows were also available for 'indoors at home' and 'indoors at nursery'.	Removed columns for home/nursery, as there was a row for each of these locations.
6b. Did any of these activities make your child 'huff and puff' or breathe harder? (Please circle)	Parents did not always feel their child would be out of breath or 'huffing and puffing', but that a range of activities would suggest this intensity of activity.	Added more examples of higher intensity activity, including makes child breathless, hot and sweaty, or need a drink or rest
6c. If yes, please state how many hours/minutes of this activity made your child 'huff and puff' or breathe harder.		
Additional comments on physical activity	No concerns	No amendments made
Part Three: Sedentary Behaviour		
7. Please state how long your child spends in screen based activities whilst in a sitting, reclining or lying position.	No concerns	No amendments made
8. Please state at what time your child last uses a screen before going to bed (e.g. if child watches a film before bed).	Screen time before bed, some children do not watch screens after morning, but parents still found this self-explanatory.	No amendments made
9. Please state how long your child spends playing and in other activities whilst sitting, reclining or lying, including quiet or carpet time .	No concerns	No amendments made

10. Please state how long your child spends seated whilst travelling.	No concerns	No amendments made
Additional comments on sedentary behaviour	No concerns	No amendments made
Part Four: Sleep		
11. Please write your child's usual bed time and wake up time.	Some parents had variation in bed, sleep, wake and out of bed time – daily reporting of these factors may increase accuracy of tool.	Changed to daily reporting of bed time, sleep time, wake time, out of bed time.
12. On which days of the week is this the case? (Please circle and provide a reason in the box if different)	Removed question due to change in format for sleep questions, meaning that consistency can be detected by the tool without this additional question.	
13. From the time that your child goes to bed, how long does it take them to fall asleep?	Some parents had variation in bed, sleep, wake and out of bed time – daily reporting of these factors may increase accuracy of tool.	Changed to daily reporting of bed time, sleep time, wake time, out of bed time.
14. From the time that your child wakes up, how long does it take them to get out of bed?	Some parents had variation in bed, sleep, wake and out of bed time – daily reporting of these factors may increase accuracy of tool.	Changed to daily reporting of bed time, sleep time, wake time, out of bed time.
15. Please state how many times and for how long each time, that your child wakes up during their night-time sleep.	No concerns	No amendments made
16. Please state how many times and for how long each time that your child naps during the day.	No concerns	No amendments made
Additional comments on sleep	No concerns	No amendments made

6.5.5.1 Format of MoveMEY

Parents consistently stated that reporting activity every day was beneficial suggesting that otherwise they would forget the activity their children engaged in. *'But certainly, I think I needed to do it every day 'cause otherwise I would forget what we'd done'* (P14). Parents discussed that due to the sporadic nature of their child's movement even recalling daily activity can have some difficulties, as their children do so much in a single day. Some also suggested that the weekly overview helped to show the reality of their children's movement which may not have been achieved through a few days of reporting *'...Erm... so, if you just picked like two or three days potentially it would have given you more of a skewed reality. Whereas the week, sort of it gives more... it gives more of a chance to show reality, if that makes sense?'* (P23). As suggested during the development of the tool (**Step 2** and **3**), weekend activity was often different and so being able to make the distinctions between the days was important to accurately capture movement on each day more readily.

Parents suggested that reporting hours and minutes of PA and SB was the best way to recall and worked well *'...cause you would need to know the time to see the impact. So, I don't think there's any... there's no point in putting session, the amount of sessions or anything. I think hours and minutes is... is the right way to go.'* (P17). The sequence of questions was well received across all participants, with parents suggesting the following: *'I think they were fairly erm... they flowed pretty well to be honest. Erm... so... no, it made a lot of sense that the order that you had them in made a lot of sense'* (P12).

One of the main concerns with the format of the diary was confusion on the question *'Please state how many hours and minutes your child spends actively playing*

outdoors in each of the following (activities may include: running around, jumping on a trampoline, climbing, skipping, throw/catch).' The first response option of this question was *'Outdoor play at school'*, however the question columns were split to report on *'home'* and *'school'* activity. Parents stated that this sometimes resulted in errors with reporting, using the columns as hours and minutes in the first instance, until they realised. Others said they went through at the beginning and crossed out all the boxes that were not relevant for particular activities e.g. *'home'* boxes for the question on outdoor and indoor play at school, *'school'* columns for the weekend and question on active play in the home. It was suggested that blocking out the irrelevant boxes may help with clarity of the tool, which has been implemented as a change to the diary following the pilot (see Table 6.4, p. 474).

The sleep section of the diary had a slightly different format, with usual bed and wake time, with an additional question on which days was this the case and any potential reasons. Many parents reported that these questions on sleep worked well as they have strict routines with their children. Others stated that there could be variation in the bed and sleep time of their children: *'I mean normally there can be a bit of fluctuation some days... so, yeah, I think that potentially given the option to say each time might have been like more reliable'*. (P23). Additionally, on receiving the diaries some parents had put average times e.g. bed time between 6.45 and 7.15pm. Further to this, parents commented on the bed and sleep time difference of their children *'Because like where wasn't a difference with bedtime, so I didn't write anything in but there was a difference with sleep time.'* (P14). In light of these comments and given the layout of the remainder of the diary, an amendment was made to this section so that these items could be reported daily to increase accuracy of the tool and help parents with this reporting (see Table 6.4, p. 474). *'Much easier to then track actually how long did it take him to fall asleep... Almost to have like a bedtime and*

sleep time kind of separate thing.' (P14). With this, there was the removal of the question *'On which days of the week is this the case? (Please circle and provide a reason in the box if different)'*, as this information could now be determined from the revised format (see Table 6.4, p. 474). The sleep questions on nap time and night wakings were daily, and these were well received. Parents stated that due to such variation between the days for these indicators, it was best to report them daily.

'Like he had... we only had one day where I hadn't... where he had a nap and that was coming back from his grandparents' house in the car.' (P14)

'...cause some nights she doesn't wake at all and other nights she wakes two times, two or three times, you know, that sort of-...' (P16).

Some parents suggested that the use of a phone application would be *'a lot easier'* and that this is something they have used voluntarily for tracking behaviours with their children, as exemplified in the extract below. The use of a tool similar to what parents have used voluntarily may increase uptake, compliance, and acceptability of tools.

'I mean I always find kind of doing things on Apps a lot easier. So, like... you know when, especially with the twins when they were younger, I used to have erm... like a nap reporting App where you could... or like a feeding App, so you could kind of track everything while it was happening. So, at least it was kind of real-time.' (P22).

There were some different interpretations on what the diary would look like, which may impact on future compliance if parents have a misinterpretation on how we define *'diary'*. For example, one parent stated: *'...like an A4 page like each day of like what we've done...'* (P20) and another stated *'...I wasn't quite sure what... how it was going to be laid*

out. Erm... so, I think maybes there was... more detail potentially than I was expecting but I liked that.' (P23).

One parent suggested that an example of how to complete the tool at the beginning may help so that it is clear what is expected. In line with this, it was requested on the tool to *'Put a 0 if the activity did not happen'*, although the majority of participants did this, in some instances the spaces were left empty. In the interviews, parents stated that they did not put a response and left spaces blank when their child had not engaged in the activity. Although this may not be too problematic, it may make it more difficult to distinguish between missing data and the activity not happening. *For clarity, an example row has been added to the beginning of the tool.*

6.5.5.2 Content of MoveMEY

The tool was comprehensible, as parents consistently reported that the tool was *'easy to understand'*, *'straightforward'*, and *'self-explanatory'*. Parents felt that it was *'well worded'*, *'well structured'*, and *'well designed'*. Parents felt that the tool was comprehensive and relevant, in that it covered all *'age appropriate activities'* that their children would engage in, as exemplified below:

'Erm... I think it... everything that a child her age does is covered and more to be honest...' (P14). *'... it all made sense. I think, you know, it was all... there was sort of different options for everything that we could possibly be doing, so that was good.'* (P17)

The questions on PA and SB had space for parents to enter other activities that were not outlined on the tool. In some instances parents did not feel it was necessary to add any additional activities as they felt everything had been covered: *'...I didn't put anything in, in*

sort of the 'Other Activity' just because everything else had... had covered-' (P17). Other parents felt it was good to have additional space to write activities that would not be relevant for everyone, for example, *'playing in the barn', 'allotment', 'friend's party with bouncy castle', and 'going to the cinema'*. This highlighted that whilst the tool was relevant and comprehensive enough to detect the majority of behaviours, there are some more family specific activities that it is also important to capture.

One parent suggested that it would be easier if the SB questions were grouped together e.g. all screen time. Others stated that having the different types of SB split helped to think more in depth about the different activities their child does.

At the end of each section there was a space for 'additional comments', some parents reported that they liked this and that they felt they could add some context around their responses. Others did not use this space, as they were not sure what sort of comments to write. As this question is to make the tool more meaningful to parents, rather than to extract certain information, these sections remain unchanged.

One question that was problematic for interpretation was the question targeting higher intensity PA *'If yes, please state how many hours/minutes of this activity made your child 'huff and puff' or breathe harder.'* Parents reported feeling as though their child *'doesn't really huff and puff, just goes a bit sweaty and hot'*. Parents felt their children were only ever breathless when they had asthma or suffering from a fever. Alternative suggestions for the wording of this question were proposed, but varied between participants, including *'make your child breathless', 'short of breath' 'get hot and sweaty', 'gets a bit tired'*. It was also suggested that example of things that children do when they reach this level of intensity may be beneficial to put so that parents interpret the question

as such: *'just like asks for a drink or goes to get the water or something like that.'* Many parents stated that 'soft play' was an example where their children really exert themselves.

As such, this question has been altered in line with these suggestions to now include the different examples so that depending on how children respond to higher intensity activity will be captured (see Table 6.4, p. 474).

During the period of measurement, a proportion of the children were suffering from illness. Parents reported that this impacted on their child's behaviours and so it was not reflective of 'typical' activity. Some parents said the question on illness was helpful to show why the behaviours may be different to what their child would usually do. However, one participant interpreted the question to mean long-term illness *'I just presumed like an illness, you know, like a long-term illness for some reason... I didn't think that, you know, him being poorly for the short spell.'* (P18). As such, this question has been amended to ensure that even when children are just poorly or unwell for the week that this is appropriately captured (see Table 6.4, p. 474).

Some parents acknowledged that some other behaviours may be missed but would potentially be difficult to capture in any tool: *'general kind of toddling about... you know, just little things like that obviously you can't really cover that.'* (P22). However, this was represented as a limitation of proxy reported tools.

Some parents felt that the terminology of 'school' was not so appropriate for their child, with 'nursery' being the most frequently used term. It was stated that although it could be confusing, parents felt as though the tool was aimed at a child their age so they assumed it meant nursery rather than school. The terminology used varied between participants, one parent suggested the following *'Maybe change school to education setting,*

because like the little ones aren't going to school' (P13). As such, the wording of the questions has changed slightly to make this distinction more clearly, and 'nursery' is written throughout rather than 'school'. There is also an explanation within the tool as before describing what is included under the headings 'home' and 'nursery'.

6.5.5.3 Practicalities of completing MoveMEY

The majority of parents suggested that the tool was not time consuming and that they would sit down at the end of the day and complete the diary for that day, with ease. For example, one parent stated: *'But yeah, I think it's all right, it's not a massive task, it's just a couple of minutes, a few minutes a day, you know, just to write a quick diary entry, isn't it?' (P18)*. However, one parent found that the tool was quite time consuming because they felt they were always thinking about what their child was doing so that they could accurately report, as demonstrated in this abstract. *'...And for me personally, you know, when you're quite busy anyway erm... it's quite a commitment to think about, you know, breaking down doing what you're doing every sort of minute or every hour, it's a lot to think about. It was quite time consuming' (P14)*.

Completing the diary when children were in the care of others was a question asked to parents. When children were in the care of grandparents, parents stated that grandparents would keep a log of what the children were doing during the day and/or would speak with the grandparents when collecting their children to ask them what they had done during the day:

'And I sort of tasked my mum with it and err so when I came back at the end of the day, she'd actually written it all on the back of an envelope what she thought of what he'd

done and then just let me do the categories of it. And then... so, we chatted over it and then just sort of made sure that I understood what she had written.' (P15).

In some instances, as nurseries had shared the information with parents, parents felt able to ask nursery if they could complete the diary during the day or to have it in mind so that they could report back to parents at the end of the day. Parents also reported that their nurseries now provide detailed feedback on what their child engages in during the day, which can be through various types of Apps. This may include how long they sat down for lunch, what they ate, how much time they spent outdoors, as well as whether the child has been sat reading stories, if they have had a nap and the times of the nap.

'And we get a diary, it's like an app on the thing and a breakdown of the day and such, so between speaking at the gates and going through the diary I can think about what... everything's there, times, how long she'd sit for her dinner and things like that... and how long.' (P16).

Further to this, parents reported the use of social media, where nurseries will upload pictures and information about what the children had done during the day.

'Err no, it was quite easy to be fair. I mean like the only thing is like I don't really know like 100 per cent what she gets up to at school. So, like we get pictures sent like every day and like we've got like a little Facebook group so like I was managing to go on Facebook and sort of like semi guess, like through the pictures like what they'd been doing...' (P20).

6.5.5.4 Scoring of MoveMEY

Although it was anticipated that children would only be engaging in one behaviour at once, speaking with parents it was apparent that children of this age may engage in multiple

behaviours simultaneously and so there may have been some replication of time in the responses. For example, eating whilst watching TV, electronic devices that read stories to children rather than parents (children engaging in other activities alongside this such as doing a jigsaw), reading whilst sat in the car, walking around whilst playing on a tablet. This may lead to replication of reporting, and the same time period being accounted for twice.

‘Erm... eating... because very often... on a morning he’d have his breakfast watching cartoons. Erm... so, that was some... there was... it was difficult not to replicate, to repeat it.’

There were also some instances where because of newer technologies there was overlap between behaviours, for example, children now watch TV on an electronic tablet rather than these being two separate activities, but parents found the tool self explanatory to just report in the section they felt most relevant, as exemplified below:

‘...Because kids nowadays don’t... I mean they do watch TV but not... it’s more iPad. You know, like they’re constantly flicking through YouTube and things. Whereas they’re not kind of... really into sitting and watching kind of loads of different kinds of programmes on the TV, it’s more films, you know but... Whereas they prefer just to watch people playing games on... on the iPad. I mean that was the only thing that I got a bit... and I thought, ‘Oh, I’m going to have to change it because I think I’ve covered it in there, but it should really be in the other section...’ (P22).

On scoring the diaries, in some instances, the overall time would exceed the 24 hours of the day. There are several reasons why this may be: 1) This may be in part due to the sleep questions as part of the tested tool being broad asking for ‘usual’ sleep rather than reported daily like the rest of the tool. However, the main amendment to the tool includes that sleep is reported daily. 2) It is likely from the scoring that parents are overestimating

the amount of time their children are engaging in PA, and in some instances SB too, which is a known challenge of proxy reported tools. Testing the diary alongside an accelerometer may help to detect the amount of overestimation of certain activities, and to be able to determine whether the diary has the ability to detect whether children are meeting the guidelines or not. 3) There may be cross-over of activity, as outlined earlier in this section, meaning that the same period of time is recorded twice due to children engaging in two behaviours simultaneously, for example, walking whilst playing on a tablet. The diaries took me on average around 10 minutes to score manually.

6.5.5.5 Feedback on participation

Parents consistently reported that they found a positive of participation in the study to be that the diary provided a reflective tool for them, that they found it *‘interesting’*, *‘rewarding’* and *‘helpful’* to be able to see how their child is spending their time. Parents felt that looking at the *‘bigger picture’* of all children’s movement was beneficial and liked this approach, as exemplified in the extract below.

‘I think it... yeah, it is really important to take everything into account because it’s just... if you’ve just got kind of child screen use and you’ve got raw data for that but then you’ve got... you don’t align that with how they sleep or you don’t align that with... they might actually be really active and sort of work hard, play hard, or they might be kind of worst case scenario they’re on their screens all the time and they never move. Erm... and it is, without looking at that bigger picture you’re... you’ll not be able to come up with useful answers on that. So...’ (P15).

A small proportion of parents stated that at first glance completing the diary felt like *‘homework’* and *‘a bit daunting’*. These parents stated that they liked the colour-coded

aspect, and once they had actually gone through the diary, they found it to be easy and straightforward. This is exemplified in the extract below. Further to this, during tool development (**Step 2**), parents had suggested that this might be the case, but that having a more detailed tool would be more meaningful and easier to complete in the long-run. Additionally, of the 14 parents who received the diary, 12 of these completed the full study protocol.

'No, nothing, no. I think when I first got it erm it was a bit daunting... I thought it was a bit... there's quite a lot here but erm... when I looked at it, it was good, even like having the colour coordinated [Unclear 16:14] sounds a bit... broke down every section and once you fill it in... it's actually quite straightforward.' (P16).

6.5.6 Final version of MoveMEY following the full development process and pilot

The final version of the MoveMEY tool following the rigorous tool development process outlined in this chapter can be found in appendix 6.16. Here I will summarise some of the key elements of MoveMEY for transparency on the format, content, and scoring of the final version of MoveMEY.

6.5.6.1 Format of MoveMEY

Following the rigorous tool development process outlined in this chapter, the final tool was a 7-day paper based activity diary. The tool was separated into four sections, firstly to capture 'general information' on health status of children, followed by sections to capture the three main behaviours: PA, SB, and sleep.

6.5.6.2 Content of MoveMEY

The first section of the tool included capturing ‘general information’ on health status of children, including disabilities, sleep disorders and illness at the time of measurement. This was a new section, which was reported by both parents/carers and topic expert researchers as important information to capture when assessing movement behaviours.

Throughout the tool development process, it was important to ensure that the content of the tool mapped on to the WHO movement behaviour guidelines (WHO, 2019). The rationale for the questions and why these were included in the provisional tool are described in section 6.5.2.2 (p. 452 onwards). Figure 6.6 outlines how the final tool questions mapped on to the guidelines.

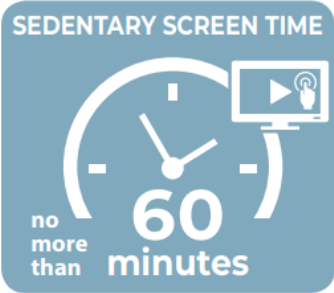
For the PA section, there were three questions, broken down by type of activity: active play (outdoor and indoor separately) and active travel. All questions on PA also included sub-questions relating to whether the activity made the child *“huff and puff”, breathe harder, breathless, hot and sweaty, or stop the activity to ask for a drink*’ and the hours and minutes that this encompassed; the purpose of which was to determine the amount of MVPA the children engaged in.

For sedentary time, four questions were proposed, which were separated to encapsulate the different elements of the guidelines, including: sedentary screen time, screen time before bed (situated in this section to follow on from the screen time question), sedentary whilst engaging in additional pursuits (e.g. reading, storytelling), and seated whilst travelling (e.g. in a car seat / pram).

For the sleep guidelines, three questions were proposed, regarding: bed time, sleep time, morning wake time and out of bed time, night time awakenings and nap times to determine sleep duration and quality.

Figure 6.6. Final tool following rigorous development and pilot mapped to guidelines

Guidelines		Final questions			Scoring
Physical Activity 	Spend at least 180 minutes in a variety of types of physical activities at any intensity.	Please state how many hours and minutes your child spends actively playing outdoors in each of the following (activities may include: running around, jumping on a trampoline, climbing, skipping, throw/catch).	Please state how many hours and minutes your child spends actively travelling, which could include travelling for leisure (e.g. to/from school, the shops, the park) each day.	Please state how many hours and minutes your child spends actively playing indoors (activities may include: dancing, running around, rough and tumble play, sit and ride push toys).	Daily PA for each question calculated by sum of all hours and minutes of activity recorded each day. The 'daily PA' time for all 3 individual questions will be summed to provide a total daily PA.
	At least 60 minutes is moderate- to vigorous intensity PA	Did any of these activities make your child 'huff and puff', breathe harder, breathless, hot and sweaty, or stop the activity to ask for a drink. (Please circle)	Did any of these activities make your child 'huff and puff', breathe harder, breathless, hot and sweaty, or stop the activity to ask for a drink. (Please circle)	Did any of these activities make your child 'huff and puff', breathe harder, breathless, hot and sweaty, or stop the activity to ask for a drink. E.g. at	Daily moderate to vigorous intensity PA calculated by the sum of the 3 questions.

				soft play (Please circle)	
	At least 60 minutes is moderate- to vigorous intensity PA.	If yes, please state how many hours/minutes of this activity made your child 'huff and puff', breathe harder, breathless, hot and sweaty, or stop the activity to ask for a drink.	If yes, please state how many hours/minutes of this activity made your child 'huff and puff', breathe harder, breathless, hot and sweaty, or stop the activity to ask for a drink.	If yes, please state how many hours/minutes of this activity made your child 'huff and puff', breathe harder, breathless, hot and sweaty, or stop the activity to ask for a drink.	
Sedentary Behaviour 	Not be restrained for more than 1 hour at a time (e.g. prams/strollers) or sit for extended periods of time	Please state how long your child spends seated whilst travelling. (Question includes rows for different types of travel e.g. in the pushchair, being carried, sat in a car).			Sum of duration of hours and minutes recorded for child being: 1) in the pushchair; 2) sat in a car; 3) other (please specify)- if this includes an activity where child would usually be restrained (please specify what activity

					this is), each day.
	Sedentary screen time should be no more than 1 hour; less is better	Please state how long your child spends in screen based activities whilst in a sitting, reclining or lying position.			Sum of duration of hours and minutes spent in each activity, each day.
	When sedentary, engaging in reading and storytelling with a caregiver is encouraged	Please state how long your child spends playing and in other activities whilst sitting, reclining or lying, including quiet or carpet time. (Question includes rows for different activities e.g. playing with toys, reading, storytelling, phonics).			Sum of duration of hours and minutes spent in each activity, each day.
Sleep 	Have 10–13 hours of good quality sleep, which may include a nap	Please write the time your child goes to bed, the time they fall asleep, the time they wake up, and the time they get out of bed each day.	Please state how many times and for how long each time, that your child wakes up during their night time sleep.	Please state how many times and for how long each time that your child naps during the day.	Duration - Total hours/mins from sleep time to wake time minus total mins awake during night + total nap duration. 'Good quality' -

					Calculated by total sleep time (total hours/mins between sleep time and wake time minus night-time waking + nap time) divided by total time in bed (total hours/minutes between bedtime and out of bedtime) x 100 'Good sleep quality' will be based on sleep efficiency – ratio of total sleep time to time in bed, $\geq 85\%$ = good sleep quality.
	Regular sleep and wake-up times	Please write the time your child goes to bed, the time they fall asleep, the time they			Consistent (classified as within 30 minutes) of

		wake up, and the time they get out of bed each day.			sleep and morning wake time for ≥ 5 days.

6.5.6.3 *Scoring of MoveMEY*

In line with the revised format of MoveMEY, the accompanying scoring system was also modified. As reporting of the finalised tool was daily, this made the scoring more straightforward as the scores inputted by parents could be summed for each individual day to provide total hours and minutes of engagement in each behaviour. Figure 6.6 shows how each question maps across to the scoring system. For the full scoring sheet please refer to appendix 6.17. For total daily PA, daily PA for each question was calculated by sum of all hours and minutes of activity recorded each day. The daily PA for the individual questions asking about PA could be summed to provide a total daily PA. From this, it was possible to determine how many days children are physically active and the duration of PA on those days to establish if they meet the recommendation. For daily MVPA, the questions asking about this activity were summed to determine how many days children engage in MVPA and their average MVPA on those days to establish if they meet the recommendation.

For sedentary time, total sedentary time could be calculated by summarising the hours and minutes spent in each of the activities each day. The total daily sedentary time from each of the questions would then be summed to provide a total daily sedentary time score. To determine individual guidelines such as screen time, the sum of duration of hours and minutes spent in sedentary screen time would be calculated for each day.

For the sleep section, total sleep duration can be determined using total time from sleep time to wake time minus total minutes awake during the night plus total nap duration. 'Good quality' sleep can be determined based on sleep efficiency, which included $\geq 85\%$ total sleep time from total time in bed to be classified as good sleep quality, based on NSF recommendations (Ohayon et al., 2017). As described in section 6.5.2.3 (p. 458) 'consistent'

wake and bed times was determined to be if these were reported as consistent for ≥ 5 days, based on existing questionnaires using this definition (Owens, Spirito and McGuinn, 2000). Previously, researchers have used the following question to detect consistency of bedtime routine previously '*Does your child have a usual bedtime routine, that is, on most nights, do the same activities occur?*' (Mindell et al., 2009). Therefore, although there is currently no consensus in the literature on what is classified as 'consistent', this definition also made logical sense with being more consistent than not. In terms of what is classified as being a 'consistent' time, I opted for the definition of within 30 minutes. There is minimal evidence that provides a classification on 'consistent', however; 30 minutes has previously been used within the literature (Meredith-Jones et al., 2019).

6.6 Discussion

6.6.1 Principal findings

This study outlines the development and content validity of Movement Measurement in the Early Years (MoveMEY), a newly developed 7 day diary that measures the behaviours of physical activity (frequency, duration, intensity, and type), sedentary behaviour (frequency, duration, and type) and sleep (duration and quality) of pre-school aged children (aged 3 or 4 years old). MoveMEY was developed with parents and carers of pre-school aged children, as well as topic relevant researchers. The primary purpose of the devised measurement tool is to determine whether children are meeting the recent (2019) WHO global guidelines for movement behaviours in this age group (WHO, 2019). The intended application of MoveMEY is for research purposes, to discriminate between those meeting versus not meeting the guidelines, assessing prevalence, and difference in movement behaviours between groups. This tool was developed as there are no existing proxy report based tools that would assess adherence to the movement behaviour

guidelines of pre-school aged children. Additionally, the target population of focus was parents and carers from lower SES backgrounds based on index of multiple deprivation. This was due to underrepresentation of these groups in measurement literature (as shown in *Chapter 3* and *Chapter 4* of this thesis), as well as a focus of interventions and initiatives aimed at this group due to inequalities in health and health outcomes across SES (see section 2.4, p.51 and 2.11.4, p.134 onwards).

This is the first available evidence on this particular measurement tool, outlining the development and content validity assessment. The tool has been developed as a 7-day paper based diary that can be administered for use by parents and carers of pre-school aged children without the need for researcher assistance. The tool is separated by home and nursery time where appropriate, and parents can report separately for every day of the week. The tool consists of four sections, with varying questions: 1) General questions: 2) Physical activity, 3) Sedentary behaviour and 4) Sleep. In general questions, parents can report information on whether their child has any disabilities, sleep problems or illnesses. In section two parents can record how much time (in hours and minutes) their child spends in different outdoor and indoor physical activities, and active travel. There is also space to record how many hours and minutes of this activity is of higher intensity. In section three parents can report how much time (in hours and minutes) their children spends engaging in different sedentary screen based activities, seated whilst travelling, and engaging in different pursuits (e.g. playing with toys, crafts) whilst seated. There is also one question where they can state the last time their child uses a screen before bed. In section four, parents report the times that their child goes to bed, to sleep, wakes up, and gets out of bed. Alongside how many times and how long each time their child wakes during the night and naps during the day to determine sleep duration, quality, and consistency of their

child's sleep. A scoring sheet accompanies the tool to be able to determine duration and frequency of each of the activities, which can subsequently be used to determine whether children meet the thresholds for the guidelines.

Parents, carers and topic expert researchers led the design, format, and content of the tool. Some main findings included a change from a brief tool to a much more detailed 7-day diary based format. This was suggested for various reasons including: daily reporting being easier to recall (and thus providing more accurate and meaningful responses), providing the ability to distinguish between weekday and weekend activity, and ensuring that there was not an over/underestimation of certain behaviours such as naps and night wakings that do not always occur consistently throughout the week. Throughout the stages of development, parents consistently stated the variability between days in terms of the activity of their young children which highlighted the importance of a tool that can capture this variation and intricacies between the days to provide a more accurate representation of pre-school children's movement behaviour. Further to this, the tool was separated so there was space to report PA and SB separately for both home and nursery, as this was recommended to make it easier to detect the nursery based activity that may otherwise not be recorded.

During the development of the tool, parents, carers, and topic expert researchers found that the provisional tool included relevant activities for children of this age, and topic expert researchers felt the items of the tool appropriately targeted the guidelines. Multiple suggestions were made to make the tool more comprehensive and easier to complete. For example, including location of outdoor active play/ PA rather than just activities; separating the different modes of seated travel; adding activities such as bath time, time spent eating;

space to add own activities; space to report why there may be differences in sleep; and changes and clarification on terminology. Some substantial additions included adding a space for 'additional comments' at the end of each section, a question on screen use before bed, and including a section on 'general information' to demonstrate whether the child has a disability, any sleep problems, or is suffering from an illness at the time of measurement.

The interviews with parents following the pilot of the tool indicated that MoveMEY contains all activities that children of this age engage in and is easy to understand when aiming to detect movement behaviours of pre-school children. Some amendments were made to the tool following the pilot, in line with parental feedback, including adding more examples to questions and crossing out/removing some irrelevant rows (e.g. crossing out 'nursery' boxes when the question was specific to play at 'home'). The most substantial amendment was the transition from questions asking about 'usual' bed, sleep, wake, and out of bed time, to daily reporting, as per the rest of the diary. This resulted in removal of a question on consistency of bed/wake time, as this was no longer necessary. The revised and finalised version of the tool was then checked by a public engagement group of parents, which confirmed that the final tool was appropriate and did not raise any further concerns.

MoveMEY is a novel tool, and has some distinct differences compared to existing parental reported tools examining PA, SB, and sleep. For example, to my knowledge, it is the first co-developed tool undergoing measurement property evaluation that would have the ability to determine the movement behaviour guidelines. The reviews presented in *Chapters 3 and 4* of this thesis demonstrated that there are currently no available tools, which have undergone measurement property evaluation that would be suitable for determining the movement behaviour guidelines of pre-school children. Existing parental reported tools

examining movement behaviours of young children (that include samples of children aged <3 years old), do not have the ability to determine the prevalence of meeting the guidelines due to questions asked and response options not aligning with the movement behaviour guidelines (Chia, Tay and Chua, 2019; Hinkley et al., 2012b; Townsend et al., 2021). For example, the '*Healthy Kids*' tool assesses outdoor PA only, screen time relating to TV and computer screens only (whilst, for example, in the current research parents reported that screen time primarily comes from other devices now such as tablets), and night time sleep only (Townsend et al., 2021). The '*SMALLQ™*', focuses primarily on digital media use with details on children's digital media habits outside of pre-school (on a typical weekday and weekend day), including frequency, duration and purpose of digital media use. This questionnaire also asks questions on indoor and outdoor play, but this includes activities such as non-screen reading and drawing, which may take place whilst seated and this distinction between whether the activities are active or seated cannot be determined from the questionnaire. These questions are asked for 'outside' of pre-school only. Whilst it is described that this tool asks for sleep duration and quality, the questions are not reported and so it is unknown how these dimensions are assessed (Chia, Tay and Chua, 2019). A further tool, the 'HAPPY' survey is a tool detailing 230 items examining correlates of pre-school children's PA; thereby having an entirely different purpose to MoveMEY and as such is not applicable to determine guideline adherence (Hinkley et al., 2012b).

Although at first intended to be a brief questionnaire, through the co-development process with parents and carers the resulting tool has formed a seven day diary based format, separated by the distinct constructs (PA, SB, and sleep). The granularity of the diary format allows for a more detailed overview of child's movement behaviours including frequency, duration, and type of activities. Whilst an advantage of this tool is the granularity

in which it can detect movement behaviours, which may be beneficial in providing more accurate and detailed results than previous parental reported tools, it is plausible that this may be quite intensive for large-scale measurement. However, there is evidence of success using the 'time use survey' (>4500 participants in the published studies after data cleaning), whereby participants self-report activity they engage in every 10 minutes for two 24 hour periods (Adams, 2010; Clifford Astbury et al., 2020). Given that the present tool was developed by parents and carers, with the majority reporting that they did not find completing the tool time consuming, there can be confidence that this is the type of tool preferred by parents and carers of pre-schoolers. However, it is necessary for future research examining the tool to be tested with larger samples to determine the applicability of the tool for large scale measurement.

6.6.2 Strengths and limitations, including limitations of the methodology used

A major strength of this study is the involvement of parents, carers, and topic expert researchers in both the development and content validity assessment of the new movement measurement tool. As parents and carers are the experts of their children's behaviour and are best suited to state which activities their children engage in over a 24 hour period, their involvement in tool development provided invaluable insights in ensuring that the tool was comprehensive and relevant. Another strength of this study is the assessment of content validity following parents completing the tool. This confirmed that the tool was comprehensive and relevant to pre-school children's 24 hour movement behaviours, that no relevant items were missing, as well as approving that the tool was understandable and easy to use. In addition, the study followed the COSMIN framework for the development and content validity assessment of measurement tools, as previously recommended by

researchers to ensure that a high quality development of tool study is conducted (Degroote et al., 2020). Although this tool was developed through a rigorous development process, including the first important step in tool development of ensuring content validity, there are some sources of error and bias with parental reported tools that may be unavoidable, as outlined in *Chapter 2* (section 2.10, p122 onwards) of this thesis. This includes that parents are not always with their children, may have to estimate time due to the sporadic and intermittent nature of children's behaviours and potential for reactivity bias from parents (e.g. taking children to the park more as a result of the behaviour being monitored/using screens less). Whilst these limitations must be acknowledged, there is a necessity for tools of this kind, and the detailed information the tool can offer in terms of the type and duration of activity across different days of the week is valuable.

As outlined in *Chapter 5*, a strength of this study was that the recruitment strategy deliberately targeted groups whose views are often underrepresented in measurement literature (as identified in *Chapters 3 and 4*). Based on the index of multiple deprivation (Ministry of Housing, Communities and Local Government, 2019), the samples included in both the development and content validity assessment of the tool largely consisted of participants living in the most deprived regions in the UK (see table 5.2, p. 396 and table 6.1, p.448). However, participants in the content validity study were very highly educated (92% educated to Bachelor degree level or higher). This highlights the importance of collecting multiple indicators to enable a more accurate representation of the demographic of the sample. To ensure readability of the items on the tool I also inputted the questions into the Flesch-Kincaid readability calculator, a tool used to show how readable text is which provides a score on the approximate educational level a person would need to be able to read a particular text easily (Flesch Kincaid Calculator, 2022). This determined that items of

MoveMEY assessing the movement behaviours were mainly suitable for those in '5th grade' (aged 10-11 years) and were determined to be 'easy to read'. However, some items, such as: *'Please state how long your child spends in screen based activities whilst in a sitting, reclining or lying position'* were scored as being suitable for those in 8th grade (aged 13—14 years) and were determined to be 'Plain English' level. Although it is unknown if the tool development would have differed if parents from higher SES groups had been the target focus, it was important to focus recruitment on lower SES groups for several reasons, including 1) The evident inequality in representation in measurement literature, with limited evidence for individuals from lower SES backgrounds (as demonstrated in *Chapter 3* and *4*) and 2) Due to persistent inequalities in health and health outcomes, interventions and initiatives aimed at improving movement (and related e.g. eating) behaviours of young children are often targeted at lower SES groups (see section 2.11.4, p.134) and so it is particularly important that tools are applicable for use within these groups.

The study has some limitations with the sample that must be recognised. Despite trying to achieve a diverse sample by recruiting settings in one of the most ethnically diverse cities in the UK (Office for National Statistics, 2011), the sample in the present study was not ethnically diverse, as such, generalisability of the findings across different ethnic groups is unknown. In addition to the plausible reasons that I did not achieve an ethnically diverse sample in this work outlined in *Chapter 5*, the recruitment and data collection for the current study took place during the Covid 19 pandemic and was all conducted remotely. As such, additional challenges were present including the inability to attend the settings and to meet parents and carers of children and discuss the study, which may have aided uptake and the ability to recruit more diverse samples. Further, the Covid-19 pandemic disproportionately impacted individuals from minority ethnic groups in the UK (Platt and

Warwick, 2020). It is plausible that concerns of contracting the virus may have led to more families staying at home to reduce exposure to infection and in turn disengaging from research that they may otherwise have taken part in. Additionally, a limitation of the study is that MoveMEY has been developed in English only. As such, only fluent English speaking parents, carers, and topic expert researchers were involved in the development and content validity assessment. The applicability of the tool to children in other countries is unknown and further research would be needed to translate the tool and conduct evaluations on the translated versions. Additionally, in future research on tool development, including further evaluation of MoveMEY, it would be important to collect data on whether children live in more than one household or in a single parent/carer household. This would be beneficial to determine if this impacts on the feasibility of the tool.

The strengths and limitations of using focus group as a research methodology in this work are discussed in length in *Chapter 5* (page 418). In addition, two further methodologies were used in this chapter (*Chapter 6*) including qualitative surveys and interviews.

Qualitative surveys using open-ended questions and response options were used to explore the opinions of topic relevant researchers, including obtaining their views to aid in the development process of the MoveMEY measurement tool. Whilst this was beneficial to obtain richer information and more detailed responses than quantitative research (e.g. asking for ratings of a range of questions), a limitation of this method is that the responses may have been more in-depth if using a method such as interviews, where researchers could speak freely on the topic. One-to-one interviews were selected as the method of choice with parents following completion of the tool due to the aim being to hear parents' experiences of completing the MoveMEY diary, and interviews being suggested as an appropriate method when aiming to understand participants' experiences on a particular

topic (Willig, 2013). A limitation of using this methodology is the presence of social desirability bias when conducting interviews, that can create complexities when interpreting the findings. During one-to-one interviews, the participant may want to provide the researcher with the response that they think is expected or that they think will be most helpful to the researcher. In the present study, I aimed to overcome this by being clear with the participant that there were no right or wrong answers and that any responses were incredibly useful to the research.

Further, as the research took place during Covid-19 it was not possible to conduct the interviews in person, and as such all interviews were conducted via the phone (although there was also a clear option of the interview via video call). This mode of interview could have additional limitations including not being able to observe body language, perhaps less natural conversation than what may happen face to face, and parents could not use the diary as a visual aid (e.g. parents could not point to examples to help support their responses or if wanting to raise additional points).

6.6.3 Implications of findings and future research

This study highlighted the importance and value of rigorous tool development with the target population, in ensuring that the tool is relevant, comprehensive, understandable, and in a suitable format for those who would be involved in future research using the tool. A key implication of this work is the methodological advancement that it provides. To date, no measurement tools for assessing movement behaviours of pre-school children have been developed and assessed for content validity with both parents of pre-schoolers and topic relevant researchers. This work therefore serves as a model for future research, including strengthening the understanding of pre-existing measurement tools, and to help inform the

procedures for future development of measurement tools used to assess movement behaviours (both in this age group and in other population samples).

A further implication of this work is that the newly developed tool, MoveMEY, which demonstrates good content validity and feasibility (in a small sample) shows promise for assessment of the international WHO movement behaviour guidelines of pre-school aged children (WHO, 2019). To my knowledge, this is the first piece of work that has aimed to develop a single tool that could be used to assess compliance with the guidelines. The implications of this are that, in the future following further evaluation (in line with the future recommendations outlined below), MoveMEY can be used to assess guideline adherence, including having the ability to detect 'type' of activity such as screen time.

There are some clear areas for future research specific to the further evaluation of MoveMEY to conclude on the quality and utility of the tool, including:

1. Further evaluation of MoveMEY in ethnically diverse samples, to ensure relevance, comprehensiveness, and understandability across different ethnic groups.
2. Examine the criterion-related validity of MoveMEY, by testing the tool alongside a range of established reference methods, such as direct observation, accelerometer and sleep diaries to determine if the tool provides similar outcomes to valid pre-existing methods.
3. Examine the test-retest reliability of MoveMEY, alongside a comparison tool to determine if the tool is stable at detecting the behaviours, as well as to determine whether any differences reported are as a result of changes attributable to the tool versus changes in behaviour

4. Examine MoveMEY with large samples of children to determine the usability of the tool at scale, particularly with sub-groups not involved in the development of the tool.
5. Once evaluation of MoveMEY is completed in the current format, translation of the tool into different languages would be necessary to ensure cross-cultural translation.

6.7 Conclusion

This study outlines the rigorous development and content validity of MoveMEY; a measurement tool developed with parents and carers of pre-school aged children, and involvement of topic expert researchers in the field of movement behaviour of pre-school aged children. MoveMEY is the first measurement tool developed to examine PA, SB, and sleep guidelines of pre-school aged children aged 3-4 years. Further research to examine the criterion-related validity and test-retest reliability of MoveMEY is warranted to conclude the utility of the measurement tool.

7 [Chapter Seven: Final discussion and conclusions](#)

7.1 [Introduction](#)

The studies outlined in *Chapters 3 to 6* of this thesis each flowed on from one another as part of the overall ‘story’ of this thesis (as can be seen in Figure 1.2, p.29), however, each chapter could also be read as a standalone piece of research. As such, each chapter included individual discussion sections, outlining main findings, strengths and limitations of each piece of work, implications of findings, areas for future research and overall conclusions. This final chapter will draw on the discussions of the chapters to summarise and reiterate some of the key findings from each chapter of the thesis, discuss some possible limitations of the research presented in this thesis, and present implications and areas for future research.

7.2 [Key Findings from Thesis](#)

7.2.1 [Chapter 3](#)

The aim of the systematic review presented in *Chapter 3* was to identify the measurement properties (validity, reliability, and feasibility) of existing measurement tools used to assess PA and SB of pre-school aged children. A large body of evidence, 69 articles describing 75 studies, was identified in this review. The key findings of the review included that a range of accelerometers, including the Actigraph (in particular GT3X versions), Actical and ActivPAL, demonstrate most reasonable validity for assessing common movement related outcomes (e.g. SB, MVPA, total PA) for free living activity of pre-school aged children. The Fitbit (Flex and Zip) also showed very promising results; however, these were based on a very limited sample of studies. The evidence base on proxy reported measurement tools was less extensive (n=13). However, some proxy reported measures

including the Pre-PAQ, *Netherland's PA questionnaire* and *nursery teachers report* can provide some valid data particularly for distinguishing between different levels of activity, alongside useful contextual information not captured by device-based measurement tools. Two studies provided very minimal information (ranging from one sentence to one paragraph) on the development of the tools, and did not provide details on the content validity of the studies. A small proportion of studies demonstrated representation of individuals from lower SES groups, and a lack of ethnic diversity was apparent in the participant samples of the included studies. Of the extensive findings of this review, the subsequent chapters focused on the findings of further qualitative work needed to help determine the acceptability and feasibility of measurement tools used to assess the PA and SB of pre-school aged children for free living activity measurement (*Chapter 5*). As well as building on the need for more in-depth evaluation work of proxy-reported measurement tools, through aiming to develop a tool for measuring PA and SB, as well as sleep, that would be applicable for capturing movement behaviour guidelines at scale, ensuring a rigorous development and content validity assessment with parents and carers of pre-school children (*Chapter 6*). The subsequent empirical work (*Chapters 5 and 6*) also deliberately drew on families from more deprived areas due to the limited measurement research with these communities as identified through this review.

7.2.2 Chapter 4

The aim of the rapid review presented in *Chapter 4* was to identify the measurement properties (validity, reliability, and feasibility) of existing measurement tools used to assess sleep of pre-school aged children. This review identified a very small evidence base of only 16 studies examining measurement properties of tools used to examine sleep, with the

majority of studies demonstrating high risk of bias. The majority of the studies examined proxy-reported tools, questionnaires and diaries, with only three studies examining the use of accelerometers. One type of sleep diary was shown to be valid for measuring sleep duration; this was assessed in one study only, but with a low risk of bias. Three questionnaires showed convergence with accelerometers, with some evidence of reliability, for measuring sleep duration of pre-school aged children. However, the tools were only assessed in one study each and all demonstrated high risk of bias. One tool, the Fitbit Ultra accelerometer, was not significantly different from polysomnography for determining night wakings; however, this was examined in one study, with a small sample size, and was determined as high risk of bias due to the method of analysis used. No identified measurement tools were appropriate for determining sleep quality. Reliability information was minimal, and feasibility was limited to the assessment of missing data. None of the studies explained the development of the proxy reported tools, and no content validity studies were identified. The studies that reported SES of the participants demonstrated that families of the children were mainly middle-upper class. Of the multiple findings of this review, the subsequent chapters focused on the findings of further qualitative work needed to help determine the feasibility of measurement tools used to assess the sleep of pre-school aged children (*Chapter 5*). As well as aiming to develop a higher quality tool integrating sleep into a tool with PA and SB, through rigorous development and content validity assessment with parents and carers of pre-school children (*Chapter 6*). The subsequent empirical work (*Chapters 5 and 6*) also deliberately drew on families from more deprived areas due to the limited measurement research with these communities as identified through this review and the review presented in *Chapter 3*.

7.2.3 Chapter 5

The aim of the research presented in *Chapter 5* was to examine the acceptability and feasibility of measurement tools used to examine movement behaviours of pre-school children, which included a sample of parents and carers from deprived communities in the UK due to underrepresentation of these communities in measurement literature, as identified in the reviews presented in *Chapter 3* and *Chapter 4*. This included demonstrations of the tools (three types of accelerometer: Actigraph GT3X+, activPAL4 micro, Actical; a sleep diary, and a questionnaire), followed by focus group discussions with parents and carers of pre-schoolers. Key findings from this study included: 1) Importance of being able to provide contextual information when reporting movement behaviours of young children (e.g. child illness, capturing different routines); 2) Practical issues associated with devices (e.g. aversion to devices being attached directly to the skin of their child; concern of larger devices during sleep time); 3) Encouraging children to wear a device (e.g. making devices attractive to children- ‘superpowers’, colourful bands and stickers); and 4) preferences on presentation of diaries and questionnaires (e.g. age appropriate activities, preference for real-time recording over recall). The issues raised regarding the feasibility and acceptability of the tools have the potential to impact on recruitment, retention and compliance in research studies. The study highlighted the importance of engaging with the target population prior to conducting a study to ensure that the study materials and equipment are acceptable and feasible, to improve successful delivery of a study. Through the work presented in this chapter, it was possible to develop a series of recommendations on how these findings can contribute and inform the design, development, and implementation of future research projects measuring the movement behaviours of pre-

school aged children. Some findings of this work, particularly relating to presentation of diaries and questionnaires fed into the development of the new tool presented in *Chapter 6*.

7.2.4 Chapter 6

The aims of the research presented in *Chapter 6* were to develop a measurement tool used to assess the movement behaviours of pre-school aged children, together with parents and carers of pre-school aged children, and topic experts. As well as to subsequently pilot the devised measurement tool and conduct interviews with parents and carers of pre-school aged children following the pilot, to determine the content validity of the devised measurement tool for assessing the movement behaviours of pre-school aged children. Findings of the study demonstrated that a seven-day diary, separated by home and nursery activity, that was able to detect weekday and weekend day differences in activity was the preferred format. The behaviours that pre-school children engage in were included in the tool, as outlined by parents, carers and topic expert researchers. The importance of also including some situational context such as illness, disability, and sleep disturbances were reported. The findings of the pilot study revealed that the tool contained relevant items, was comprehensive in that it contained age-appropriate activities and covered all behaviours that the pre-school children engage in, and the tool and instructions were comprehensible. Parents liked the format of the tool using daily recall of activities, felt the sequence of the questions presented in the diary and response options were all suitable. Although the tool was well received with no major concerns raised, based on the parental feedback, some amendments were made to wording and layout of some of the questions and response options, to ensure that these were in line with what parents felt would further improve the tool. The findings of the study highlighted the importance of including the

target population and a group of relevant experts in measurement tool development, to ensure that the tool is appropriate, contains relevant items, is comprehensive and comprehensible for use. As will be discussed later in this chapter (section 7.5, p.528), future research to finalise evaluation of the tool in comparison with reference methods is warranted prior to further use.

7.3 Limitations

The limitations associated with proxy-reported measurement tools have been discussed throughout this thesis (see section 2.7.10 p.92, section 2.8.8 p.112, and section 2.10 p.122). The development of the new measurement tool followed the recommended COSMIN framework to ensure sound development and the assessment of content validity of the newly developed tool, which is stated to be the first most important step that is rarely undertaken. The lack of tools that have undergone thorough development and content validity assessment was a finding of the reviews presented in *Chapters 3 and 4* of this research, as well as stated previously by others (e.g. Degroote et al., 2020; Hidding et al., 2017). Despite this, MoveMEY and other proxy reported measurement tools are susceptible to certain biases and errors commonly associated with tools of this kind. These may include: issues associated with recalling behaviour of a child whom a parent is not always with, busy lifestyles meaning that exact timings cannot always be taken, difficulties with determining intensity of activity, and potential for social desirability bias. However, this tool is designed to provide an estimate of behaviour, to discriminate between children meeting the movement behaviour guidelines versus those who do not. This type of measurement tool may not be suitable for some study protocols where, for example, measuring exact energy expenditure of a child is required; but this type of tool may be suitable and ideal for studies

aiming to view an overall picture of a child's movement behaviours, including type and context of activity, and to be used at scale.

A further limitation when aiming to measure movement is 'other' behaviours that children engage in during the day, such as standing, that may not be captured by the majority of measurement tools, including MoveMEY. Early evidence suggests that the ActivPAL accelerometer shows promise for determining posture allocation of pre-school children but does not always show this with optimal accuracy (Davies et al., 2012); therefore, there is much more work needed in this area, and it may not be possible to detect these behaviours with the majority of measurement tools.

A main focus of this work was to target lower SES groups and to try achieve ethnically diverse samples, given the underrepresentation of these groups in measurement research (as identified in *Chapters 3 and 4*). The recruitment for the empirical work presented in this thesis (*Chapters 5 and 6*) involved recruiting settings in the top quintile of deprivation (based on index of multiple deprivation) in the UK, as well as specifically trying to achieve a diverse sample by recruitment of settings in one of the most ethnically diverse cities in the UK (Office for National Statistics, 2011). Despite this, the sample in the research presented here was primarily (96%) White British. The claims made from this research are limited to primarily White British population samples at this time, as the findings across different ethnic groups are unknown. The lack of ethnic diversity in the research presented here is a major limitation of the work, as this lack of representation limits generalisability of the findings across different ethnic groups. It is crucial that the next steps of evaluating MoveMEY (including assessing validity in comparison with reference methods, test-retest reliability, and feasibility in larger samples) involve pre-school children and parents from

ethnic minority groups. The ways in which this may be achieved would include several strategies: 1) Trying to access a gatekeeper with whom individuals from different ethnic backgrounds may feel more able to identify with, 2) Public engagement and involvement with individuals of different ethnic origin, to identify if the recruitment materials (e.g. recruitment posters) can be made more inclusive, 3) Aiming for face to face recruitment in settings in ethnically diverse communities. Further to this, although the majority of families included in the empirical research (78%) were classified as living in the most deprived quintiles in the UK, the pilot study involved a group of highly educated people, with 92% of the sample being educated to Bachelor degree level or higher.

For the systematic and rapid reviews conducted as part of this research (*Chapter 3 and Chapter 4*), a main limitation was the risk of bias assessment tools used. There are currently no risk of bias assessment tools available that are applicable for studies of measurement properties on the range of tools used in PA, SB and sleep measurement. Although the COSMIN risk of bias checklist (Mokkink et al., 2018b; Prinsen et al., 2018; Terwee et al., 2018a) is suitable for appraising studies on self/proxy reported measurement tools, it was not devised for appraising risk of bias of studies examining other measurement tools. Risk of bias assessment tools devised specifically for assessing studies evaluating measurement properties of a range of tools would be beneficial to appropriately assess risk within these types of studies.

Further, it was not possible to conduct meta-analyses in either of the reviews (*Chapter 3 and 4*) presented in this thesis due to heterogeneity between the included studies. This was due to the studies examining multiple different measurement tools (and versions of measurement tools e.g. multiple versions of accelerometer were examined in

the review presented in *Chapter 3*, and it was apparent that the different versions of tools could produce different outcomes), difference in comparison measurement tools used, the lack of multiple studies assessing the same measurement tool, and the different outcome measures being assessed in the studies. Rather a narrative synthesis was conducted using a pragmatic approach to combine the results of studies, however, being able to combine the statistics from multiple studies may have increased precision of the results of these reviews.

The empirical work conducted as part of this thesis consisted of qualitative research methodologies (including focus groups, qualitative surveys, and interviews). Whilst this was a strength in terms of being novel for the measurement literature field, helping to achieve richer information and more in-depth responses on the understudied measurement properties of feasibility, acceptability and content validity, there are some limitations of qualitative research that need to be addressed. Firstly, there is a higher likelihood of social desirability bias being present when using qualitative methods as there is direct contact between participant and researcher when the participant is recalling their responses. Additionally, it is likely that self-selection bias was present with parents who have more of an interest in movement behaviour of their pre-schoolers choosing to take part in the research. However, self-selection bias is a limitation of research more broadly. Finally, with analysis of qualitative research the findings of the research are to an extent restricted to the researcher's interpretation of participants' accounts.

Additional limitations specific to the conditions in which the research took place (during the Covid-19 pandemic including multiple lockdowns) are to be acknowledged. This resulted in heightened responsibilities for early years settings which restricted the degree to which settings could engage in research due to increased restrictions (e.g. wanting to limit

unnecessary contact with anyone for whom it was not essential to be in the early years settings), staff shortages including increase in staff absences due to Covid. This also limited the extent to which recruitment and data collection could take place face to face.

7.4 Implications of Findings

It has been stated that to appropriately tackle factors high on the public health agenda in the UK, including obesity prevention, there is the need to understand the full spectrum of movement behaviour of pre-school children (Reilly et al., 2020; WHO, 2017). The rates of childhood obesity in England in children aged 4 and 5 years old have risen substantially in recent years, from 9.9% in 2019/20 to 14.4% in 2020/21 (NHS Digital, 2021). The prevalence of childhood obesity is considerably higher in children aged 4 and 5 years old from the most deprived regions in the UK (20.3%) than those in the least deprived (7.8%), with a clear gradient throughout levels of deprivation (NHS Digital, 2021). Thus, emphasising the importance of examining movement behaviours, and needing appropriate and acceptable tools in order to do this.

The recent Covid-19 pandemic has exacerbated pre-existing inequalities in society (Marmot et al., 2020b). As discussed at length in section 2.12 (p. 134 onwards) of this thesis, covid-19 has had profound impacts on the health and development of young children, due to loss of learning and socialisation, which have increased inequalities (e.g. see Hefferon et al., 2021; Marmot et al., 2020b). Movement behaviour of pre-school children has been impacted by the covid-19 pandemic, including children engaging in more sedentary screen time, less PA and time outdoors and later bed and wake times (e.g. see López-Bueno et al., 2020; Okely et al., 2021a). In the covid recovery phase, there is now an increased need for intervention and approaches for improving health through addressing and establishing

healthy movement behaviours and reducing health inequalities in children from the earliest years (Marmot et al., 2020b). It is important to be able to appropriately monitor and evaluate changes in these behaviours, underpinned by suitable methods of measurement. In line with this, an implication of the research presented in this thesis is the finding that there are currently no existing tools that are valid, reliable, and feasible for assessing the movement behaviours of pre-school aged children and no existing measurement tools that could be used at scale to determine guideline adherence. This is an important implication for researchers, practitioners, and policy makers; it is crucial that tools are rigorously developed for the purpose of assessing national and international guidelines, and subsequently thoroughly evaluated to ensure that they assess the behaviours accurately and are feasible for use by participants.

The reviews presented in this thesis that collated the current available evidence provide resources that can contribute to national policy recommendations for surveillance and monitoring of physical activity and related behaviours of pre-school aged children. Further, the reviews represent key reference points for researchers and a criterion when interpreting and using findings from studies examining these behaviours. The reviews in this thesis highlighted that different tools can produce significantly different outcomes, and to produce high-quality evidence moving forward, measurement issues should be a central consideration.

In line with the shift towards a movement behaviour approach and to develop appropriate interventions and approaches, and monitor changes in effectiveness, it is important to have measurement tools available that can be used at scale and are equally acceptable across groups of different SES. However, with the persistent and evident

inequalities, and subsequent intervention targeted at these groups, it is crucial that there is a specific focus on lower SES groups, with this being an important consideration needed within the measurement field. The research outlined in this thesis has provided a sound basis on the quality of existing measurement tools for assessing movement behaviour of pre-school children, including highlighting the underrepresentation of lower SES and ethnic minority groups in measurement research. A wider implication of this work is that within the movement behaviour measurement field there are no proxy-reported tools that have been evaluated in a representative range of ethnic minority groups, so similarly to low SES communities, representation is currently very limited. This has implications for the tools that are currently being used, ranging from tools used as reference methods in validation studies, and also tools being used in research more generally including intervention studies or studies examining associations between movement behaviours and health and developmental outcomes. The quality (including validity, reliability, and feasibility) of tools across ethnic groups should be a priority for future research in this area, to ensure that the tools are equally as useful and accurate across groups.

The thesis also provides important insights from parents and carers, from deprived communities in the UK, on the acceptability and feasibility of a range of measurement tools for the assessment of all movement behaviours of pre-school children. A primary implication of this work is that it enabled the development of a series of recommendations for the measurement of movement that can be used in the design, development and implementation of future research projects. The findings offer practical issues, alongside useful strategies, that may help with recruitment and retention within research projects; researchers and practitioners can use the findings of the research to determine which tools to use and ways in which they can be used to increase compliance and enjoyment.

Additionally, the findings may be useful to policy makers. To date, feasibility information on measurement tools used to assess movement behaviours has been very limited. This novel piece of work provides the first qualitative insights into a range of tools for assessing the full spectrum of movement behaviour. Policy makers may use these findings when considering which tools may be most feasible for use, particularly for surveillance and monitoring. Finally, these findings are useful for future tool development and may be helpful to manufacturers in ensuring that tools are applicable and appropriate for young children.

This research also offers a novel co-developed tool (MoveMEY) suitable for assessing movement behaviours of pre-school aged children, ready for further validation work. MoveMEY has demonstrated feasibility (in a small sample) and content validity (including containing relevant items, being comprehensive of children's movement behaviours and understandable to parents). The implications of this work will extend across future research examining the movement behaviours of pre-school children, as it provides the first piece of work that has developed a single measurement tool suitable for assessing the international movement behaviour guidelines of pre-school aged children. When assessing interventions and strategies, and monitoring movement behaviours and the associations with health outcomes, measurement tools that are valid, reliable, and feasible for use at scale are paramount. This will help to advance the field of movement behaviour of pre-school children, which will generate more evidence to help with future policy development.

Finally, the tool development and content validity assessment process outlined in this thesis is methodologically influential. To date, no measurement tools for assessing movement behaviours of pre-school children have been developed and assessed for content validity with both parents of pre-schoolers and topic relevant researchers. This work

therefore serves as a model for future research, including strengthening the understanding of pre-existing measurement tools, and to help inform the procedures for future development of measurement tools used to assess movement behaviours (both in this age group and in other population samples).

What is already known on this topic

- Paradigm shift suggesting that movement behaviours (physical activity, sedentary behaviour and sleep) should be viewed collectively rather than in isolation.
- Early years are an important age group to examine and establish healthy movement behaviours, demonstrated through introduction of international movement behaviour guidelines for this age group.
- Examining movement behaviours in pre-school aged children is important for public health concerns, such as obesity prevention, and is a critical time in which inequalities can be addressed.
- Limited evidence examining combinations of movement behaviours with health and development indicators in young children. A suggested reason concerns measurement issues, with a lack of valid, reliable, and feasible methods used to assess the whole spectrum of movement behaviours.

What this study adds

- Identified that there are currently no existing tools that are valid, reliable, and feasible for assessing the movement behaviours of pre-school aged children. No existing measurement tools that could be used at scale to determine guideline adherence.
- Identified underrepresentation of individuals from low socioeconomic status backgrounds and ethnic minorities in existing measurement literature in this age group.
- Provided qualitative insights from parents and carers, from deprived communities in the UK, on the acceptability and feasibility of existing measurement tools (accelerometers, diaries, and questionnaires) for the assessment of all movement behaviours of pre-school children.
- Provides a newly co-developed tool (MoveMEY) in the format of a 7-day activity diary
- Evaluation of MoveMEY demonstrating feasibility (in a small sample), showing promise for assessing movement behaviours of pre-school children at scale and content validity assessment including demonstrating that MoveMEY has relevant items, is comprehensive of children's movement behaviours and understandable to parents.

Figure 7.1. Key findings from thesis: what is already known on this topic and what does this research add.

The wider implications of the work in the context of broader public health concerns and the field of movement behaviour measurement of pre-school children have been

discussed above, however, below I outline key implications for each piece of research presented in this thesis:

The two reviews presented in this thesis (*Chapter 3 and 4*) provide detailed information on the current evidence surrounding the measurement properties (validity, reliability, feasibility) of tools used to assess physical activity, sedentary behaviour and sleep of pre-school children. These reviews provide useful resources for researchers, practitioners and pre-school educators when selecting a measurement tool to examine the PA and SB, and/or sleep of pre-school children. A key long-term implication of this work is that the reviews outline high-quality information on the appropriateness of tools used to measure physical activity, sedentary behaviour, and sleep of pre-school children, and represent key reference points for researchers and a criterion for ensuring confidence in the outcomes of studies examining these behaviours. With this, the reviews emphasised that the choice of tool can impact substantially on measurement outcomes, that careful consideration is needed when interpreting and understanding results of studies based on which measurement tools were used and highlighted the importance of choosing a tool with good measurement properties in research and practice to develop appropriate policies and programmes for behaviour change. Further, the reviews can be used by policy makers for recommendations for the surveillance of physical activity and related behaviours, both in the UK and internationally.

The qualitative results on the acceptability and feasibility of tools (*Chapter 5*) have enabled the development of a series of recommendations for the measurement of movement behaviours of pre-school aged children that may be used in the design, development and implementation of research projects. These recommendations provide an

important contribution to the literature given the lack of information on the acceptability and feasibility of a range of measurement tools in this age group. The long-term implications of this include researchers being able to use the findings to aid their own research process including overcoming some of the practical issues and using some of the strategies outlined to help with recruitment and retention in research studies. Additionally, the research can be used by policy makers and practitioners when considering some of the feasibility aspects of measurement movement in practice. Finally, the findings may be particularly useful to manufacturers in making device-based tools applicable to the pre-school population. These recommendations are presented in an infographic (see Figure 5.3, p.421) and are particularly valuable given that the recommendations are drawn directly from views of parents and carers of young children.

The results of *Chapter 6* highlight the importance of rigorous tool development with the target population of interest, in ensuring that tools are accepted and appropriate for those who will be using them. The results show that a seven day activity diary, with the ability to detect home and nursery activity, and weekday and weekend activity was the preferred way of obtaining movement behaviours of pre-school children through proxy report. The implications of this chapter are also that the newly developed tool, MoveMEY has good content validity, is relevant, comprehensive and comprehensible for assessing the movement behaviours of pre-school children. The work is internationally relevant, presenting the first available tool that aligns with the WHO movement behaviour guidelines (WHO, 2019). MoveMEY is a tool that can help to estimate the movement behaviours that pre-school children engage in and may also be able to discriminate between children who do and do not meet each of the individual guidelines. The longer- term implications of this are that, in the future following further evaluation, MoveMEY can be used to determine

guideline adherence and has the ability to detect ‘type’ of activity such as screen time. This will act as a very useful tool to help inform further guideline and policy development, as well as future interventions. In addition to this, the work serves as a model for future tool development work. To date, no measurement tools for assessing movement behaviours of pre-school children have been developed with both parents and topic relevant researchers. Therefore, this work has methodological implications within the measurement field; acting as a template to strengthen the understanding of pre-existing tools, in terms of if they are measuring what they intend to measure, and for thorough and in-depth development of new tools.

7.5 Future Research

In each chapter of this thesis I have outlined some areas for future research specific to the individual piece of work, with some chapters (*Chapters 5 and 6*) drawing on some gaps identified through the reviews presented in earlier chapters (*Chapters 3 and 4*). However, a number of further additional questions have emerged that, if addressed, will help enhance the field of measuring movement behaviours of pre-school children. This includes an exploration of accelerometers for measuring 24-hour movement, in particular a focus on accelerometers used in PA and SB research for the assessment of sleep, against criterion methods (polysomnography for sleep research). Further to this, research is needed to examine the impact of accelerometer decisions on the outcomes for 24-hour movement measurement, in particular placement of devices and algorithm development specific to pre-school aged children. This will help with the use of accelerometers for 24-hour movement measurement in a range of research studies, as well as providing an important basis for future validation work of proxy-reported tools against accelerometers.

Future work specific to evaluating MoveMEY is required to further examine the utility of the tool for measuring movement behaviours of pre-school children. This will include:

- Conducting a study to examine the criterion validity of MoveMEY, through comparison of the tool with a range of established reference methods, such as direct observation, accelerometers, and sleep diaries.
- Conducting a test-retest reliability study of MoveMEY, alongside a comparison tool to determine if the tool is stable at detecting the behaviours, as well as to determine whether any differences reported are as a result of changes attributable to the tool or changes in behaviour.
- Testing MoveMEY with large samples of children to determine the usability of the tool at scale, particularly with sub-groups not involved in the development of the tool. This may be done in conjunction with local authorities to determine if the tool could be administered at this level.
- Examining the applicability of the tool with individuals in ethnically diverse groups. A way in which this may be achieved may be through working with established gatekeepers in settings with whom individuals from different ethnic backgrounds feel they can identify with and trust.
- Determine whether MoveMEY, a proxy reported tool, has the ability to detect composition of the 24 hour day, by analysing the data using compositional data analysis techniques.
- Explore variation in tool administration, including online (app) versus paper based tools. This would apply to MoveMEY, but also as an area of research more

generally to examine the effectiveness and accuracy of tools in these different formats.

7.6 Overall conclusion

Movement Measurement in the Early Years (MoveMEY) is a novel tool, the first co-developed (with parents and carers) tool for use to assess the full spectrum of movement behaviours of pre-school aged children. MoveMEY is a promising tool for the measurement of movement behaviours of pre-school aged children; demonstrated through co-development and parental views on acceptability of the tool, piloting of the tool and from demonstrating good content validity (including being relevant, comprehensive, and comprehensible). Further evaluation work of MoveMEY is required to determine the validity (in comparison with other tools) and reliability of the tool. Additionally, work is required to ensure usability and feasibility of the tool at scale, and with different sub-groups not involved in the development and content validity assessment of the tool.

Appendix A: Abstracts of published research

Phillips et al. *International Journal of Behavioral Nutrition and Physical Activity*
(2021) 18:141
<https://doi.org/10.1186/s12966-021-01132-9>

International Journal of Behavioral
Nutrition and Physical Activity

REVIEW

Open Access



A systematic review of the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children

Sophie M. Phillips^{1,2*}, Carolyn Summerbell^{1,2}, Matthew Hobbs³, Kathryn R. Hesketh⁴, Sonia Saxena⁵, Cassey Muir^{2,6} and Frances C. Hillier-Brown^{2,6,7,8}

Abstract

Physical activity (PA) and sedentary behaviour (SB) of pre-school aged children are associated with important health and developmental outcomes. Accurate measurement of these behaviours in young children is critical for research and practice in this area. The aim of this review was to examine the validity, reliability, and feasibility of measurement tools used to assess PA and SB of pre-school aged children.

Searches of electronic databases, and manual searching, were conducted to identify articles that examined the measurement properties (validity, reliability or feasibility) of measurement tools used to examine PA and/or SB of pre-school aged children (3–7 years old). Following screening, data were extracted and risk of bias assessment completed on all included articles.

A total of 69 articles, describing 75 individual studies were included. Studies assessed measurement tools for PA ($n = 27$), SB ($n = 5$), and both PA and SB ($n = 43$). Outcome measures of PA and SB differed between studies (e.g. moderate to vigorous activity, step count, posture allocation). Most studies examined the measurement properties of one measurement tool only ($n = 65$). Measurement tools examined included: calorimetry, direct observation, combined heart rate and accelerometry, heart rate monitors, accelerometers, pedometers, and proxy report (parent, carer or teacher reported) measures (questionnaires or diaries). Studies most frequently assessed the validity (criterion and convergent) ($n = 65$), face and content validity ($n = 2$), test-retest reliability ($n = 10$) and intra-instrument reliability ($n = 1$) of the measurement tools. Feasibility data was abstracted from 41 studies.

Multiple measurement tools used to measure PA and SB in pre-school aged children showed some degree of validity, reliability and feasibility, but often for different purposes. Accelerometers, including the Actigraph

* Correspondence: sophie.m.phillips@durham.ac.uk

¹Department of Sport and Exercise Sciences, Durham University, Durham City, UK

²The Centre for Translational Research in Public Health (Fuse), Newcastle upon Tyne, UK

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



The Validity, Reliability, and Feasibility of Measurement Tools Used to Assess Sleep of Pre-school Aged Children: A Systematic Rapid Review

Sophie M. Phillips^{1,2*}, Carolyn Summerbell^{1,2}, Helen L. Ball^{2,3}, Kathryn R. Hesketh^{4,5}, Sonia Saxena⁶ and Frances C. Hillier-Brown^{2,7,8,9}

¹ Department of Sport and Exercise Sciences, Durham University, Durham City, United Kingdom, ² The Centre for Translational Research in Public Health (Fuse), Newcastle Upon Tyne, United Kingdom, ³ Department of Anthropology, Infancy and Sleep Centre, Durham University, Durham City, United Kingdom, ⁴ Department of Population Policy & Practice Research and Teaching, UCL Great Ormond Street Institute of Child Health, London, United Kingdom, ⁵ Centre for Diet and Activity Research and MRC Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom, ⁶ School of Public Health, Imperial College London, London, United Kingdom, ⁷ Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, United Kingdom, ⁸ Human Nutrition Research Centre, Newcastle University, Newcastle Upon Tyne, United Kingdom, ⁹ Newcastle University Centre of Research Excellence in Healthier Lives, Newcastle University, Newcastle Upon Tyne, United Kingdom

OPEN ACCESS

Edited by:

Patrick De Boever,
University of Antwerp, Belgium

Reviewed by:

Vincenzo Natale,
University of Bologna, Italy
Amrita Dasgupta,
University of California, San Diego,
United States

*Correspondence:

Sophie M. Phillips
sophie.m.phillips@durham.ac.uk

Specialty section:

This article was submitted to
Children and Health,
a section of the journal
Frontiers in Pediatrics

Received: 03 September 2021

Accepted: 22 October 2021

Published: 26 November 2021

Citation:

Phillips SM, Summerbell C, Ball HL,
Hesketh KR, Saxena S and
Hillier-Brown FC (2021) The Validity,
Reliability, and Feasibility of
Measurement Tools Used to Assess
Sleep of Pre-school Aged Children: A
Systematic Rapid Review.
Front. Pediatr. 9:770262.
doi: 10.3389/fped.2021.770262

Background: Sleep of pre-school aged children is important for their health and development, but there are currently no standards for measuring sleep in this age group. We aimed to examine the validity, reliability and feasibility of tools used to assess sleep of pre-school aged children.

Methods: Studies were eligible for inclusion if they examined the validity and/or reliability and/or feasibility of a measurement tool used to examine sleep of pre-school aged children (aged 3–7 years). We systematically searched six electronic databases, grey literature and trial registries. We manually searched topic specific journals, reference and citations of included studies, and reference lists of existing reviews. We extracted data and conducted a risk of bias assessment on the included studies using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) risk of bias checklist. We used a narrative synthesis to present the results.

Results: Sixteen studies met the inclusion criteria: these explored accelerometers ($n = 3$) and parental reported tools ($n = 13$; nine questionnaires, six diaries). Studies assessed construct validity ($n = 3$), criterion validity ($n = 1$), convergent validity ($n = 13$), test-retest reliability ($n = 2$), internal consistency ($n = 4$) and feasibility ($n = 12$). Most studies assessed the convergent validity of questionnaires and diaries compared with accelerometers, but the validity of accelerometers for sleep in this age group is unknown. Of studies with a low risk of bias, one sleep diary was shown to be valid for measuring sleep duration. No measurement tools were appropriate for determining sleep quality. Reporting of reliability and feasibility was minimal.



Article

Parental Views on the Acceptability and Feasibility of Measurement Tools Used to Assess Movement Behaviour of Pre-School Children: A Qualitative Study

Sophie M. Phillips ^{1,2,*}, Carolyn Summerbell ^{1,2}, Kathryn R. Hesketh ^{3,4}, Sonia Saxena ⁵ and Frances C. Hillier-Brown ^{2,6,7,8}

- ¹ Department of Sport and Exercise Sciences, Durham University, Durham DH1 3HN, UK; carolyn.summerbell@durham.ac.uk
 - ² The Centre for Translational Research in Public Health (Fuse), Newcastle University, Newcastle upon Tyne NE1 7RU, UK; frances.hillier-brown@newcastle.ac.uk
 - ³ MRC Epidemiology Unit, University of Cambridge, Cambridge CB2 0QQ, UK; krh40@medschl.cam.ac.uk
 - ⁴ Population Policy & Practice Research and Teaching Department, UCL Great Ormond Street Institute of Child Health, London WC1N 1EH, UK
 - ⁵ School of Public Health, Imperial College London, London W6 8RP, UK; s.saxena@imperial.ac.uk
 - ⁶ Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne NE2 4AX, UK
 - ⁷ Human Nutrition Research Centre, Newcastle University, Newcastle upon Tyne NE2 4HH, UK
 - ⁸ Newcastle University Centre of Research Excellence in Healthier Lives, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
- * Correspondence: sophie.m.phillips@durham.ac.uk

Citation: Phillips, S.M.; Summerbell, C.; Hesketh, K.R.; Saxena, S.; Hillier-Brown, F.C. Parental Views on the Acceptability and Feasibility of Measurement Tools Used to Assess Movement Behaviour of Pre-School Children: A Qualitative Study. *Int. J. Environ. Res. Public Health* **2022**, *19*, x. <https://doi.org/10.3390/xxxx>

Academic Editor: Paul B. Tchounwou

Received: 18 February 2022

Accepted: 17 March 2022

Published: 21 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: Movement behaviours (physical activity, sedentary behaviour, and sleep) are important for the health and development of pre-school children (aged 3–4 years). There is limited qualitative research examining the acceptability and feasibility of tools used to assess movement behaviours in pre-schoolers. This study explored parental views on various measurement tools in three deprived areas in England, UK (West Yorkshire, County Durham and Northumberland). The study consisted of a demonstration of the different tools (accelerometers, a diary and a questionnaire), directly followed by focus group discussions. Three focus group discussions with a total of eleven parents and carers were transcribed verbatim and analysed using thematic analysis. Findings revealed four main themes: (1) importance of contextual information when using any measurement tool (e.g., child illness, capturing different routines); (2) practical issues associated with devices (e.g., aversion to devices being attached directly to the skin of their child; concern of larger devices during sleep time); (3) encouraging children to wear a device (e.g., making devices attractive to children —‘superpowers’); and (4) presentation of diaries and questionnaires (e.g., age-appropriate movement activities, preference for real-time recording over recall). Practical recommendations for the use of the tools to measure movement behaviours of pre-school children are provided.

Keywords: feasibility; acceptability; movement behaviours; measurement; qualitative research; pre-school children

RESEARCH

Open Access



A critique of the English national policy from a social determinants of health perspective using a realist and problem representation approach: the 'Childhood Obesity: a plan for action' (2016, 2018, 2019)

Naomi Griffin^{1,2*}, Sophie M. Phillips^{1,2}, Frances Hillier-Brown^{1,3}, Jonathan Wistow^{1,4}, Hannah Fairbrother⁵, Eleanor Holding⁶, Katie Powell⁶ and Carolyn Summerbell^{1,2}

Abstract

Background: The UK government released Chapter 1 of the 'Childhood Obesity: a plan for action' (2016), followed by Chapter 2 (2018) and preliminary Chapter 3 was published for consultation in 2019 (hereon collectively '*The Policy*'). The stated policy aims were to reduce the prevalence of childhood obesity in England, addressing disparities in health by reducing the gap (*approximately two-fold*) in childhood obesity between those from the most and least deprived areas.

Methods: Combining a realist approach with an analysis of policy discourses, we analysed the policies using a social determinants of health (SDH) perspective (focusing on socio-economic inequalities). This novel approach reveals how the framing of policy 'problems' leads to particular approaches and interventions.

Results: While recognising a social gradient in relation to obesity measures, we critique obesity problem narratives. *The Policy* included some upstream, structural approaches (e.g. restrictions in food advertising and the soft-drinks industry levy). However, the focus on downstream individual-level behavioural approaches to reduce calorie intake and increase physical activity does not account for the SDH and the complexity and contestedness of 'obesity' and pays insufficient attention to how proposals will help to reduce inequalities. Our findings illustrate that individualising of responsibility to respond to what wider evidence shows is structural inequalities, can perpetuate damaging narratives and lead to ineffective interventions, providing caution to academics, practitioners and policy makers (local and national), of the power of problem representation. Our findings also show that the problem framing in *The Policy* risks reducing important public health aims to encourage healthy diets and increase opportunities for physical activity (and the physical and mental health benefits of both) for children to weight management with a focus on particular children.

Conclusions: We propose an alternative conceptualisation of the policy 'problem', that obesity rates are illustrative of inequality, arguing there needs to be policy focus on the structural and factors that maintain health inequalities,

*Correspondence: naomi.c.griffin@durham.ac.uk

² Department of Sport and Exercise Sciences, Durham University, Durham, UK

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this

Appendix B: Training log

Committee Member	Date
National Institute for Health Research, School for Public Health Research, Researcher Network (ResNET), Fuse representative/committee member	Sept 2021 to present
Chief Medical Officers Physical Activity Surveillance Committee	Jan 2020-Sept 2021
Training Courses	
National Institute for Health Research Applied Research Collaboration <i>'Introduction to qualitative evaluation'</i>	August 2021
National Institute for Health Research, Centre for engagement and dissemination <i>'Dissemination guidance and public involvement and engagement'</i>	April 2021
Sport Structures <i>'Equality, Diversity and Inclusion Training'</i>	March 2021
National Institute for Health Research School for Public Health Research/ School for Primary Care Research <i>'Improving diversity in research'</i>	February 2021
National Institute for Health Research <i>'Public involvement and engagement event'</i>	November 2020
Communication Training <i>'Delivery skills'</i>	November 2020
Communication Training <i>'Content Creation'</i>	November 2020
Social Research Association	May 2019

<i>Focus Group Training</i>	
Researcher Development Course, Durham University <i>'Endnote Training'</i>	January 2019
Researcher Development Course, Durham University <i>'Long Documents in Word'</i>	January 2019
Postgraduate Training, Department of Sociology, Durham University <i>'Getting published'</i>	December 2018
Fuse Workshop, Teesside University <i>Social Media Workshop</i>	December 2018
Researcher Development Course, Durham University <i>Effective Search Strategies</i>	October 2018
Conferences Attended	
International Society of Physical Activity and Health Conference (Online)	October 2021
Public Health England conference (Online)	May 2021
National Institute for Health Research, School for Public Health Research Annual Scientific Meeting (Online)	May 2021
National Institute for Health Research Academy Conference <i>'Research for everyone'</i> (Online)	November 2020
International Society of Physical Activity and Health Conference (Online)	October 2020
International/European Congress on Obesity 2020 (Online)	Sept 2020
Prospective Physical Activity, Sitting, and Sleep Consortium 3 rd Annual Meeting (Online)	June 2020
UK Congress on Obesity, Leeds	September 2019
International Conference of Ambulatory Monitoring and Physical Activity Measurement, Maastricht	June 2019

National Institute for Health Research, School for Public Health Research Annual Scientific Meeting, Newcastle	March 2019
Seminars/Webinars Attended	
International Society of Physical Activity and Health <i>'Building an international database of sedentary behaviours'</i>	February 2022
Newcastle Health Innovation Partners seminar series <i>'Improving health, reducing inequalities. We need to start much earlier!'</i>	January 2022
Outdoor Play and Learning <i>'The case for play in schools'</i>	October 2021
Fuse Research Meeting <i>'From aspirations to action: how to develop meaningful integrated approaches between research and practice to address health inequalities in early life'</i>	June 2021
National Institute for Health Research, School for Public Health Research <i>Researcher Network Meeting</i>	June 2021
Participatory action research with children and young people seminar	June 2021
National Institute for Health Research, Applied Research Collaboration <i>'Child poverty and children's health'</i>	April 2021
The Waterloo Foundation CoDA project <i>'Children's composition of movement related behaviours in relation to a series of health outcomes'</i>	April 2021
World Obesity Policy Lab <i>'Physical inactivity and obesity: how can we break the cycle?'</i>	April 2021
International Society of Physical Activity and Health and World Health Organization joint seminar series <i>'World Health Organization guidelines on physical activity and sedentary behaviour: translating science to practice'</i>	March 2021

Sport England <i>'New strategy launch event'</i>	January 2021
National Health and Nutrition Examination Survey (NHANES) <i>'Devices in physical activity surveillance'</i>	January 2021
Newcastle University <i>'The international surveillance of physical of physical fitness among children and young people'</i>	January 2021
Fuse Physical Activity Workshop <i>'From science to the real world: how can we improve physical activity practices locally and nationally'</i>	January 2021
Understanding inequalities <i>'Succeeding against all odds- which factors matter?'</i>	December 2020
International Society of Physical Activity and Health <i>'The Science Behind the World Health Organization guidelines on PA'</i>	December 2020
Department of Sociological Sciences, The University of Sheffield <i>'How can intersectionality further understand on health inequalities?'</i>	November 2020
University of Wollongong, Early Start <i>'Early childhood education and care in a shifting landscape'</i>	September 2020
Association for the Study of Obesity Webinar Series <i>'Obesity and Stigma'</i>	August 2020
MRC Epidemiology Seminar Series <i>'Need for standardisation of global surveillance seminar'</i>	August 2020
'MOVING' policy database launch on physical activity and nutrition	June 2020

<i>'Movement behaviours and cognitive development in early childhood: Evidence, insights and interventions from South Africa'</i> presented by Catherine Draper	May 2020
Department of Sport and Exercise Sciences, Durham University <i>'Introduction to dual-energy X-ray absorptiometry (DXA)'</i>	May 2019
National Institute for Health Research, School for Public Health Research <i>Researcher Network Meeting</i>	March 2019
Department of Sociology, Durham University- Higher education and social inequalities seminar <i>'What counts as socioeconomically disadvantaged?'</i>	March 2019
Fuse Research Meeting <i>'Complex Systems Research Meeting'</i>	February 2019
Fuse Research Meeting <i>'Complexities and Fluidities in Health Research Seminar'</i>	December 2018
Fuse Research Meeting <i>'Physical Activity Workshop'</i>	December 2018
Department of Sport and Exercise Sciences, Durham University <i>Weekly research seminars</i>	October 2018 onwards

Appendix C: Knowledge exchange log

Conference and research meeting presentations	Date
NIHR SPHR Annual Scientific Meeting <i>Three minute thesis video</i>	May 2021
Faculty of Social Sciences and Health conference, Durham University – ‘Global challenges’ <i>Oral presentation- Using the social determinants of health to explore health inequalities in national policy on childhood obesity</i>	November 2020
International / European Congress on Obesity <i>Oral presentation- A systematic review of the validity, reliability, and feasibility of the measurement tools used to assess physical activity and sedentary behaviour of pre-school aged children</i>	September 2020
Department of Sport and Exercise Science Department Durham University, <i>Oral Presentation- PhD Research Progress and Plans</i>	June 2019
Wolfson Early Career Research Conference, Durham University <i>Poster Pitch Presentation – A systematic review of the validity, reliability, and feasibility of measurement tools used to assess physical activity and sedentary behaviour: A systematic review protocol</i>	June 2019
Fuse Early Years and Adolescence Research Meeting, Durham Town Hall <i>Poster Presentation - Acceptability and feasibility of measurement tools</i>	May 2019
NIHR SPHR Annual Scientific Meeting, Newcastle University <i>Poster Presentation – PhD Research Plans</i>	March 2019
Invited Speaker Presentations	
Reproduction, Development and Child Health Theme Meeting, Newcastle University <i>‘Measuring movement behaviours at population scale’</i>	December 2021

Child Health Unit Meeting, Imperial College London <i>'What do we know about the quality of tools used to measure physical activity and sedentary behaviour of pre-schoolers?'</i>	November 2021
MSc Health Psychology, Staffordshire University <i>'Reflections on entering a PhD...'</i>	April 2021
Department of Sport and Exercise Science, Durham University <i>'Experiences of being embedded in the local active partnership'</i>	March 2021
Chief Medical Officers Physical Activity guideline surveillance committee meeting <i>Presented on validity and reliability of measurement tools for pre-school aged physical activity</i>	June 2020
Division of Health Psychology Annual Conference <i>Masters Research</i>	July 2019
Child Health Unit Meeting, Imperial College London <i>'Measurement of movement related behaviours of pre-school children'.</i>	November 2019
Born in Bradford, Bradford Institute for Health Research Seminar Series <i>Measurement of movement-related behaviours of pre-school children at scale: PhD research plans.</i>	September 2019
Presentations to non-academic audiences	
County Durham Sport, Sport England funded Active Partnership. <i>Physical activity inequalities in County Durham. Findings from an analysis of the Active Lives Survey</i>	June 2021
Evidence Based Approach group, Durham County Council <i>What factors contribute towards activity/inactivity in County Durham? Methods and preliminary findings from 4 year Active Lives analysis.</i>	May 2021
County Durham Sport, Sport England funded Active Partnership. <i>'Early years physical activity'.</i>	March 2021

Research briefings	
NIHR SPHR Research briefing. <i>A critique of the 'Childhood Obesity: a plan for action' English national policy using a social determinants of health perspective.</i> Accessed here .	Dec 2021
NIHR SPHR Research briefing. <i>'Measuring 24 hour movement of pre-school children: how do we measure sleep?'</i> Accessed here .	Nov 2021
Fuse Research Briefing. <i>'How should we measure physical activity and sedentary behaviour of pre-school aged children?'</i> Accessed here .	Nov 2021
Blogs, infographics, other articles	
Explanatory Infographic of early years physical activity. <i>'Let's Get Little Ones Moving!'</i> . Accessed here .	August 2021
Fuse blog <i>'Working together to tackle inequalities and inactivity.'</i> Accessed here .	April 2021
News Article <i>PhD researcher on secondment at County Durham Sport.</i> Accessed here .	March 2021
Awards	
UKRI Strategic Priority Funding (see Appendix E) <i>Secured three month secondment working at County Durham Sport, in collaboration with colleagues in the Durham Research Methods Centre and Department of Sport and Exercise Science, Durham University</i>	Jan-March 2021
SPHR Summer Intern Funding (See Appendix D) <i>Secured funding to host a summer intern</i>	Funding secured April 2021, intern

	supervised July- August 2021
Wolfson Early Career Researcher Conference <i>Best Poster Pitch Prize</i>	June 2019

Appendix D: Additional projects worked on alongside PhD

Alongside my PhD research I have been involved with several additional projects:

National Institute for Health Research, School for Public Health Research (NIHR SPHR) funded project examining effectiveness of Forest School on health and development of pre-school children (2021)

I secured funding from NIHR SPHR to work with a summer intern on a project that involved examining the effectiveness of Forest School on the health and development of pre-school children. The purpose of this project was to examine what is known from the existing literature about the effectiveness of Forest School on health and development of pre-school children, as well as identifying accessibility to the programme. A large focus of the work was to engage with stakeholders and practitioners who lead these programmes in the community. The outputs of the work can be accessed here: a [blog](#), [research briefing](#), [full research report](#).

NIHR SPHR funded project examining policy context for reducing health inequalities among children and young people (2020-2021)

I was involved with a further NIHR SPHR project examining the policy context for reducing health inequalities among children and young people. The part of the project I worked on involved a multidisciplinary and multi-institutional team of researchers, examining the English childhood obesity policies. My role included data extraction of policy documents using two analytical frameworks, analysis of the policy, writing up the paper for publication (with the lead and third author), co-writing a research briefing, and I presented the work at local conference. The peer-reviewed journal article can be found [here](#). The associated research briefing can be found [here](#).

Appendix E: Placement in practice

UKRI funded secondment with County Durham Sport (2020)

I completed a three month UKRI funded secondment with County Durham Sport, a Sport England funded local active partnership. This involved collaborations with County Durham Sport and Durham County Council. During my time, I worked on three different projects, these included: developing a report on early years physical activity and movement; 2) health data analysis of available data to identify groups most in need of physical activity intervention in County Durham; 3) analysis of the Active Lives Survey data to understanding sociodemographic inequalities in community-based physical activity. Through this work I was able to develop collaborations with external partners, input into local physical activity policy, provide evidence for a funding bid for an intervention to reduce physical activity inequalities, and work with large-data sets included extensive data cleaning and interpretation of statistics.

Chapter Appendices

Appendix 3.1: PRISMA Checklist for Systematic Review

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	137
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	23
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	138-139
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	140
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	140-141
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	144-145
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	141-143
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	142-143
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	144-145

Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	145-146
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	145-146
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	146-148
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	152-154
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	155-156

Page 1 of 2

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	146-148
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	155-156
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	157
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	158-162
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	163-164, appendix 3.6 and 3.7
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	164-280
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	164-280

Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	163-164
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	279-280
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	280-285
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	286-287
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	289-290
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	14

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

Appendix 3.2: PROPSERO Systematic Review Protocol



CRD42019133613_PR
OSPERO Systematic R

Appendix 3.3: Inclusion checklist

Systematic Review Inclusion Checklist

A systematic review to identify and summarise the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviours of pre-school children.

General Information:

Study Title	
Authors of Study:	
Year of Study:	
Citation:	

Inclusion Checklist:

Does the study explore a measurement tool that is used to assess physical activity and/or sedentary time in a pre-school population?	Yes ☐ No ☐ → Exclude
Is the primary aim of the study to explore the quality (validity and/or reliability and/or feasibility) of the measurement tool for assessing physical activity and/or sedentary behaviours?	Yes ☐ No ☐ → Exclude
Does the study include typically developing pre-school children aged 3-7 years, mean age of ≤6 years old?	Yes ☐ No ☐ → Exclude
Does the study include individuals over the age of 7 years?	Yes ☐ → Exclude No ☐
Does the study include individuals under the age of 3 years? (0-12 months, 1 year or 2 years old).	Yes ☐ → Exclude No ☐
Does the study include children with conditions/chronic diseases (i.e cerebral palsy)?	Yes ☐ → Exclude No ☐
Is the measurement tool being explored specific to a certain type of activity only (i.e gardening)?	Yes ☐ → Exclude No ☐
Is the measurement tools designed for a targeted pre-school population only (i.e children from care homes)?	Yes ☐ → Exclude No ☐
Does the study explore the cut-points or wear time of devices only?	Yes ☐ → Exclude No ☐
Is the study a primary piece of research (i.e not a review)?	Yes ☐ → Exclude No ☐

Summary of Assessment for Inclusion:

Include in review ☐	Exclude from review ☐
Please provide reasons why the study has been included/excluded in the review →	
Independently assessed, and then compared? Yes ☐ No ☐	Differences resolved Yes ☐ No ☐
Notes:	

Appendix 3.4: Data Extraction Form

Systematic Review Data Extraction Form

A systematic review to identify and summarise the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviours of pre-school children.

General Information:

Study ID (e.g 01):	Data extractor:	Date form completed:
Authors of Study:	Year of Study:	Country of Study:
Citation:		
Publication type :Journal Article ☒ Abstract ☐ Other (specify e.g. book chapter) _____		
Funding source of study:	Potential conflict of interest from funding? Y / N / unclear	

Study Information:

Study Characteristics	
Study Design/Method:	
Aim of Study:	
Study Setting (consider laboratory versus real world application):	
Total study duration:	
Study scale (did the study explore the methods on a large scale?):	
Does the study meet the criteria for inclusion?	Yes ☐ No ☐ → Exclude If no, please provide details:

Participant Information:

Participant Recruitment	
Method of Recruitment of Participants:	
Inclusion/exclusion criteria for participation in study:	

Representativeness of sample: Are participants in the study likely to be representative of the target population? (Consider SES)	
What percentage of selected individuals agreed to participate?	
Sample Size	
Attrition Rates	
Number and reason for (and sociodemographic differences of) withdrawals and exclusions from study	
What percentage of participants successfully provided data used the measurement tool being evaluated?	

Participant Characteristics	
Age range of participants:	
Mean age of participants:	
Sex of participants:	
Ethnicity of participants:	
Socioeconomic status of participants:	
Other sociodemographics (eg. Educational level, literacy level, first language)	
Are participants described as having a specific condition/chronic disease, such as cerebral palsy?	Yes ☐ No ☐ Unclear ☐ Details:
Are participants defined as a group having specific social or cultural characteristics?	Yes ☐ No ☐ Unclear ☐ Details:
Additional details of participants:	
Subgroups	<i>Enter a description of any participant subgroups from this paper to be analysed in the review.</i>
<i>Do the participants meet the criteria for inclusion?</i>	Yes ☐ No ☐ → Exclude If no, please provide details:

Outcomes:

Measurement Tools	
Assessment Tool(s) Explored:	
Comparator/Criterion Measure:	
If no comparator/criterion measure used, please provide details why:	
Cut points used (if applicable):	
Epoch length used (if applicable):	

Placement of devices (if applicable):	
Outcome Measure	
Outcome Behaviour Measured (Physical activity and/or sedentary behaviour)	
Type/Details of Outcome Measure (e.g step counts, MVPA, energy expenditure etc.)	
Type of Evaluation	
Type of Evaluation: Reliability (e.g test-retest)	
Type of Evaluation: Validity (e.g face, content, criterion)	
Type of Evaluation: Feasibility	
Additional Evaluation- provide details:	
Please provide additional details in relation to: assessment tools explored, outcome measures and/or type of evaluation conducted.	
Do the outcome measures meet the criteria for inclusion?	Yes ☐ No ☐ → Exclude If excluded, please provide details why:

Results:

Results of the Study:	
Methods of validity/reliability/feasibility testing used and appropriateness of these methods	
Were results relating to unintended effects/undesirable findings described? Include any evidence.	
Potential for author conflict <i>i.e. evidence that author or data collectors would benefit if results favoured the assessment tool under study or the control</i>	
Key conclusions of the study authors	

Other Potentially Relevant Information:

Could the inclusion of this study potentially bias the generalisability of the review? <i>Equity pointer: Remember to consider whether disadvantaged populations may have been excluded from the study.</i>	
Issues affecting directness <i>(Note any aspects of population, intervention, etc. that affect this study's direct applicability to the review question)</i>	

References to other relevant studies	
Additional notes by review authors	
Request further details/study information? Yes ☐ No ☐	
Contact details of authors:	
Date contacted:	
Information Requested/Obtained:	

Summary of Assessment for Inclusion:

Independently assessed, and then compared? ☐ Yes ☐ No ☐	Differences resolved Yes ☐ No ☐
Notes:	

Appendix 3.5: List of excluded studies with reasons

Total excluded following full paper screening (n=77).

(Exclusion criteria 1) Study did not examine the utility of the tool (n=18):

1. De Craemer, M., De Decker, E., Verloigne, M., De Bourdeaudhuij, I., Manios, Y., Cardon, G., & ToyBox-study group. (2016). The effect of a cluster randomised control trial on objectively measured sedentary time and parental reports of time spent in sedentary activities in Belgian preschoolers: the ToyBox-study. *International Journal of Behavioral Nutrition and Physical Activity*, 13(1), 1.
2. Downing, K. L., Janssen, X., Cliff, D. P., Okely, A. D., & Reilly, J. J. (2019). Energy expenditure associated with posture transitions in preschool children. *PloS one*, 14(4).
3. Flynn, J. I., Coe, D. P., Larsen, C. A., Rider, B. C., Conger, S. A., & Bassett, J. D. (2014). Detecting indoor and outdoor environments using the ActiGraph GT3X+ light sensor in children. *Medicine and science in sports and exercise*, 46(1), 201-206.
4. LOGAN, N., REILLY, J. J., GRANT, S., & PATON, J. Y. (2000). Resting heart rate definition and its effect on apparent levels of physical activity in young children. *Medicine & science in sports & exercise*, 32(1), 162.
5. Montgomery, C., Reilly, J. J., Jackson, D. M., Kelly, L. A., Slater, C., Paton, J. Y., & Grant, S. (2004). Relation between physical activity and energy expenditure in a representative sample of young children. *The American journal of clinical nutrition*, 80(3), 591-596.
6. Mouratidou, T., Miguel, M. L., Androutsos, O., Manios, Y., De Bourdeaudhuij, I., Cardon, G., ... & Payr, A. (2014). Tools, harmonization and standardization procedures of the impact and outcome evaluation indices obtained during a kindergarten-based, family-involved intervention to prevent obesity in early childhood: the ToyBox-study. *obesity reviews*, 15, 53-60.
7. O'Connor, T. M., Cerin, E., Robles, J., Lee, R. E., Kerr, J., Butte, N., ... & Baranowski, T. (2013). Feasibility study to objectively assess activity and location of Hispanic preschoolers: a short communication. *Geospatial health*, 7(2), 375.
8. Puyau, M. R., Adolph, A. L., Liu, Y., Wilson, T. A., Zakeri, I. F., & Butte, N. F. (2016). Energy cost of activities in preschool-aged children. *Journal of Physical Activity and Health*, 13(s1), S11-S16.

9. Reilly, J. J., Janssen, X., Cliff, D. P., & Okely, A. D. (2015). Appropriateness of the definition of 'sedentary' in young children: Whole-room calorimetry study. *Journal of science and medicine in sport*, 18(5), 565-568.
10. Saris, W. H. M., Snel, P., & Binkhorst, R. A. (1977). A portable heart rate distribution recorder for studying daily physical activity. *European journal of applied physiology and occupational physiology*, 37(1), 17-25.- Age of children unknown, also does not appear to be evaluating the tools in relation to PA and/or SB, but rather just looking at heart rate.
11. Song, K. M., Bjornson, K. F., Cappello, T., & Coleman, K. (2006). Use of the StepWatch activity monitor for characterization of normal activity levels of children. *Journal of Pediatric Orthopaedics*, 26(2), 245-249.

Explored epoch length only (n=3)

1. Colley, R. C., Harvey, A., Grattan, K. P., & Adamo, K. B. (2014). Impact of accelerometer epoch length on physical activity and sedentary behaviour outcomes for preschool-aged children. *Health Rep*, 25(1), 3-9.
2. Kim, Y., Beets, M. W., Pate, R. R., & Blair, S. N. (2013). The effect of reintegrating Actigraph accelerometer counts in preschool children: Comparison using different epoch lengths. *Journal of science and medicine in sport*, 16(2), 129-134.
3. Leeger-Aschmann, C. S., Schmutz, E. A., Zysset, A. E., Kakebeeke, T. H., Messerli-Bürgy, N., Stübli, K., ... & Puder, J. J. (2019). Accelerometer-derived physical activity estimation in preschoolers—comparison of cut-point sets incorporating the vector magnitude vs the vertical axis. *BMC public health*, 19(1), 513.

Tool explored PA environment, but not PA (n=2)

1. Doerr, C. S., Bock, C., Fischer, J. E., & De Bock, F. (2014). Preschools' Friendliness Toward Physical Activity: Item Battery and Two Scores Developed by Mixed Methods. *American journal of preventive medicine*, 46(6), 593-604.
2. Park, S. H., Park, C. G., McCreary, L., & Norr, K. F. (2017). Cognitive Interviews for Validating the Family Nutrition Physical Activity Instrument for Korean-American Families With Young Children. *Journal of pediatric nursing*, 36, 1-6

Inter-observer reliability only (n=2)

1. Brown, W. H., Pfeiffer, K. A., McIver, K. L., Dowda, M., Almeida, J. M., & Pate, R. R. (2006). Assessing preschool children's physical activity: the Observational System for Recording Physical Activity in children-preschool version. *Research quarterly for exercise and sport*, 77(2), 167-176.

2. McIver, K. L., Brown, W. H., Pfeiffer, K. A., Dowda, M., & Pate, R. R. (2009). Assessing children's physical activity in their homes: The observational system for recording physical activity in children-home. *Journal of applied behavior analysis*, 42(1), 1-16.

(Exclusion Criteria 2) Age of children outside range (n=31):

Included children under 3 years (n=10)

1. Bingham, D., Collings, P., Clemes, S., Costa, S., Santorelli, G., Griffiths, P., & Barber, S. (2016). Reliability and validity of the early years physical activity questionnaire (EY-PAQ). *Sports*, 4(2), 30.
2. Burdette, H. L., Whitaker, R. C., & Daniels, S. R. (2004). Parental report of outdoor playtime as a measure of physical activity in preschool-aged children. *Archives of pediatrics & adolescent medicine*, 158(4), 353-357.
3. De, F. B., Menze, J., Becker, S., Litaker, D., Fischer, J., & Seidel, I. (2010). Combining accelerometry and HR for assessing preschoolers' physical activity. *Medicine and science in sports and exercise*, 42(12), 2237-2243.
4. Henriksson, H., Forsum, E., & Löf, M. (2014). Evaluation of Actiheart and a 7 d activity diary for estimating free-living total and activity energy expenditure using criterion methods in 1-5-and 3-year-old children. *British Journal of Nutrition*, 111(10), 1830-1840
5. Klesges, L. M., & Klesges, R. C. (1987). The assessment of children's physical activity: a comparison of methods. *Medicine and Science in Sports and Exercise*, 19(5), 511-517.
6. KLESGES, R. C., Klesges, L. M., Swenson, A. M., & Pheley, A. M. (1985). A validation of two motion sensors in the prediction of child and adult physical activity levels. *American journal of epidemiology*, 122(3), 400-410.
7. Rice, K. R., Joschtel, B., & Trost, S. G. (2013). Validity of family child care providers' proxy reports on children's physical activity. *Childhood Obesity*, 9(5), 393-398.
8. Sarker, H., Anderson, L. N., Borkhoff, C. M., Abreo, K., Tremblay, M. S., Lebovic, G., ... & Birken, C. S. (2015). Validation of parent-reported physical activity and sedentary time by accelerometry in young children. *BMC research notes*, 8(1), 735.
9. Vale, S., Santos, R., Silva, P., Soares-Miranda, L., & Mota, J. (2009). Preschool children physical activity measurement: importance of epoch length choice. *Pediatric exercise science*, 21(4), 413-420.
10. Wiseman, N., Harris, N., & Downes, M. (2017). Validation of an iPad activity to measure preschool children's food and physical activity knowledge and preferences. *international journal of behavioral nutrition and physical activity*, 14(1), 11.

Included children over 7 years (n=17)

1. Aadland, E., Andersen, L. B., Ekelund, U., Anderssen, S. A., & Resaland, G. K. (2018). Reproducibility of domain-specific physical activity over two seasons in children. *BMC public health*, 18(1), 821.
2. Bender, J. M., Brownson, R. C., Elliott, M. B., & Haire-Joshu, D. L. (2005). Children's physical activity: using accelerometers to validate a parent proxy record. *Medicine and science in sports and exercise*, 37(8), 1409.
3. Busser, H. J., Ott, J., Van Lummel, R. C., Uiterwaal, M., & Blank, R. (1997). Ambulatory monitoring of children's activity. *Medical engineering & physics*, 19(5), 440-445.
4. Cale, L. (1993). *Monitoring physical activity in children* (Doctoral dissertation, Loughborough University).
5. Camargo, D. M., Santisteban, S., Paredes, E., Flórez, M. A., & Bueno, D. (2015). Confiabilidad de un cuestionario para medir la actividad física y los comportamientos sedentarios en niños desde preescolar hasta cuarto grado de primaria. *Biomédica*, 35(3), 347-356.
6. Cohen, A., McDonald, S., McIver, K., Pate, R., & Trost, S. (2014). Assessing physical activity during youth sport: the observational system for recording activity in children: youth sports. *Pediatric exercise science*, 26(2), 203-209.
7. Duncan, J. S., Badland, H. M., & Schofield, G. (2009). Combining GPS with heart rate monitoring to measure physical activity in children: A feasibility study. *Journal of Science and Medicine in Sport*, 12(5), 583-585.
8. Going, S. B., Levin, S., Harrell, J., Stewart, D., Kushi, L., Cornell, C. E., ... & Sallis, J. (1999). Physical activity assessment in American Indian schoolchildren in the Pathways study. *The American journal of clinical nutrition*, 69(4), 788S-795S.
9. Hendrick, P., Boyd, T., Low, O., Takarangi, K., Paterson, M., Claydon, L., & Milosavljevic, S. (2010). Construct validity of RT3 accelerometer: A comparison of level-ground and treadmill walking at self-selected speeds. *Journal of Rehabilitation Research & Development*, 47(2).
10. Mehdizadeh, A., Shafiee, M., Khadem-Rezaian, M., Sardar, M. A., Vatanparast, H., Rose, E., ... & Nematy, M. (2019). Evidence for the Validity of the Children's Attraction to Physical Activity (CAPA) Scale in Iranian Preschool Children. *Journal of pediatric nursing*, 44, e52-e57.
11. Nascimento-Ferreira, M. V., De Moraes, A. C. F., Toazza-Oliveira, P. V., Forjaz, C. L., Aristizabal, J. C., Santaliesra-Pasías, A. M., ... & Moreno, L. A. (2018). Reliability and validity of a questionnaire for physical activity assessment in South American children and adolescents: the SAYCARE study. *Obesity*, 26, S23-S30.

12. Nguyen, D. M. T., Lecoultre, V., Sunami, Y., & Schutz, Y. (2013). Assessment of physical activity and energy expenditure by GPS combined with accelerometry in real-life conditions. *Journal of Physical Activity and Health*, 10(6), 880-888.
13. Ojiambo, R., Konstabel, K., Veidebaum, T., Reilly, J., Verbestel, V., Huybrechts, I., ... & Bammann, K. (2012). Validity of hip-mounted uniaxial accelerometry with heart-rate monitoring vs. triaxial accelerometry in the assessment of free-living energy expenditure in young children: the IDEFICS Validation Study. *Journal of Applied Physiology*, 113(10), 1530-1536.
14. Rowe, P. J., Schuldheisz, J. M., & Van Der Mars, H. (1997). Validation of SOFIT for measuring physical activity of first-to eighth-grade students. *Pediatric Exercise Science*, 9(2), 136-149.
15. Rowlands, A.V. (1998). The measurement of physical activity in children, (Doctoral dissertation, Prifysgol Bangor University).
16. Van Camp, C. M., & Berth, D. (2018). Further evaluation of observational and mechanical measures of physical activity. *Behavioral Interventions*, 33(3), 284-296.
17. Van Camp, C. M., & Hayes, L. B. (2017). Identifying beneficial physical activity during school recess: Utility and feasibility of the Fitbit. *Journal of Behavioral Education*, 26(4), 394-409.

Mean of children >6 years (n=3)

1. Basterfield, L., Adamson, A. J., Parkinson, K. N., Maute, U., Li, P. X., Reilly, J. J., & Gateshead Millennium Study Core Team. (2008). Surveillance of physical activity in the UK is flawed: validation of the Health Survey for England Physical Activity Questionnaire. *Archives of disease in childhood*, 93(12), 1054-1058.
2. Tanaka, C., Tanaka, S., Kawahara, J., & Midorikawa, T. (2007). Triaxial accelerometry for assessment of physical activity in young children. *Obesity*, 15(5), 1233-1241.
3. Sekine, M., Yamagami, T., Chen, X., Hayashikawa, Y., Hamanishi, S., & Kagamimori, S. (2002). Validity of a questionnaire evaluating physical activity level in young children. *Environmental health and preventive medicine*, 6(4), 264-267.

Age not specified (n=1)

1. Ridgers, N. D., Stratton, G., & McKenzie, T. L. (2010). Reliability and validity of the System for Observing Children's Activity and Relationships during Play (SOCARP). *Journal of Physical Activity and Health*, 7(1), 17-25.

(Exclusion Criteria 3): Children with conditions or chronic disease (n=1):

1. Anderson, D. R., Field, D. E., Collins, P. A., Lorch, E. P., & Nathan, J. G. (1985). Estimates of young children's time with television: a methodological comparison of parent reports with time-lapse video home observation. *Child development*, 1345-1357.

(Exclusion Criteria 4) Wear time only (n=13):

1. Aadland, E., & Johannessen, K. (2015). Agreement of objectively measured physical activity and sedentary time in preschool children. *Preventive medicine reports*, 2, 635-639.
2. Addy, C. L., Trilk, J. L., Dowda, M., Byun, W., & Pate, R. R. (2014). Assessing preschool children's physical activity: how many days of accelerometry measurement. *Pediatric exercise science*, 26(1), 103-109-
3. Lima, R. A., Barros, S. S. H., Cardoso Júnior, C. G., Silva, G., Farias Júnior, J. C. D., Andersen, L. B., & Barros, M. V. G. D. (2014). Influence of number of days and valid hours using accelerometry on the estimates of physical activity level in preschool children from Recife, Pernambuco, Brazil. *Revista Brasileira de Cineantropometria & Desempenho Humano*, 16(2), 171-181.
4. Byun, W., Beets, M., & Pate, R. (2015). Sedentary behavior in preschoolers: How many days of accelerometer monitoring is needed?. *International journal of environmental research and public health*, 12(10), 13148-13161.
5. DuRant, R. H., Baranowski, T., Davis, H. A. R. R. Y., Rhodes, T. H. O. M. A. S., Thompson, W. O., Greaves, K. A., & Puhl, J. A. C. Q. U. E. L. I. N. E. (1993). Reliability and variability of indicators of heart-rate monitoring in children. *Medicine and Science in Sports and Exercise*, 25(3), 389-395.
6. Durant, R. H., Baranowski, T., Davis, H. A. R. R. Y., Thompson, W. O., Puhl, J. A. C. Q. U. E. L. I. N. E., Greaves, K. A., & Rhodes, T. H. O. M. A. S. (1992). Reliability and variability of heart rate monitoring in 3-, 4-, or 5-yr-old children. *Medicine and science in sports and exercise*, 24(2), 265-271.
7. DuRANT, R. H., Baranowski, T., Puhl, J., Rhodes, T., Davis, H., Greaves, K. A., & Thompson, W. O. (1993). Evaluation of the Children's Activity Rating Scale (CARS) in young children. *Medicine & Science in Sports & Exercise*
8. Hinkley, T., O'CONNELL, E. O. I. N., Okely, A. D., Crawford, D., Hesketh, K., & Salmon, J. (2012). Assessing volume of accelerometry data for reliability in preschool children. *Medicine & Science in Sports & Exercise*, 44(12), 2436-2441.

9. Hislop, J., Law, J., Rush, R., Grainger, A., Bulley, C., Reilly, J. J., & Mercer, T. (2014). An investigation into the minimum accelerometry wear time for reliable estimates of habitual physical activity and definition of a standard measurement day in pre-school children. *Physiological Measurement*, 35(11), 2213.
10. Kang, M., Bjornson, K., Barreira, T. V., Ragan, B. G., & Song, K. (2014). The minimum number of days required to establish reliable physical activity estimates in children aged 2–15 years. *Physiological measurement*, 35(11), 2229.
11. Penpraze, V., Reilly, J. J., MacLean, C. M., Montgomery, C., Kelly, L. A., Paton, J. Y., ... & Grant, S. (2006). Monitoring of physical activity in young children: how much is enough?. *Pediatric Exercise Science*, 18(4), 483-491.
12. Ricardo, L. I. C., Wendt, A., Galliano, L. M., de Andrade Muller, W., Niño Cruz, G. I., Wehrmeister, F., ... & Silva, C. M. (2020). Number of days required to estimate physical activity constructs objectively measured in different age groups: Findings from three Brazilian (Pelotas) population-based birth cohorts.
13. Toschke, J. A., von Kries, R., Rosenfeld, E., & Toschke, A. M. (2007). Reliability of physical activity measures from accelerometry among preschoolers in free-living conditions. *Clinical nutrition*, 26(4), 416-420.

(Exclusion Criteria 5) Calibration/prediction equation only (n=10):

1. Ahmadi, M. N., Brookes, D., Chowdhury, A., Pavey, T., & Trost, S. G. (2019). Free-living Evaluation of Laboratory-based Activity Classifiers in Preschoolers. *Medicine and science in sports and exercise*.
2. Bornstein, D. B., Beets, M. W., Byun, W., Welk, G., Bottai, M., Dowda, M., & Pate, R. (2011). Equating accelerometer estimates of moderate-to-vigorous physical activity: in search of the Rosetta Stone. *Journal of Science and Medicine in Sport*, 14(5), 404-410.
3. Butte, N. F., Wong, W. W., Lee, J. S., Adolph, A. L., Puyau, M. R., & Zakeri, I. F. (2014). Prediction of energy expenditure and physical activity in preschoolers. *Medicine and science in sports and exercise*, 46(6), 1216.
4. Johansson, E., Larisch, L. M., Marcus, C., & Hagströmer, M. (2016). Calibration and validation of a wrist-and hip-worn actigraph accelerometer in 4-year-old children. *PloS one*, 11(9), e0162436.
5. Li, S. (2019). *Machine Learning Approaches to Calibrate Wrist-Worn Accelerometry for Physical Activity Assessment in Preschoolers* (Doctoral dissertation, The University of Texas at San Antonio).

6. Roscoe, C. M., James, R. S., & Duncan, M. J. (2017). Calibration of GENEActiv accelerometer wrist cut-points for the assessment of physical activity intensity of preschool aged children. *European journal of pediatrics*, 176(8), 1093-1098.-calibration?
7. Tanaka, C., Hikiyama, Y., Ando, T., Oshima, Y., Usui, C., Ohgi, Y., ... & Tanaka, S. (2019). Prediction of Physical Activity Intensity with Accelerometry in Young Children. *International journal of environmental research and public health*, 16(6), 931.
8. Trost, S. G., Cliff, D., Ahmadi, M., Van Tuc, N., & Hagenbuchner, M. (2018). Sensor-enabled activity class recognition in preschoolers: Hip versus wrist data. *Medicine and science in sports and exercise*, 50(3), 634-641.
9. van Cauwenberghe, V., Labarque, V., Trost, S. G., De Bourdeaudhuij, I., & Cardon, G. (2011). Calibration and comparison of accelerometer cut points in preschool children. *International Journal of Pediatric Obesity*, 6(sup3), e582-589.
10. Zakeri, I. F., Adolph, A. L., Puyau, M. R., Vohra, F. A., & Butte, N. F. (2013). Cross-sectional time series and multivariate adaptive regression splines models using accelerometry and heart rate predict energy expenditure of preschoolers. *The Journal of nutrition*, 143(1), 114-122.

(Exclusion Criteria 6) Review (n=1):

1. Rowlands, A. V., Eston, R. G., & Ingledew, D. K. (1997). Measurement of physical activity in children with particular reference to the use of heart rate and pedometry. *Sports Medicine*, 24(4), 258-272.

(Exclusion Criteria 7) Insufficient information available (n=1):

1. Downing, K. L., Janssen, X., & Reilly, J. J. (2019). Feasibility of wearable cameras to assess screen time and time spent restrained in children aged 3 to 5 years: a study protocol. *BMJ open*, 9(5), e028265.

(Exclusion Criteria 8) Thesis, but associated papers had already been included (n=2):

1. Kelly, L. (2005). "Objectively measured physical activity and sedentary behaviour in young children."
2. Byun, W. (2013). Sedentary behavior in preschool children: Studies of reliability, level, and effect of objectively measured sedentary behavior, ProQuest Information & Learning. 73

Appendix 3.6: Downs and Black risk of bias assessment on all studies

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Study	Year	Is the aim clearly described?	Are the main outcomes to be measured clearly described in the introduction and methods section?	Reporting Are the characteristics of the participants included in the study clearly described?	Are the exposures of interest clearly described?	Are the main findings of the study clearly described?	Does the study provide estimates of the random variability in the data for the main outcomes?	Have the characteristics of patients lost to follow up been described?	Have actual probability values been reported for main outcomes, except where the	External validity Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	Were those subjects who were prepared to participate representative of the entire population from which they were	Were the staff, places, and facilities where participants were tested, representative of the testing that the	Internal validity- bias If any of the results of the study were based on 'data dredging' - was it made clear?	Do the analyses adjust for the time lag between the 2 measurements or different lengths of follow-up?	Were the statistical tests used to assess main outcomes appropriate?	Were the main outcome measures used accurate (valid and reliable)?	Internal validity- confounding Was there adequate adjustment for confounding in the analyses from which the main	Were the losses to participant follow-ups between measurements taken into account?
Adolph et al.	2012	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	No	Unclear	Not applicable	Yes	Yes	Yes	Unclear
Alghareed et al.	2010	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Alhassan et al.	2017	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Bacardi-Gascon et al.	2010	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Bar-On et al. (Study 1)	1998	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Yes
Bar-On et al. (Study 2)	1998	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Not applicable	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Blüchke	2014	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	No	Yes	Yes	Yes	Yes	Yes	Unclear
Bunn et al.	2018a	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Unclear
Bjorn et al.	2018b	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Cardon et al.	2007	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Chen et al.	2002	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Chow & Au	2009	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Corder et al.	2009	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	No	Yes	Yes	Yes	Yes
Davies et al. (Study 1)	2010	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Davies et al. (Study 2)	2010	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Not applicable	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
De Craemer et al.	2015	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
De Decker et al.	2010	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Difuria et al.	2010	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Dobell et al.	2019	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Deger et al.	2011	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Not applicable	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Elkous et al.	1991	Yes	Yes	Yes	Yes	Yes	Not applicable	No	Unclear	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ertinece et al.	2018	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Fairweather et al. (Study 1)	1999	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Unclear	Yes	Not applicable	Yes	Yes	Yes	Unclear
Fairweather et al. (Study 2)	1999	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Finn & Sprenkle	2000	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Fortini et al.	2015	Yes	No	No	No	No	Yes	Unclear	No	Unclear	Unclear	No	Yes	Not applicable	Yes	No	Unclear	Unclear
Gonzalez-Gil et al.	2014	Yes	No	No	No	Yes	Yes	Yes	Not applicable	No	Unclear	Yes	Yes	Yes	Yes	No	No	Yes
Hamdi & Larkin	2006	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Unclear	Unclear
Histop et al.	2012a	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Histop et al.	2012b	Yes	Yes	Yes	Yes	No	Yes	Unclear	No	Unclear	Unclear	Yes	No	Not applicable	Yes	Yes	Yes	Unclear
Histop et al.	2013	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Janssen et al.	2013a	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Unclear
Janssen et al.	2013b	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Yes
Janssen et al.	2013c	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Not applicable	Unclear	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Yes
Janssen et al.	2014	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Not applicable	Unclear	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Yes
Janssen et al.	2015	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Not applicable	Unclear	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Yes
Janz et al.	2005	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	Unclear	Yes	Yes	Yes	Yes	No	Yes
Kahan et al. (Study 1)	2010	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Kahan et al. (Study 2)	2010	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Kelly et al.	2004	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	Unclear	Yes	Not applicable	Yes	Yes	Yes	Yes
Kiesges et al.	1985	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Not applicable	Unclear	Yes	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Larson et al.	2011	Yes	Yes	Yes	No	Yes	Yes	Unclear	No	Unclear	Unclear	No	Yes	Not applicable	Unclear	Yes	Yes	Unclear
Lee et al.	2014	Yes	No	No	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	No	Unclear
Liggett et al.	2012	Yes	Yes	Yes	No	Unclear	Yes	Yes	No	Unclear	Unclear	No	Yes	Not applicable	Unclear	Yes	Yes	Yes
Lopez-Alarcon et al.	2004	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Louie & Chan	2003	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	Unclear	Yes	Not applicable	Yes	Yes	Yes	Yes
Manios et al.	1990	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear
Martini et al.	2011	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
McEneaney et al.	2005	Yes	Yes	Yes	Yes	No	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Not applicable	Unclear	Yes	Yes	Unclear
Mendoza et al.	2010	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murray	2009	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear
Nichols et al.	1982	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Noland et al.	1990	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	No	Yes	Yes	Yes	Unclear
Ngström et al.	2017	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Oliver et al.	2007	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Oortwijn et al.	2009	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Not applicable	Unclear	Unclear	No	Yes	Not applicable	not applicable	Yes	Yes	Yes
Pagels et al.	2011	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Pate et al.	2006	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	Partially	Yes	Not applicable	Yes	Yes	Yes	Yes
Pfeiffer et al.	2006	Yes	Yes	Yes	Yes	No	Yes	Unclear	Yes	Unclear	Unclear	Partially	Yes	Not applicable	Yes	Yes	Yes	Unclear
Pohl et al.	1990	No	Yes	No	Yes	Yes	Yes	Unclear	Yes	Not applicable	Unclear	Unclear	Yes	Not applicable	not applicable	Yes	No	Unclear
Riedl et al.	2006	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Unclear
Riedl et al.	2003	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Unclear	Unclear	No	Not applicable	No	Yes	Yes	Unclear
Saris & Binkhorst	1977	Yes	No	No	No	Yes	Yes	Unclear	No	Unclear	Unclear	Yes	No	Not applicable	Unclear	No	No	Unclear
Sharma et al. (Study 1)	2011	Yes	Yes	Yes	No	Yes	Yes	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear
Sharma et al. (Study 2)	2011	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	No
Sharp et al.	2017	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Shin	2015	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Yes
Siguma et al.	2010	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Sirard et al. (Study 2)	2005	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Sirard et al. (Study 1)	2005	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Yes
Steenhoek et al.	2019	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Not applicable	Unclear	Unclear	No	Yes	Not applicable	Yes	Yes	Yes	Yes
Telford et al.	2004	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Van Cawenberg et al.	2012	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Vanderloo et al.	2016	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes
Wen et al.	2010	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Not applicable	Yes	Yes	Yes	Yes

Appendix 3.7: COSMIN risk of bias assessment on proxy reported tools

Study	Year	Construct validity					Reliability								
		Is it clear what the comparator instrument measures?	Were the measurement properties of the comparator instrument sufficient?	Was the statistical method appropriate for the hypothesis to be tested?	Were there any other important flaws in the design or statistical methods of the study?	Overall construct (convergent) validity	Were subjects stable in the interim period on the construct?	Was the time interval appropriate?	Were the test conditions similar for the measurements? (administration, environment, instructions)	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	For dichotomous/ordinal scores: Was kappa calculated?	For ordinal scores: Was a weighted kappa calculated?	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Were there any other important flaws in the design or statistical methods of the study?	Overall reliability quality score
Bacardi-Gascón et al.	2012	Very good	inadequate	Very good	doubtful	Inadequate	Adequate	Very good	Adequate	Doubtful	Inadequate	N/A	N/A	Very good	Inadequate
Dwyer et al.	2011	Very good	doubtful	Very good	Very good	Doubtful	Adequate	Very good	Adequate	Adequate	Adequate	N/A	N/A	Very good	Adequate
Janz et al.	2005	Very good	Very good	Very good	inadequate	Inadequate	Inadequate	Inadequate	Doubtful	Adequate	Very good	Very good	Very good	Very good	Inadequate
Manios et al.	1998	Very good	inadequate	Very good	Very good	Inadequate	Adequate	Very good	Adequate	N/A	Doubtful	N/A	N/A	Very good	Doubtful
González-Gil et al.	2014					Inadequate	Inadequate	Very good	Adequate	Very good	Doubtful	N/A	N/A	Very good	Inadequate
Corder et al.	2009	Very good	inadequate	Very good	inadequate	Inadequate	Adequate	Doubtful	Very good	Adequate	N/A	N/A	N/A	Very good	Doubtful
Telford et al.	2004	Very good	doubtful	Very good	Very good	Doubtful	Adequate	Very good	Adequate	Adequate	N/A	N/A	N/A	Very good	Adequate
Mendoza et al.	2013	Very good	doubtful	Very good	doubtful	Doubtful	Adequate	Very good	Very good	Adequate	N/A	N/A	N/A	Very good	Adequate
Nishikido et al.	1982	Very good	inadequate	Very good	doubtful	Inadequate									
Wen et al.	2010	Very good	Very good	Very good	doubtful	Doubtful									
Chen et al.	2002	Very good	inadequate	adequate	doubtful	Inadequate									
Chow & Au	2009	Very good	inadequate	Very good	doubtful	Inadequate									
Noland et al.	1990	Very good	Very good	adequate	doubtful	Doubtful									

Appendix 4.1: PRISMA Checklist for Rapid Review

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	292
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	24
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	293-295
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	295-296
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	301
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	297-300
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	299-300
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	302
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	302-303
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	302-303

Section and Topic	Item #	Checklist item	Location where item is reported
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	302-303
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	303-304
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	306-307
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	302
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	302-303
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	306-307
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	306-307
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	N/A
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	N/A
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	303-304
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	303-304
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	308-309

Section and Topic	Item #	Checklist item	Location where item is reported
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Appendix 4.3
Study characteristics	17	Cite each included study and present its characteristics.	310-313
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	314
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	355-357
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	315-353
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	315-353
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	N/A
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N/A
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	314
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	314
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	358-363
	23b	Discuss any limitations of the evidence included in the review.	358-363
	23c	Discuss any limitations of the review processes used.	363-364
	23d	Discuss implications of the results for practice, policy, and future research.	364-365

Section and Topic	Item #	Checklist item	Location where item is reported
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	296
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	296
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N/A
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	14
Competing interests	26	Declare any competing interests of review authors.	N/A
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	N/A

Appendix 4.2: PROPSERO Rapid Review Protocol



PROSPERO Sleep
Measurement Review.

Appendix 4.3: Excluded studies with reasons

Part 1: Databases**Children outside of age range (n=38)****Children <3 years old (n=8)**

1. Belanger ME, Bernier A, Paquet J, Simard V, Carrier J. Validating Actigraphy as a Measure of Sleep for Preschool Children. *Journal of Clinical Sleep Medicine*. 2013;9(7):701-6.
2. Eaton CK, Henning E, Lam J, Paasch V. Actigraphy technology: Informing assessment and intervention for sleep disturbances in young children. *Clinical Practice in Pediatric Psychology*. 2019;7(4):347-57.
3. Goodlin-Jones BL, Sitnick SL, Tang K, Liu J, Anders TF. The Children's Sleep Habits Questionnaire in toddlers and preschool children. *Journal of Developmental & Behavioral Pediatrics*.29(2):82-8.
4. Hatch B, Nordahl CW, Schwichtenberg AJ, Ozonoff S, Miller M. Factor Structure of the Children's Sleep Habits Questionnaire in Young Children with and Without Autism. *Journal of Autism and Developmental Disorders*.12.
5. Henderson JA, Jordan SS. Development and preliminary evaluation of the bedtime routines questionnaire. *Journal of Psychopathology and Behavioral Assessment*. 2010;32(2):271-80.
6. Minde K, Popiel K, Leos N, Falkner S, Parker K, Handleyderry M. THE EVALUATION AND TREATMENT OF SLEEP DISTURBANCES IN YOUNG-CHILDREN. *Journal of Child Psychology and Psychiatry*. 1993;34(4):521-33.
7. Sneddon P, Peacock GG, Crowley SL. Assessment of sleep problems in preschool aged children: an adaptation of the children's sleep habits questionnaire. *Behavioral Sleep Medicine*.11(4):283-96.
8. Steur LMH, Visser EH, Grootenhuis MA, Terwee CB, Kaspers GJL, van Litsenburg RRL. Psychometric properties and Dutch norm values of the Children's Sleep Habits Questionnaire in toddlers. *Sleep Medicine*.34:57-63.

Children >7 years old (n=20)

1. Bevans KB, Meltzer LJ, De La Motte A, Kratchman A, Viel D, Forrest CB. Qualitative Development and Content Validation of the PROMIS Pediatric Sleep Health Items. *Behavioral Sleep Medicine*.17(5):657-71.
2. Blanc F, Merklen F, Blanchet C, Mondain M, Akkari M. Respiratory polygraphy in children: Feasibility in everyday practice in an ENT department and value of automatic detection of respiratory events. *European annals of otorhinolaryngology, head & neck diseases*.136(4):235-40.
3. Certal V, de Lima FF, Winck JC, Azevedo I, Costa-Pereira A. Translation and cross-cultural adaptation of the Pediatric Sleep Questionnaire into Portuguese language. *International Journal of Pediatric Otorhinolaryngology*.79(2):175-8.
4. Dollinger SJ, Schum RL. The Missouri Children's Picture Series: Utility of the sleep disturbance scale with a clinical sample. *Journal of Pediatric Psychology*. 1979;4(2):113-8.

5. Forrest CB, Meltzer LJ, Marcus CL, De La Motte A, Kratchman A, Buysse DJ, et al. Development and validation of the PROMIS Pediatric Sleep Disturbance and Sleep-Related Impairment item banks. *Sleep*. 2018;41(6).
6. Goodwin JL, Enright PL, Kaemingk KL, Rosen GM, Morgan WJ, Fregosi RF, et al. Feasibility of using unattended polysomnography in children for research--report of the Tucson Children's Assessment of Sleep Apnea study (TuCASA). *Sleep*. 24(8):937-44.
7. Hyde M, O'Driscoll DM, Binette S, Galang C, Tan SK, Verginis N, et al. Validation of actigraphy for determining sleep and wake in children with sleep disordered breathing. *Journal of Sleep Research*. 2007;16(2):213-6.
8. Ioan I, Weick D, Schweitzer C, Guyon A, Coutier L, Franco P. Feasibility of parent-attended ambulatory polysomnography in children with suspected obstructive sleep apnea. *Journal of Clinical Sleep Medicine*. 2020;16(7):1013-9.
9. Iwasaki M, Iemura A, Oyama T, Matsuishi T. A Novel Subjective Sleep Assessment Tool for Healthy Elementary School Children in Japan. *Journal of Epidemiology*. 2010;20:S476-S81.
10. Kuwada A, Mohri I, Asano R, Matsuzawa S, Kato-Nishimura K, Hirata I, et al. Japanese Sleep Questionnaire for Elementary Schoolers (JSQ-ES): validation and population-based score distribution. *Sleep Medicine*. 2018;41:69-77.
11. Marcus CL, Traylor J, Biggs SN, Roberts RS, Nixon GM, Narang I, et al. Feasibility of comprehensive, unattended ambulatory polysomnography in school-aged children. *Journal of Clinical Sleep Medicine*. 10(8):913-8.
12. Owens JA, Spirito A, McGuinn M. The Children's Sleep Habits Questionnaire (CSHQ): psychometric properties of a survey instrument for school-aged children. *Sleep*. 23(8):1043-51.
13. Perdahli Fiş N, Arman A, Ay P, Topuzoğlu A, Güler AS, Gökçe Imren S, et al. Çocuk uyku alışkanlıkları anketinin Türkçe geçerliliği ve güvenilirliği = The validity and the reliability of Turkish Version of Children's Sleep Habits Questionnaire. *Anadolu Psikiyatri Dergisi*. 2010;11(2):151-60.
14. Schlarb AA, Schwerdtle B, Hautzinger M. Validation and psychometric properties of the German version of the Children's Sleep Habits Questionnaire (CSHQ-DE). *Somnologie*. 2010;14(4):260-6.
15. Smith C, Galland B, Taylor R, Meredith-Jones K. ActiGraph GT3X+ and Actical Wrist and Hip Worn Accelerometers for Sleep and Wake Indices in Young Children Using an Automated Algorithm: Validation With Polysomnography. *Frontiers in Psychiatry*. 2020;10.
16. Spruyt K, Gozal D, Dayyat E, Roman A, Molfese D. Sleep assessments in healthy school-aged children using actigraphy: Concordance with polysomnography. *Journal of Sleep Research*. 2011;20(1):223-32.
17. Tamura T, Togawa T, Murata M. A bed temperature monitoring system for assessing body movement during sleep. *Clinical Physics & Physiological Measurement*. 9(2):139-45.
18. Waumans RC, Terwee CB, Van den Berg G, Knol DL, Van Litsenburg RR, Gemke RJ. Sleep and sleep disturbance in children: Reliability and validity of the Dutch version of the Child Sleep Habits Questionnaire. *Sleep*. 33(6):841-5.

19. Werner H, LeBourgeois MK, Geiger A, Jenni OG. Assessment of chronotype in four- to eleven-year-old children: Reliability and validity of the Children's ChronoType Questionnaire (CCTQ). *Chronobiology International*. 2009;26(5):992-1014.
20. Yosae S, Gharamaleki AS, Zamani A, Khosravi A, Djafarian K. Validation of self-reported sleep against actigraphy. *Journal of Zanjan University of Medical Sciences and Health Services*. 2013;21(88).

Children <3 years and >7 years (n=8)

1. Bastida-Pozuelo MF, Sánchez-Ortuño MM. Preliminary analysis of the concurrent validity of the Spanish translation of the BEARS sleep screening tool for children. *Journal of Psychiatric and Mental Health Nursing*. 2016;23(8):513-20.
2. Bonuck KA, Goodlin-Jones BL, Schechter C, Owens J. Modified Children's sleep habits questionnaire for behavioral sleep problems: A validation study. *Sleep Health*. 2017;3(3):136-41.
3. Lebourgeois MK. Validation of the Children's Sleep-Wake Scale: ProQuest Information & Learning; 2003.
4. LeBourgeois MK, Harsh JR. Development and psychometric evaluation of the Children's Sleep-Wake Scale. *Sleep Health*. 2(3):198-204.
5. McGreavey JA, Donnan PT, Pagliari HC, Sullivan FM. The Tayside children's sleep questionnaire: A simple tool to evaluate sleep problems in young children. *Child: Care, Health and Development*. 2005;31(5):539-44.
6. Rosen D. Many parents report their child's breathing and sleep patterns during overnight sleep study as atypical. *Clinical Pediatrics*. 2010;49(8):764-7.
7. Shirouzu S, Seno Y, Tobioka K, Yagi T, Takahashi T, Sasaki M, et al. For Children's Sleep Assessment: Can we trace the change of sleep depth based on ECG data measured at their respective home with a wearable device? 2016 2016. 208-11 p.
8. Wu XM, Bennett DH, Lee K, Cassady DL, Ritz B, Hertz-Picciotto I. Feasibility of using web surveys to collect time-activity data. *Journal of Exposure Science and Environmental Epidemiology*. 2012;22(2):116-25.

Children outside age range / sleep disordered breathing- unable to distinguish the pre-school aged children without SDB (n=2)

1. Meltzer LJ, Walsh CM, Traylor J, Westin AM. Direct comparison of two new actigraphs and polysomnography in children and adolescents. *Sleep*. 35(1):159-66.
2. Meltzer LJ, Walsh CM, Peightal AA. Comparison of actigraphy immobility rules with polysomnographic sleep onset latency in children and adolescents. *Sleep and Breathing*. 2015;19(4):1415-23.

Children with disability/limiting illness (n=1)

1. Lesser AD, Luczynski KC, Hood SA. Evaluating motion detection to score sleep disturbance for children: A translational approach to developing a measurement system. *Journal of Applied Behavior Analysis*. 2019;52(2):580-99.

Purpose to determine sleep disturbances/diagnoses (n=4)

1. Poels PJP, Schilder AGM, Van Den Berg S, Hoes AW, Joosten KFM. Evaluation of a New Device for Home Cardiorespiratory Recording in Children. *Archives of Otolaryngology - Head and Neck Surgery*. 2003;129(12):1281-4.
2. Romeo DM, Bruni O, Brogna C, Ferri R, Galluccio C, De Clemente V, et al. Application of the sleep disturbance scale for children (SDSC) in preschool age. *European Journal of Paediatric Neurology*.17(4):374-82.
3. Shimizu S, Kato-Nishimura K, Mohri I, Kagitani-Shimono K, Tachibana M, Ohno Y, et al. Psychometric properties and population-based score distributions of the Japanese Sleep Questionnaire for Preschoolers. *Sleep Medicine*.15(4):451-8.
4. Gill AI, Schaughency E, Gray A, Galland BC. Reliability of home-based physiological sleep measurements in snoring and non-snoring 3-year olds. *Sleep & Breathing*.17(1):147-56.

Did not explore utility of tool (n=3)

1. Jafar NK, Tham EKH, Eng DZH, Goh DYT, Teoh OH, Lee YS, et al. The association between chronotype and sleep problems in preschool children. *Sleep Medicine*. 2017;30:240-4.
2. Lo MJ. Relationship between sleep habits and nighttime sleep among healthy preschool children in taiwan. *Annals of the Academy of Medicine Singapore*. 2016;45(12):549-56.
3. Tey C, Wake M, Campbell M, Hampton A, Williams J. The Light Time-Use Diary and preschool activity patterns: exploratory study. *International Journal of Pediatric Obesity*.2(3):167-73.

Normative sleep data parameters/ sleep wake patterns (n=3)

1. Scholle S, Beyer U, Bernhard M, Eichholz S, Erler T, Graness P, et al. Normative values of polysomnographic parameters in childhood and adolescence: quantitative sleep parameters. *Sleep Medicine*.12(6):542-9.
2. Scholle S, Wiater A, Scholle HC. Normative values of polysomnographic parameters in childhood and adolescence: cardiorespiratory parameters. *Sleep Medicine*.12(10):988-96.
3. Acebo C, Sadeh A, Seifer R, Tzischinsky O, Hafer A, Carskadon MA. Sleep/wake patterns derived from activity monitoring and maternal report for healthy 1-to 5-year-old children. *Sleep*. 2005;28(12):1568-77.

Not measuring sleep (n=1)

1. Chuang SY, Teng A, Butler JE, Gandevia SC, Selvadurai H, Jaffe A. Validation of a quantitative method to measure neural respiratory drive in children during sleep. *Respiratory Physiology & Neurobiology*.239:75-80.

Parent/child interactions rather than sleep duration/quality (n=3)

1. Alfano CA, Smith VC, Reynolds KC, Reddy R, Dougherty LR. The Parent-Child Sleep Interactions Scale (PSIS) for Preschoolers: Factor Structure and Initial Psychometric Properties. *Journal of Clinical Sleep Medicine*. 2013;9(11):1153-60.
2. Barrios CS, Jay SY, Smith VC, Alfano CA, Dougherty LR. Stability and Predictive Validity of the Parent-Child Sleep Interactions Scale: A Longitudinal Study Among Preschoolers. *Journal of Clinical Child & Adolescent Psychology*.47(3):382-96.
3. Smith VC, Leppert KA, Alfano CA, Dougherty LR. Construct validity of the parent-child sleep interactions scale (PSIS): associations with parenting, family stress, and maternal and child psychopathology. *Sleep Medicine*.15(8):942-51.

Number of days monitoring – measuring stability of the behaviour rather than measurement properties of the tool (n=2)

1. Acebo C, Sadeh A, Seifer R, Tzischinsky O, Wolfson AR, Hafer A, et al. Estimating sleep patterns with activity monitoring in children and adolescents: how many nights are necessary for reliable measures? *Sleep*.22(1):95-103.
2. Taylor RW, Williams SM, Farmer VL, Taylor BJ. The stability of sleep patterns in children 3 to 7 years of age. *Journal of Pediatrics*.166(3):697-702.e1.

Abstract with insufficient information available (n=5)

1. Appleyard K, Galland B, Gill A, Lobb C, Schaughency E. An objective assessment of the reliability of parental reports of snoring in preschool children. *Journal of Sleep Research*. 2011;20:17-8.
2. DiLeo HA, Umlauf MG, Makris CM, Orji C. Assessment of the reliability and validity of one brand of actigraphy device in young children. *Sleep*. 2005;28:A318-A.
3. Gomes AM, Figueiredo S, Rocha C, Ferreira E, Silva B, Caldas AC, et al. SLEEP QUALITY AND CHRONOTYPE OF PORTUGUESE SCHOOL-AGED CHILDREN. 2017 2017. 1-16 p.
4. Mellecker RR, McManus AM, Matsuzaka A. Validity and Reliability of the Sedentary Behavior and Sleep Scale (SBSS) in Young Hong Kong Chinese Children. *Asian Journal of Exercise & Sports Science*. 2012;9(1):21-36.
5. Randler C, Truc Y. Adaptation of the Composite Scale of Morningness for parent report and results from kindergarten children. *Swiss Journal of Psychology*. 2014;73(1):35-9.

Part 2: Trial registries

Children <3 years and >7 years (n=1)

[ICTRP Search Portal \(ictrptest.azurewebsites.net\)](https://ictrptest.azurewebsites.net/)

Part 3: Citation searches/reference lists of included studies

Children outside age range

Children > 7 years (n=9)

1. Berger, I., Obeid, J., Timmons, B. W., & DeMatteo, C. (2017). Exploring accelerometer versus self-report sleep assessment in youth with concussion. *Global pediatric health*, 4, 2333794X17745973
2. Colley, R. C., Wong, S. L., Garriguet, D., Janssen, I., Gorber, S. C., & Tremblay, M. S. (2012). Physical activity, sedentary behaviour and sleep in Canadian children: parent-report versus direct measures and relative associations with health risk. *Health reports*, 23(2), A1.;
3. Dayyat E, Spruyt K, Molfese D, Gozal D. Sleep estimates in children: parental versus actigraphic assessments. *Nat Sci Sleep*. (2011) 3:115–23. 10.2147/NSS.S25676;
4. Goodwin, J. L., Silva, G. E., Kaemingk, K. L., Sherrill, D. L., Morgan, W. J., & Quan, S. F. (2007). Comparison between reported and recorded total sleep time and sleep latency in 6-to 11-year-old children: the Tucson Children's Assessment of Sleep Apnea Study (TuCASA). *Sleep and Breathing*, 11(2), 85-92;
5. Holley, S., Hill, C. M., & Stevenson, J. (2010). A comparison of actigraphy and parental report of sleep habits in typically developing children aged 6 to 11 years. *Behavioral sleep medicine*, 8(1), 16-27.

6. Lockley SW, Skene DJ, Arendt J. Comparison between subjective and actigraphic measurement of sleep and sleep rhythms. *Journal of Sleep Research*. 1999;9:175–183;
7. Loureiro, H. C., Pinto, T. R., Pinto, J. C., Pinto, H. R., & Paiva, T. (2013). Validation of the children sleep habits questionnaire and the sleep self report for Portuguese children. *Sleep Science*, 6(4), 151-158
8. Markovich, A. N., Gendron, M. A., & Corkum, P. V. (2015). Validating the Children's Sleep Habits Questionnaire against polysomnography and actigraphy in school-aged children. *Frontiers in psychiatry*, 5, 188.
9. Mazza, S., Bastuji, H., & Rey, A. E. (2020). Objective and Subjective Assessments of Sleep in Children: Comparison of Actigraphy, Sleep Diary Completed by Children and Parents' Estimation. *Frontiers in Psychiatry*, 11.

Children <3 years (n=4)

1. Bélanger, M. È., Bernier, A., Paquet, J., Simard, V., & Carrier, J. (2013). Validating actigraphy as a measure of sleep for preschool children. *Journal of Clinical Sleep Medicine*, 9(7), 701-706
2. Covington, L. B., Rogers, V. E., & Black, M. M. (2019). Sleep Measurement in Toddlers From Low-income Families. *Journal of Developmental & Behavioral Pediatrics*, 40(3), 219-223.
3. Liu Z, Wang G, Tang H, Wen F, Li N. Reliability and validity of the Children's sleep habits questionnaire in preschool-aged Chinese children. *Sleep Biol Rhythms*. 2014; 12(3):187–193
4. Ward T, Gay C, Anders TF, Alkon A, Lee KA. Sleep and napping patterns in 3- to 5-year-old children attending full-day childcare centers. *Journal of Pediatric Psychology*. 2007;33:666–672

Children with clinically diagnosed disturbances/disorders (n=3)

1. Sadeh, A., Lavie, P., Scher, A., Tirosh, E., & Epstein, R. (1991). Actigraphic home-monitoring sleep-disturbed and control infants and young children: a new method for pediatric assessment of sleep-wake patterns. *Pediatrics*, 87(4), 494-499.
2. Toon, E., Davey, M. J., Hollis, S. L., Nixon, G. M., Horne, R. S., & Biggs, S. N. (2016). Comparison of commercial wrist-based and smartphone accelerometers, actigraphy, and PSG in a clinical cohort of children and adolescents. *Journal of Clinical Sleep Medicine*, 12(3), 343-350.
3. Osterbauer, B., Koempel, J. A., Ward, S. L. D., Fisher, L. M., & Don, D. M. (2016). A comparison study of the Fitbit activity monitor and PSG for assessing sleep patterns and movement in children. *Journal of Otolaryngology Advances*, 1(3), 24.

Did not explore utility of tool (n=6)

1. Belísio, A. S., Kolodiuk, F. F., Louzada, F. M., Valdez, P., & Azevedo, C. V. (2017). Sleep–Wake Cycle, Daytime Sleepiness, and Attention Components in Children Attending Preschool in the Morning and Afternoon Shifts. *Mind, Brain, and Education*, 11(1), 10-20.
2. Iwata, S., Iwata, O., Iemura, A., Iwasaki, M., & Matsuishi, T. (2012). Sleep architecture in healthy 5-year-old preschool children: associations between sleep schedule and quality variables. *Acta Paediatrica*, 101(3), e110-e114.;
3. Sadeh, A. (2008). Commentary: Comparing actigraphy and parental report as measures of children's sleep. *Journal of pediatric psychology*, 33(4), 406-407
4. Silva E, Simões P, Macedo M, Duarte J, Silva D. Parents' perception of the sleep habits and quality of preschool-aged children. *Rev Enfermagem Ref*. (2018) IV:63–70. 10.12707/RIV17103;
5. Tininenko, J. R. (2008). Actigraphic evaluation of sleep disturbance in young children (Doctoral dissertation, University of Oregon).

6. Touchette, E., Mongrain, V., Petit, D., Tremblay, R. E., & Montplaisir, J. Y. (2008). Development of sleep-wake schedules during childhood and relationship with sleep duration. *Archives of pediatrics & adolescent medicine*, 162(4), 343-349

Insufficient information available (n=2)

Lotjonen, J. M. (2001). Sleep/wake detection using an active security device. *Sleep*, 24, S399.
Doi, Y., Ishihara, K., Uchiyama M., Takimoto, H. (2013). Development of the Japanese Version of Children's Chronotype Questionnaire. *Japan Journal Sleep Medicine*. 7:486-93 (In Japanese).

Part 4: Journal articles

Did not explore utility of tool (n=1)

Lionetti, F., Dellagiulia, A., Verderame, C., Sperati, A., Bodale, G., Spinelli, M., & Fasolo, M. (2021). The Children's Sleep Habits Questionnaire: identification of sleep dimensions, normative values, and associations with behavioral problems in Italian preschoolers. *Sleep Health*.

Children <3 years (n=1)

Olivia Cicalese, Ariel Williamson, Jodi Mindell, Lisa Meltzer, 592 Objective vs. Subjective Sleep Data in Early Childhood: Implications for Health Disparities Research, *Sleep*, Volume 44, Issue Supplement_2, May 2021, Page A233, <https://doi.org/10.1093/sleep/zsab072.590>

Appendix 4.4: Risk of bias assessment of included studies

Study	Structural Validity (Box 3)					Internal consistency (box 4)							Reliability (test-retest) (Box 6)							Criterion Validity (Box 8)					Convergent Validity (Box 9)					
	For CTT: Was exploratory or confirmatory factor analysis performed?	For IRT/Rasch: does the chosen model fit to the research?	Was the sample size included in the analysis adequate?	Were there any other important flaws in the design or statistical methods of the study?	Overall construct (structural) validity quality assessment?	Was an internal consistency statistic calculated for each unidimensional scale or subscale separately?	For continuous scores: Was Cronbach's alpha or omega calculated?	For dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	For IRT-based scores: Was standard error of the theta (SE (θ)) or reliability coefficient of estimate of latent trait calculated (index of subject separation)?	Were there any other important flaws in the design or statistical methods of the study?	Overall internal consistency quality assessment?	Were subjects stable in the interim period on the construct?	Was the time interval appropriate?	Were the test conditions similar for the measurements? (administration, environment, instructions)	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	For dichotomous/nominal scores: Was kappa calculated?	For ordinal scores: Was a weighted kappa calculated?	For ordinal scores: Was the weighting scheme described? e.g., linear, quadratic?	Were there any other important flaws in the design or statistical methods of the study?	Overall reliability quality assessment?	For continuous variables, were correlations or AUC-ROC calculated?	For dichotomous scores, were sensitivity and specificity determined?	Were there any other important flaws in the design or statistical methods of the study?	Overall criterion validity quality assessment?	Is it clear what the comparison instrument measures?	Were the measurement properties of the comparison instrument sufficient?	Was the statistical method appropriate for the hypothesis tested?	Were there any other important flaws in the design or statistical methods of the study?	Overall construct (convergent) validity quality assessment?	
Duraccio et al., 2017 [34]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Inadequate	Very good	Doubtful	Inadequate	
Galland et al., 2016 [59]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Doubtful	Very good	Doubtful	Doubtful
Ishihara et al., 2014 [52]	N/A	N/A	N/A	N/A	N/A	Very good	N/A	Very good	N/A	Very good	Very good	Adequate	Very good	Very good	N/A	Inadequate	Inadequate	Adequate	Very good	Inadequate	N/A	N/A	N/A	N/A	Very good	Inadequate	Very good	Very good	Inadequate	
Iwasaki et al., 2010 [35]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Inadequate	Very good	Doubtful	Inadequate
Jones and Ball, 2014 [60]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Doubtful	Very good	Doubtful	Doubtful
Kushnir and Sadeh, 2013 [53]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Inadequate	Very good	Doubtful	Inadequate
Lam et al., 2011 [57]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Inadequate	Doubtful	Very good	Inadequate
Lucas de la Cruz et al., 2016 [54]	Adequate	N/A	Very good	Doubtful	Doubtful	Very good	N/A	Very good	N/A	Very good	Very good	Inadequate	Adequate	Very good	N/A	Inadequate	Inadequate	Adequate	Very good	Inadequate	N/A	N/A	N/A	N/A	Very good	Inadequate	Very good	Doubtful	Inadequate	
Meltzer et al., 2015 [51]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Inadequate	N/A	Very good	Inadequate	N/A	N/A	N/A	N/A	N/A	N/A
Perpétuo et al., 2020 [55]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Inadequate	Very good	Doubtful	Inadequate
Sekine et al., 2002 [36]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Doubtful	Very good	Very good	Doubtful
Staples et al., 2019 [58]	Adequate	N/A	Adequate	Very good	Adequate	Very good	Very good	N/A	N/A	Very good	Very good	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Tan et al., 2018 [56]	Very good	N/A	Adequate	Adequate	Adequate	Very good	N/A	Very good	N/A	Very good	Very good	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Tikotzky and Sadeh, 2001 [62]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Very good	Very good	Very good	Very good	
Tracy et al., 2021 [61]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Doubtful	Doubtful	Doubtful	Doubtful	
Werner et al., 2008 [37]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Very good	Inadequate	Very good	Doubtful	Inadequate	

Appendix 5.1: Recruitment poster

Would you like to be involved with research on the movement-related behaviours (physical activity, sitting behaviours and sleep) of your pre-school child?



If you are a parent or carer of a pre-school child aged 3 or 4 years, we want to hear from you!

What is the purpose of the research?
To explore your opinions on the different ways of measuring physical activity, sitting behaviours and sleep in your pre-school child.

Why is this research important?
To identify the best ways of measuring the movement-related behaviours of children, so that associations between these behaviours and health and development can be better understood.

What will the research consist of?  Approx 1.5 hours

```

graph LR
    A[Provide demographic information] --> B[Workshop - A demonstration and explanation of various measurement instruments]
    B --> C[Opportunity to trial out some of the different measurement instruments]
    C --> D[Focus Group Discussion (6-8 parents/carers per session, and children may be present)]
  
```

If you are interested in hearing more about, or taking part in the research, please contact me:

 **Sophie Phillips**, sophie.m.phillips@durham.ac.uk

 **Durham University**  **fuse**  **NIHR** | School for Public Health Research

This project is funded by/ supported by the National Institute for Health Research (NIHR) School for Public Health Research (Grant Reference Number PD-SPH-2015-10025). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

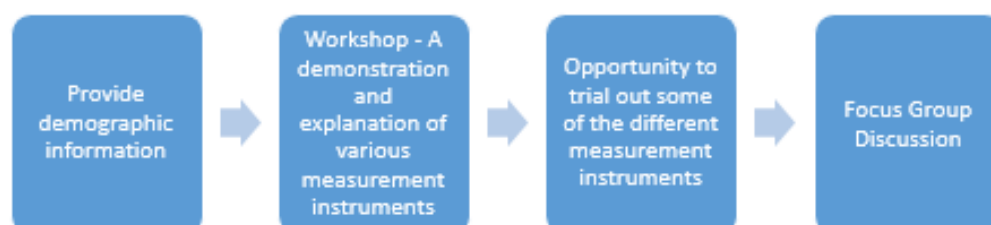
Appendix 5.2: Parent/Carer Recruitment Letter

Dear parent or carer,



I am Sophie Phillips, a PhD student at Durham University. I am writing to invite you to take part in a research project that I am conducting exploring opinions on various measurement tools used to assess the physical activity, sitting time and sleep of pre-school children. We are conducting this research to identify the best ways to measure the movement-related behaviours, so that associations between these behaviours and health can be better understood. You have been invited to participate as you are the parent or carer of a child aged 3 or 4 years old at {name of school/centre}.

What will the research consist of?



If you decide to participate in the study, you will be invited to {name of school or centre}, that your child attends. The whole session will take approximately 1 hour and 30 minutes and will include the presence of between 6 and 8 parents or carers of pre-school children. Your children may also be present during this time. You are able to leave the room at any time during the focus group session, and you do not have to answer specific questions if you do not wish. This discussion would be recorded on a voice-recorder, however, you would remain anonymous by being asked to provide a pseudonym (false name) of which you would be referred to throughout the whole of this recording, and in any subsequent write up of the research.

I have attached an information sheet that will provide further details about what the study would entail.

Remember, this is completely voluntary. You can choose to be in the study or not. If you'd like to participate or have any questions about the research, please feel free to contact me on: sophie.m.phillips@durham.ac.uk. At this point we can discuss dates and times that would be convenient for the session to take place. Once a suitable number of parents or carers have responded, I will identify which dates and times are most convenient for all and circulate this information using the contact details that you have provided me with.

I very much look forward to hearing from you.

Yours sincerely,

Sophie Phillips (Lead Researcher & PhD Student, Department of Sport and Exercise Science, Durham University)- sophie.m.phillips@durham.ac.uk

Supervised by Dr Frances Hillier-Brown (Teaching Fellow- Department of Sport and Exercise Science, Durham University)- frances.hillier-brown@durham.ac.uk

Appendix 5.3: Information sheet

Participant Information Sheet



Exploring the opinions of parents and carers of measurement tools used to assess the movement-related behaviours (physical activity, sedentary behaviour and sleep) of pre-school children.

Who is conducting this research?

I am Sophie Phillips, a PhD student at Durham University. I am conducting research exploring the opinions of parents and carers on various measurement tools used to assess the physical activity, sitting time and sleep of pre-school children, as part of my PhD research. Before you decide whether to agree to take part **it is important for you to understand the purpose of the research and what is involved as a participant.** Please read the following information carefully. Please note ethical approval has been obtained from the Department of Sport and Exercise Science at Durham University.

Why have I been invited to take part?

You have been invited because you are a parent or carer of a pre-school child (child aged 3-4 years), and we are interested in your views regarding ways of measuring the behaviours (physical activity, sitting time and sleep) of children of this age.

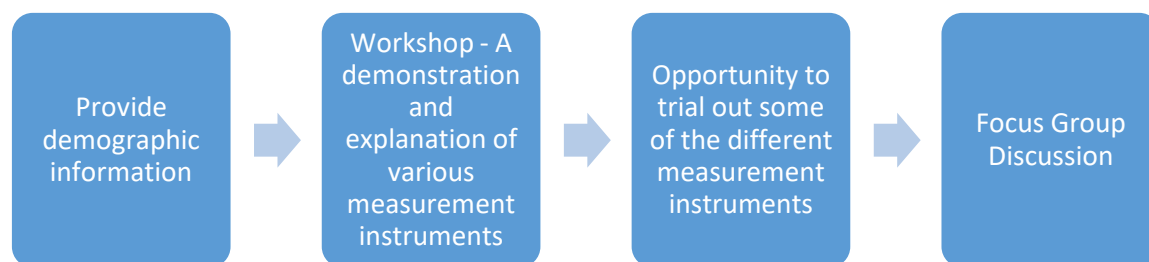
What is the purpose of the study?

This study is being conducted in order to explore the opinions of parents and carers of pre-school children on the various measurement tools that are used to assess physical activity, sitting behaviours and sleep of pre-school children. This will help further research in this area, so that associations between these behaviours and health and development can be better understood. There are two primary aims of the research:

1. To devise and select survey-based measures suitable for assessing the movement-related behaviours of pre-school children.
2. Determine which measurement instruments used to assess the physical activity, sitting behaviours and sleep of pre-school children will be most appropriate in subsequent studies as part of this PhD project, and potentially in wider research.

Do I have to take part?

Your participation is completely voluntary and you do not have to agree to take part. During the session, you do not have to answer any questions you do not feel comfortable with, and you can leave or take a break from the focus group whenever you wish. Additionally, if you agree to take part, you can withdraw at any time, without giving a reason. You are also able to withdraw your responses for up to 7 days following the focus group discussion by contacting the researcher on the email address provided with your pseudonym.

What will happen to me if I take part?

The session outline is detailed above. The whole session will take approximately 1.5 hours, at {name of children's centre/school}. The whole session will include the presence of between 6 and 8 parents or carers of pre-school children. Your children may also be present during this time. You are able to leave the room at any time during the focus group session, and you do not have to answer specific questions if you do not wish. The focus group discussion will be recorded using a voice recorder and then transcribed by the researcher following the session. The discussion is being recorded to ensure that we do not miss any information, and so that all opinions are captured.

Will my data be kept confidential?

All information obtained during the study will be kept confidential. The focus group discussion will be recorded using a voice recorder. This will subsequently be transcribed and no identifiable information will be included in the transcript. These files, along with the demographic information, will be transferred to a password protected research laptop, and saved in encrypted files. Once the information has been transferred, the original files will be deleted and shredded. Consent forms will be stored in a locked cabinet in the researcher's office. The raw data will only be seen by the lead researcher (Sophie Phillips) and her supervisory team for purposes of analysing the data.

If the findings of the study are published, it will be entirely anonymous and you will not be able to be identified. It is possible that the outputs of this research may include direct quotes from the focus groups, but your identity will be kept anonymous at all times. You will only ever be referred to by your pseudonym and your name will not be used in any outputs related to this research.

It is important that whatever is said during the focus group does not leave the room, and that you do not share any information or identities of other people whom are taking part in the focus group discussion. Please ensure that you treat all other members of the group with full respect at all times, it is possible that you may disagree with the opinions of other people and these differences can be discussed amicably.

How is this research being funded?

The lead researcher, Sophie Phillips, is supported by the National Institute for Health Research School for Public Health Research (NIHR SPHR). The NIHR SPHR is a partnership between the Universities of Sheffield; Bristol; Cambridge; Imperial; and University College London; The London School for Hygiene and Tropical Medicine (LSHTM); LiLaC – a collaboration between the Universities of Liverpool and Lancaster; and Fuse- The Centre for Translational Research in Public Health a collaboration between Newcastle, Durham, Northumbria, Sunderland and Teesside Universities. The views to be expressed will be those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

We hope that this information sheet has been helpful. However, please feel free to contact us at any time using the details below.

Sophie Phillips (Lead Researcher & PhD Student, Department of Sport and Exercise Science, Durham University) (sophie.m.phillips@durham.ac.uk)

Supervised by **Dr Frances Hillier-Brown** (Teaching fellow, Department of Sport and Exercise Science, Durham University) (frances.hillier-brown@durham.ac.uk)

Appendix 5.4: Privacy notice

Privacy Notice

**PART 1 – GENERIC PRIVACY NOTICE**

Durham University has a responsibility under data protection legislation to provide individuals with information about how we process their personal data. We do this in a number of ways, one of which is the publication of privacy notices. Organisations variously call them a privacy statement, a fair processing notice or a privacy policy.

To ensure that we process your personal data fairly and lawfully we are required to inform you:

- Why we collect your data
- How it will be used
- Who it will be shared with
-

We will also explain what rights you have to control how we use your information and how to inform us about your wishes. Durham University will make the Privacy Notice available via the website and at the point we request personal data.

Our privacy notices comprise two parts – a generic part (ie common to all of our privacy notices) and a part tailored to the specific processing activity being undertaken.

Data Controller

The Data Controller is Durham University. If you would like more information about how the University uses your personal data, please see the University's [Information Governance webpages](#) or contact Information Governance Unit:

Telephone: (0191 33) 46246 or 46103

E-mail: information.governance@durham.ac.uk

Information Governance Unit also coordinate response to individuals asserting their rights under the legislation. Please contact the Unit in the first instance.

Data Protection Officer

The Data Protection Officer is responsible for advising the University on compliance with Data Protection legislation and monitoring its performance against it. If you have any concerns regarding the way in which the University is processing your personal data, please contact the Data Protection Officer:

Jennifer Sewel

University Secretary

Telephone: (0191 33) 46144

E-mail: university.secretary@durham.ac.uk

Your rights in relation to your personal data

Privacy notices and/or consent

You have the right to be provided with information about how and why we process your personal data. Where you have the choice to determine how your personal data will be used, we will ask you for consent. Where you do not have a choice (for example, where we have a legal obligation to process the personal data), we will provide you with a privacy notice. A privacy notice is a verbal or written statement that explains how we use personal data.

Whenever you give your consent for the processing of your personal data, you receive the right to withdraw that consent at any time. Where withdrawal of consent will have an impact on the services we are able to provide, this will be explained to you, so that you can determine whether it is the right decision for you.

Accessing your personal data

You have the right to be told whether we are processing your personal data and, if so, to be given a copy of it. This is known as the right of subject access. You can find out more about this right on the University's [Subject Access Requests webpage](#).

Right to rectification

If you believe that personal data we hold about you is inaccurate, please contact us and we will investigate. You can also request that we complete any incomplete data.

Once we have determined what we are going to do, we will contact you to let you know.

Right to erasure

You can ask us to erase your personal data in any of the following circumstances:

- We no longer need the personal data for the purpose it was originally collected
- You withdraw your consent and there is no other legal basis for the processing
- You object to the processing and there are no overriding legitimate grounds for the processing
- The personal data have been unlawfully processed
- The personal data have to be erased for compliance with a legal obligation
- The personal data have been collected in relation to the offer of information society services (information society services are online services such as banking or social media sites).

Once we have determined whether we will erase the personal data, we will contact you to let you know.

Right to restriction of processing

You can ask us to restrict the processing of your personal data in the following circumstances:

- You believe that the data is inaccurate and you want us to restrict processing until we determine whether it is indeed inaccurate
- The processing is unlawful and you want us to restrict processing rather than erase it
- We no longer need the data for the purpose we originally collected it but you need it in order to establish, exercise or defend a legal claim and
- You have objected to the processing and you want us to restrict processing until we determine whether our legitimate interests in processing the data override your objection.

Once we have determined how we propose to restrict processing of the data, we will contact you to discuss and, where possible, agree this with you.

Retention

The University keeps personal data for as long as it is needed for the purpose for which it was originally collected. Most of these time periods are set out in the [University Records Retention Schedule](#).

Making a complaint

If you are unsatisfied with the way in which we process your personal data, we ask that you let us know so that we can try and put things right. If we are not able to resolve issues to your satisfaction, you can refer the matter to the Information Commissioner's Office (ICO). The ICO can be contacted at:

Information Commissioner's Office Wycliffe House Water Lane Wilmslow Cheshire SK9 5AF

Telephone: 0303 123 1113

Website: [Information Commissioner's Office](#)

PART 2 – TAILORED PRIVACY NOTICE

This section of the Privacy Notice provides you with the privacy information that you need to know before you provide personal data to the University for the particular purpose(s) stated below.

Project Title:

Exploring the opinions of parents and carers of measurement tools used to assess the movement-related behaviours (physical activity, sedentary behaviour and sleep) of pre-school children.

Type(s) of personal data collected and held by the researcher and method of collection:

Personal data will be collected through a consent form, demographic questionnaire and focus group (which will be voice recorded).

- The consent form will include the name of the participant.
- The demographic questionnaire will include: age, sex and ethnicity of participant, age of child, parent/carers' educational status and home postcode.
- The focus group discussion will be voice recorded and include views on measurement instruments used to assess physical activity, sitting behaviour and sleep of pre-school children.

Lawful Basis

The present research is taking place under the 'University's public task', whereby collection and use of personal data is carried out under the University's public task, which includes teaching, learning and research.

- For further information see <https://durham.ac.uk/research.innovation/governance/ethics/governance/dp/legalbasis/>

How personal data is stored:

All personal data will be held securely and strictly confidential to the research team. All data (both the voice recording from the focus groups and quantitative from the demographic questionnaire) will be inputted and stored electronically on to a Durham University password protected research laptop in password protected files - this electronically stored data will only refer to participants by their pseudonym. The voice recording and paper copies of the demographic information will be deleted and shredded once this information is stored on the laptop. Consent forms will be kept in a locked cabinet in the researcher's office, separate to the project data.

How personal data is processed:

The demographic questionnaire information will be entered into a database for analysis, to contextualise the responses of the focus groups. The recorded focus group discussion will be transcribed by the researcher, and personal information will be anonymised. All information inputted for analysis will refer to participants by their pseudonym and never their real name.

Withdrawal of data

Participants will be able to withdraw participation at any point throughout the focus group discussion and up to 7 days after, by contacting the lead researcher and providing their pseudonym. The 7 day cut off point is due to data analysis starting soon after the focus groups will have taken place, due to the timescales of the project.

Who the researcher shares personal data with:

Data will only be available to the lead researcher and her supervisory team for purposes of analysis of this project.

How long personal data is held by the researcher:

All data will be kept for 10 years, the electronic data will stay stored electronically and manual data will be kept within the university. After 10 years the data will be discarded appropriately (deletion of electronic files or shredding of manual files), unless used for a publishable journal, in which case, the data will be discarded 10 years from publishing.

How to object to the processing of your personal data for this project:

If you have any concerns regarding the processing of your personal data, or you wish to withdraw your data from the project, contact: **Sophie Phillips** (Lead Researcher & PhD Student, Department of Sport and Exercise Science, Durham University) on sophie.m.phillips@durham.ac.uk or her supervisor: **Dr Frances Hillier-Brown** (Teaching Fellow, Department of Sport and Exercise Science, Durham University) on frances.hillier-brown@durham.ac.uk.

Appendix 5.5: Consent form

Participant Number: ____

Consent Form

Exploring the opinions of parents and carers of measurement tools used to assess the movement-related behaviours (physical activity, sitting behaviour and sleep) of pre-school children.



Contact Details:

Researcher(s): Sophie Phillips (sophie.m.phillips@durham.ac.uk); Department of Sport and Exercise Science, Durham University.

Primary Supervisor's name: Dr Frances Hillier-Brown (frances.hillier-brown@durham.ac.uk)

This form is to confirm that you understand the purposes of the project, what is involved and that you are happy to take part. Please initial each box to indicate your agreement:

☐	I understand that this research is being conducted as part of a PhD project at Durham University. I confirm that the task has been explained to me fully, I have read and understand the information sheet and privacy notice for the above project. I have been given the opportunity to ask questions about the research project, and I am satisfied with the answers I have been given.	
☐	I agree to provide some demographic information, and to take part in the focus group, and understand that this will be with the researcher and other participants.	
☐	I give my permission for the focus group to be audio recorded and transcribed.	
☐	I understand that all data (demographic questionnaire and voice recording) will be stored securely, that when the recording has been transcribed it will be destroyed, and that any identifiable information about myself or others will not be included in the transcript.	
☐	I understand that the data obtained from this study, including my words, may be quoted in publications, reports, and other research outputs, but that my name will not be used and my identity will be kept anonymous at all times, in any outputs related to this research.	
☐	I understand that what is discussed in the focus group will be kept confidential by the researcher, but that if the interviewer feels that I or somebody else is at risk of serious harm, they may need to disclose this to relevant agencies.	
☐	I will not discuss the identities of participants or anything they express with anyone else after the focus group is over, and I commit to being respectful towards other focus group members at all times.	
☐	I am aware that the researcher has asked all members of the focus group to commit to not discussing the identities of other participants or anything they say outside of the focus group setting, but understand that this cannot be guaranteed.	
☐	I understand that I have the right to not answer any questions I do not feel comfortable with, and that I can leave or take a break from the focus group at any time.	

Participant Number: __

I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason, and for up to 7 days following the time of data collection.	
I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project.	



Having read the information sheet and consent form, I confirm that I understand what is required of me for this research study and that I am happy to take part.

Participant's Signature_____ Date_____
(NAME IN BLOCK LETTERS)_____
Researcher's Signature_____ Date_____
(NAME IN BLOCK LETTERS)_____

Appendix 5.6: Demographic information questionnaire

Participant Number: ____

Focus Group: Demographic Details Questionnaire

Exploring the opinions of parents and carers of measurement tools used to assess the movement-related behaviours (physical activity, sedentary behaviour and sleep) of pre-school children.



Please answer the following questions in the spaces provided, or tick the most appropriate options.

1. Please provide your age: __
2. Please provide your sex (please tick).
☐ Male ☐ Female
3. Please provide the age of your child (years/months): __ / __
4. Which of the following best describes your ethnicity? (Please tick)
 - ☐ White British
 - ☐ White Irish
 - ☐ Any other White background (Please specify: _____)
 - ☐ Mixed White and Black: Caribbean
 - ☐ Mixed White and Black: African
 - ☐ Mixed White and Asian
 - ☐ Any other Mixed background (Please specify: _____)
 - ☐ Asian or Asian British: Indian
 - ☐ Asian or Asian British: Pakistani
 - ☐ Asian or Asian British: Bangladeshi
 - ☐ Any other Asian background (Please specify: _____)
 - ☐ Black or Black British: Caribbean
 - ☐ Black or Black British: African
 - ☐ Any other Black background (Please specify: _____)
 - ☐ Chinese
 - ☐ Any other ethnic group (Please specify: _____)
 - ☐ Prefer not to say
 - ☐ Don't know

Participant Number: ____

5. What is the highest degree or level of school you have completed? If currently enrolled, please state highest degree received.

- ☐ No formal qualifications
- ☐ GCSEs/ O-levels/ CSEs
- ☐ Vocational qualifications such as Apprenticeships or City and Guilds (= NVQ1+2)
- ☐ A-level, Scottish Higher or equivalent (=NVQ3)
- ☐ Diplomas in higher education, HNC/HND/BTEC Higher or equivalent
- ☐ Bachelor Degree or equivalent (=NVQ4-6)
- ☐ Masters/ PhD or equivalent (=NVQ7+8)
- ☐ Other (Please specify _____)
- ☐ Don't know
- ☐ Prefer not to say

7. Which of the following questions best describes your current employment status?

- ☐ Not working, but actively looking for paid work
- ☐ Not working and not looking for paid work
- ☐ Working part-time including self-employment (up to 15 hours)
- ☐ Working part-time including self-employment (16-30 hours)
- ☐ Working full-time including self-employment (30+ hours)
- ☐ Retired early and claiming benefits (ESA or JSA)
- ☐ Retired early and not claiming benefits
- ☐ Retired and receiving state pension
- ☐ Student
- ☐ Unable to work
- ☐ Looking after the home
- ☐ Other (Please specify _____)
- ☐ Prefer not to say

Participant Number: ____

8. Which of the following groups best represents your household's total take-home income, after tax and other deductions?

	Per week	Per month	Per year
☐	Under £48	Under £208	Under £2,500
☐	£48 - £96	£208 - £416	£2,500 - £4,999
☐	£97 - £192	£417 - £833	£5,000 - £9,999
☐	£193 - £288	£834 - £1,249	£10,000 - £14,999
☐	£289 - £385	£1,250 - £1,667	£15,000 - £19,999
☐	£386 - £481	£1,668 - £2,083	£20,000 - £24,999
☐	£482 - £577	£2,084 - £2,500	£25,000 - £29,999
☐	£578 - £673	£2,501 - £2,917	£30,000 - £34,999
☐	£674 - £769	£2,918 - £3,333	£35,000 - £39,999
☐	£770 - £865	£3,334 - £3,750	£40,000 - £44,999
☐	£866 - £962	£3,751 - £4,167	£45,000 - £49,999
☐	£963 - £1,442	£4,168 - £6,250	£50,000 - £74,999
☐	£1,443 - £1,923	£6,251 - £8,333	£75,000 - £99,999
☐	£1,924 or more	£8,334 or more	£100,000 or more
☐	Nothing		
☐	Don't know		
☐	Prefer not to say		

9. Please state your current home postcode: _____

Please state your chosen pseudonym (False name) for this discussion:

Please provide a pseudonym (false name) that you will be referred to: throughout the session, focus group, in the focus group transcript and would be used in any written report of the work.	

Would you be happy to be contacted regarding any further research in this area? Please circle response: Yes/No

Thank you for taking the time to complete this questionnaire.

Appendix 5.7: Accelerometer information

Actigraph GT3X+ (Actigraph, Pensacola, USA) is a triaxial accelerometer. The size of the Actigraph GT3X+ is 4.6 cm × 3.3 cm × 1.5 cm and weighs 19g.



ActivPAL4 micro (PAL Technologies Ltd., Glasgow, UK) is a uniaxial accelerometer and inclinometer. The size of the ActivPAL is 2.35cm x 4.3cm x 0.5 cm and weighs 9.5g.



Actical (Philips Respironics Inc., Murrysville PA, USA) is an omnidirectional accelerometer. The size of the Actical is 2.8 cm x 2.7cm x 1 cm and weighs 17 g.



Appendix 5.8: Accelerometer instructions

**What is it?**

It is called an 'Actigraph' and it measures the physical movement of your child, including when they are physically active, sitting down or lying down.

How does it work?

It measures how much and how fast your child's body moves up and down. It does not need to be switched on or off as it is already set and ready to go.

Don't worry if it flashes or not it will be working anyhow!

How does my child wear it?

It should be worn around your child's waist for 7 days including a weekend. Try to make sure that the monitor sits on the right side of your child's hip and it is next to their skin and under their clothes.

When does my child wear it?

It is very important for the monitor to be worn 24 hours of the day. The monitor is waterproof but it is best to take off and put it in a safe place if your child has a bath or a shower or goes swimming. Please try to remember to put it back on immediately afterwards, and try to remember to record the time it is taken off and put back on in the table on the next page.

What if I have any questions about the monitor?

If you have any questions, you can contact Sophie Phillips (sophie.m.phillips@durham.ac.uk).





What is it?

It is called an 'Actical' and it measures the physical movement of your child, including when they are physically active, sitting down or lying down.



How does it work?

It measures how much and how fast your child's body moves up and down. It does not need to be switched on or off as it is already set and ready to go.

How does my child wear it?

It should be worn around your child's waist for 7 days including a weekend. Try to make sure that the monitor sits on the right side of your child's hip and it is next to their skin and under their clothes.

When does my child wear it?

It is very important for the monitor to be worn 24 hours of the day. The monitor is waterproof but it is best to take off and put it in a safe place if your child has a bath or a shower or goes swimming. Please try to remember to put it back on immediately afterwards, and try to remember to record the time it is taken off and put back on in the table on the next page.

What if I have any questions about the monitor?

If you have any questions, you can contact Sophie Phillips (sophie.m.phillips@durham.ac.uk).





What is it?

It is called an 'ActivPAL' and it measures the physical movement of your child, including when they are physically active, sitting down or lying down.

How does it work?

It measures how much and how fast your child's body moves up and down. It does not need to be switched on or off as it is already set and ready to go. The monitor will flash green every 6 seconds- this means it's working!

How does my child wear it?

The monitor should be worn on the front of the thigh, roughly 1/3 of the way between the hip and knee with the stick man standing up. The monitor should be kept in the waterproof sleeve, and attached directly onto the skin, and covered with an adhesive patch. It should be worn for 7 days including a weekend.

When does my child wear it?

It is very important for the monitor to be worn 24 hours of the day. The monitor is waterproof so your child can wear it whilst showering and bathing, but it is best to take it off if your child goes swimming. Please try to remember to put it back on immediately afterwards, and try to remember to record the time it is taken off and put back on in the table on the next page.

What happens if I need to change the adhesive patch?

1. Remove the thigh monitor and peel off the adhesive patch.
2. Wipe down the monitor and the area of your thigh with the alcohol prep pad
3. Position the monitor in the same place, ensuring the stick man on the front is standing up (as can be seen in the image on this sheet).|
4. Place the new adhesive patch over the monitor and gently press this down to ensure it is secure.

What if I have any questions about the monitor?

If you have any questions, you can contact Sophie Phillips (sophie.m.phillips@durham.ac.uk)



Appendix 5.9: Accelerometer logs

Accelerometer Log... We need your help!
 We need to know when your child wore the device.
PLEASE COMPLETE THE TABLE BELOW

Today is... (Please write)		Has the device been worn today?		What time did your child get out of bed?	Was the device removed today?			Why was the device removed?	What time did your child go to bed?
		Yes	No	Yes	No	If yes, please provide the time it was taken off and time put back on.			
<i>Example</i>	<i>Monday</i>	✓		<i>8am</i>	✓		<i>Time off: 5.30pm Time on: 6.30pm</i>	<i>Swimming lesson</i>	<i>8pm</i>
Day 1									
Day 2									
Day 3									
Day 4									
Day 5									
Day 6									
Day 7									

Accelerometer Diary... We need your help!

We need to know when your child wore the device.

PLEASE COMPLETE THE TABLE BELOW

Today is... (Please write)		Has the device been worn today?		Was the device removed today?			Why was the device removed?
				Yes	No	If yes, please provide the time it was taken off and time put back on .	
<i>Example</i>	<i>Monday</i>	✓		✓		Time off: 5.30pm Time on: 6.30pm	Swimming lesson
Day 1							
Day 2							
Day 3							
Day 4							
Day 5							
Day 6							
Day 7							

Accelerometer Diary... We need your help!
 We need to know when your child wore the device.
PLEASE COMPLETE THE TABLE BELOW



Date	Time device was removed	Time device was put back on	Reason



Appendix 5.10: Sleep diary

My Pre-School Child's Sleep Diary

Child's Name: _____ Date Starting Diary: __/__/__ Unique Number: ____



		<i>Example</i>	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Day of Week:		<i>Monday</i>							
What time did your child go to bed?		<i>7pm</i>							
What time did your child fall to sleep?		<i>7.30pm</i>							
What time did your child wake up?		<i>6am</i>							
Did your child wake up during the night?	How many times did they wake up during the night?	<i>2</i>							
	How long were they awake for during the night?	<i>30 minutes</i>							
What was the total <u>night time</u> sleep duration of your child?		<i>9.5 hours</i>							

Additional comments about night time sleep		Moved to parents bed during the night							
Did your child nap during the day?	How many naps did your child take?	1							
	On average, how long was your child asleep for during nap?	45 minutes							
What was the total <u>day time</u> sleep duration of your child?		45 minutes							
Additional comments about day time sleep		No comments							

Appendix 5.11: Measurement of Movement in the Early Years Provisional Questionnaire

Measurement of Movement in the Early Years



This questionnaire asks about the movement behaviours of your pre-school child.

There are three parts to this questionnaire:

Part 1: **Physical activity**- Bodily movement, including when your child is engaged in active play.

Part 2: **Sedentary behaviour**- When your child is sitting, reclining or lying.

Part 3: **Sleep**

Please answer all questions as accurately as possible. There are no right or wrong answers and all responses are confidential.



This project is funded by/ supported by the National Institute for Health Research (NIHR) School for Public Health Research (Grant Reference Number PD-SPH-2015-10025). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Part 1- Physical Activity

1a. In the past week, how many days and for how long each day would you say your child has spent actively playing outdoors (e.g. playing in the garden/yard, playing at the park, riding a bike, jumping on trampoline, climbing, skipping, throw/catch, football sessions)?



Please answer including outdoor play at home, at preschool/nursery or part of organised sessions.



How often	For how long each day (Please tick one box or write a time if more than one hour daily)				
Number of days this week.	Up to 15 mins	16-30 minutes	31-45 minutes	46-60 minutes	More than an hour per day. Please estimate time.
____/7					Hours: ____ Mins: ____

1b. Did any of this activity/active play make your child sweat or breathe harder? (Please circle)

Yes	No (Please proceed to Q2)

1c. If yes, how many hours/minutes of this activity/active play made your child sweat or breathe harder?

Hours: ____ Minutes: ____

2a. In the past week, how many days and for how long each day would you say your child has spent actively playing indoors (e.g. messy play, dancing, running around, rough and tumble play, sit and ride/push toys, indoor play area, active computer games such as Wii Sports, gymnastics/swimming classes)?

Please answer including indoor play at home, at preschool/nursery or part of organised sessions.

How often	For how long each day (Please tick one box or write a time if more than one hour daily)				
Number of days this week	Up to 15 mins	16-30 minutes	31-45 minutes	46-60 minutes	More than an hour per day. Please estimate time.
____/7					Hours: ____ Mins: ____

2b. Did any of this activity/active play make your child sweat or breathe harder? (Please circle)

Yes	No (Please proceed to Q3)

2c. If yes, how many hours/minutes of this activity/active play made your child sweat or breathe harder?

Hours: ____ Minutes: ____

3a. In the past week, how many days and for how long each day would you say your child has spent travelling by walking or cycling (e.g. to and from school)?

How often	For how long each day				
	(Please tick one box or write a time if more than one hour daily)				
Number of days this week	Up to 15 mins	16-30 minutes	31-45 minutes	46-60 minutes	More than an hour per day. Please estimate time.
____/7					Hours: ____ Mins: ____

3b. Did any of this time spent travelling make your child sweat or breathe harder? (Please circle)

Yes	No (Please proceed to Q4)

3c. If yes, how many hours/minutes of this time spent travelling made your child sweat or breathe harder?

Hours: ____ Minutes: ____



Part 2- Sedentary Behaviour

4. In the past week, how many days and for how long each day would you say your child has spent: in their pushchair, being carried, sat in the car/on public transport?

How often	For how long each day (Please tick one box or write a time if more than one hour daily)				
Number of days this week	Up to 15 ~~mins~~	16-30 minutes	31-45 minutes	46-60 minutes	More than an hour per day. Please estimate time.
____/7					Hours: ____ ~~Mins~~ : ____

5. In the past week, how many days and for how long each day would you say your child has spent engaged in screen based activities (e.g. watching TV, using a computer, tablet or mobile phone, playing computer games) whilst sitting, reclining or lying?

Please answer including both time at home and at preschool/nursery.

How often	For how long each day (Please tick one box or write a time if more than one hour daily)				
Number of days this week	Up to 15 mins	16-30 minutes	31-45 minutes	46-60 minutes	More than an hour per day. Please estimate time.
____/7					Hours: ____ Mins: ____

6. In the past week, how many days and for how long each day would you say your child has spent engaged in other activities (e.g. reading, storytelling, colouring, playing with toys), whilst sitting, reclining or lying?



Please answer including both time at home and at preschool/nursery.

How often	For how long each day (Please tick one box or write a time if more than one hour daily)				
Number of days this week	Up to 15 mins	16-30 minutes	31-45 minutes	46-60 minutes	More than an hour per day. Please estimate time.
____/7					Hours: ____ Mins: ____

Part 3- Sleep

7a. Please write your child's usual bed time and wake up time.

Bed time: _____ : _____ am/pm

Wake-up time: _____ : _____ am/pm

7b. On which days of the week is this usually the case? (Please circle)

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday

8. On average, from the time that your child goes to bed, how long does it take them to fall asleep?

Minutes: _____



9. On average, how many times and for how long each time does your child wake up during their night time sleep?

How often	For how long each time (Please tick one box or write a time if more than one hour)				
Number of times per night	Up to 15 mins	16-30 minutes	31-45 minutes	46-60 minutes	More than an hour. Please estimate time.
					Hours: ____ Mins: ____

10. On average, how many times and for how long each time does your child nap during the day?

How often	For how long each time (Please tick one box or write a time if more than one hour)				
Number of times per day	Up to 15 mins ~~0000~~	16-30 minutes	31-45 minutes	46-60 minutes	More than an hour. Please estimate time.
					Hours: ____ ~~Mins:~~ ____

Thank you for completing this questionnaire.



MoveMEY
By Sophie Phillips

Appendix 5.12: Focus group discussion guide

Focus Group: Topic/Discussion Guide

Exploring the opinions of parents and carers of measurement tools used to assess the movement-related behaviours (physical activity, sedentary behaviour and sleep) of pre-school children.

**Facilitator's welcome, introduction and instructions to participants**

Welcome and thank you for volunteering to take part in this focus group. I am a PhD student at Durham University interested in improving child health, I am specifically looking at physical activity, sitting behaviours and sleep of young children. You have been asked to participate as you are the experts where your children are concerned, and I am interested in your point of view in regards to some measurement instruments that are used to measure physical activity, sitting time and sleep in the pre-school population. I realise you are busy and I really appreciate your time.

Introduction: This focus group discussion is designed to explore your opinions on various measurement instruments used to assess the physical activity, sedentary (sitting) behaviours and sleep of your pre-school children, to explore the practicality of the various tools and to identify which may be most appropriate to use in the subsequent stages of my research. The focus group discussion will take around 1 hour, and the whole discussion will be taped on the voice-recorder.

Anonymity: Despite being audio recorded, I would like to assure you that the discussion will be anonymous. The files will be stored in an encrypted folder on a password protected research laptop. The voice recordings will be deleted from the device once transcription has been completed. The transcribed notes of the focus group will contain no information that would allow individual subjects to be linked to specific statements. You should try to answer and comment as accurately and truthfully as possible. I and the other focus group participants would appreciate it if you would refrain from discussing the comments of other group members outside the focus group. If there are any questions or discussions that you do not wish to answer or participate in, you do not have to do so; however please try to be as involved as possible.

Ground rules

- The most important rule is that only one person speaks at a time. There may be a temptation to jump in when someone is talking but please wait until they have finished.
- There are no right or wrong answers
- You do not have to speak in any particular order
- When you do have something to say, please do so. There are many of you in the group and it is important that I obtain the views of each of you
- You do not have to agree with the views of other people in the group
- Does anyone have any questions before we begin?

Warm up

- First, I'd like to go round the table and for everyone to introduce themselves, and say something about their child (this can be anything – age, favourite activity, personality).
- Following this, we will look at each measurement instrument individually and have a discussion around the feasibility of the measure. Some broader questions regarding how these instruments might be used in future studies will then be discussed. We will also look at comparisons between some survey based measures to identify which may be the best measure to monitor these behaviours.

Guiding questions

- How practical, acceptable and suitable the method may be for assessing the behaviours of their pre-school child?
- How likely is it that the tool would be completed in line with study protocol?
- Is the tool suitable in different contexts (i.e. if going away for the weekend) or with different routines?
- Would the tool be easy to wear/administer into existing life?
- How convenient or inconvenient would the tool be? Why might that be, and how does it compare to other tools?
- Are there any issues with the understanding, comprehension, readability or clarity in regards to questionnaire based tools?

- Are the instructions that accompanying the tools clear and easy to understand? Would it be an option for the tools to be administered remotely?
- Are there any specific problems you envisage occurring with the use of the tool for measuring the behaviours in your pre-school child?
- Are there any suggestions for improvements or amendments that could be made to the tools that would make them a more applicable measure?

Comparison between the survey-based methods

- Are there any issues with the understanding, comprehension, readability or clarity in regards to questionnaire-based tools? Is it unclear on any of the questions what behaviours are being targeted?
- Which would be easier to respond to?
- Are there any suggestions for improvements or amendments that could be made to the tools that would make them a more applicable measure?
- Do you think the questions are relevant to the constructs being explored? If not, what would make them more relevant?
- Do you think the questions are presented in a suitable sequence?

Additional questions...

- Would feedback on the behaviours of their child be something desired when taking part in this kind of research?
- How would you feel about the height, weight and waist circumference of your child been taken?

Concluding questions

1. Of all the measurement instruments we have explored and discussed today, are there any particular measurement tools that you think would be best to measure: 1) physical activity; 2) sedentary behaviour; 3) sleep of your pre-school child? Why?
2. Of all the measurement tools, we have explored and discussed today, are there any particular measurement tool that you think would not be suitable to measure: 1) physical activity; 2) sedentary behaviour; 3) sleep of your pre-school child? Why?

3. If you had to choose one method for assessing your child's: 1) physical activity; 2) sedentary behaviour; 3) sleep of the measurement instruments we have explored and discussed today, which measures would you choose? Why?

Conclusion

- Thank you for participating. This has been a very successful discussion. We hope you have found the discussion interesting, and your opinions will be a valuable asset to the research.
- If there is anything you are unhappy with or wish to complain about, please contact me using the email address provided on the debrief sheet.
- I would like to remind you that any comments featuring in this report will be completely anonymous. You have the right to withdraw your responses for up to 7 days after the focus group has been completed.

Appendix 5.13: Debrief

Debriefing Sheet



Exploring the opinions of parents and carers of measurement tools used to assess the movement-related behaviours (physical activity, sedentary behaviour and sleep) of pre-school children.

Thank you very much for taking part in this study. This research was conducted to help devise and select survey-based measures that are deemed suitable to assess the movement-related behaviours of pre-school children. In addition to this, the research will help determine which measurement instruments used to assess the physical activity, sitting behaviours and sleep of pre-school children will be most appropriate in subsequent studies as part of this PhD project, and potentially in wider research. The focus group discussion as part of this research will be used in order to answer these research questions. The demographic information will be reported to give context to the responses, and the group of people who the information came from.

The data you have provided is automatically anonymised (by the use of the pseudonyms), and cannot be traced back to your identity. Any outputs as a result of this research will be completely anonymised, and your individual raw data will not be available to anyone outside of the research team. Please remember that this focus group discussion was confidential, and to be respectful towards other focus group members at all times, including not disclosing any of the information that was stated during the discussion.

If you wish to withdraw your data then you are able to do so by contacting the lead researcher (Sophie Phillips), using the contact details provided. You are able to withdraw your responses, without reason, for up to 7 days from today by stating your pseudonym.

If you would like further information about the study or would like to know the findings when all the data has been collected and analysed, then please contact me using the contact details provided. The findings of all participants will be collated, and as such, it will not be possible to provide individual feedback for this study.

If taking part in this study has raised any specific concerns, or you have any further questions then then please do feel free to contact me.

Contact Details:

Researcher: Sophie Phillips (sophie.m.phillips@durham.ac.uk); Department of Sport and Exercise Science, Durham University

Primary Supervisor's name: Dr Frances Hillier-Brown (frances.hillier-brown@durham.ac.uk)

Appendix 5.14: Risk assessment

POTENTIAL RISKS TO PARTICIPANTS

You should think carefully about the risks that participating in your research poses to participants. Be aware that some subjects can be sensitive for participants even if they are not dealing explicitly with a 'sensitive' topic. Please complete this section as fully as possible and continue on additional pages if necessary.



What risks to participants may arise from participating in your research?	How likely is it that these risks will actually happen?	How much harm would be caused if this risk did occur?	What measures are you putting in place to ensure this does not happen (or that if it does, the impact on participants is reduced)?
1. Despite the research not specifically focusing on obesity, the behaviours are associated with obesity and this may be a sensitive topic for some individuals.	Low risk	May cause worry or concern for participants.	• Signposts to diet and activity recommendations will be provided to individuals if they have any concerns regarding obesity. • The discussions will focus on health, rather than obesity. • Language of the researcher will be used very carefully- with the use of person first language if obesity related conversations arise as part of the discussion.

2. One of the accelerometers is fixed to an elasticated belt. It is possible that individuals may swing these and catch someone with them.	Low risk	May cause mild physical harm.	• Instructions on how to safely use the accelerometer will be provided to participants prior to distributing the monitors.
3.			

POTENTIAL RISKS TO RESEARCHERS

You should think carefully about any hazards or risks to you as a researcher that will be present because of you conducting this research. Please complete this section as fully as possible and continue on additional pages if necessary. Please include an assessment of any health conditions, injuries, allergies or intolerances that may present a risk to you taking part in the proposed research activities (including any related medication used to control these), or any reasonable adjustments that may be required where a disability might otherwise prevent you from participating fully within the research.



1. Where will the research be conducted/what will be the research site?

What hazards or risks to you as a researcher may arise from conducting this research?	How likely is it that these risks will actually happen?	How much harm would be caused if this risk did happen?	What measures are being put in place to ensure this does not happen (or that if it does, the impact on researchers is reduced)?
1. The researcher may be in a vulnerable position due to the research taking place in children's centres and schools in deprived communities in the United Kingdom, whereby children under 5 will be present.	Low risk	The researcher could be in a vulnerable position, however this is perceived to be very low risk due to provisions being put in place prior to the research commencing.	- The lead researcher will provide supervisors with detailed information on the dates, time and location of data collection to ensure safety of the researcher. - The lead researcher has completed a DBS check and any contact with the children will be in the presence of their parent, carer or a member of the pre-school staff, to protect the researcher.

Appendix 6.1: COSMIN guidelines

Standard/requirements to meet 'very good' standard	How this study achieved the standard
1a	
1. Is a clear description provided of the construct to be measured?	Yes, see section 6.4.1 of this thesis.
2. Is the origin of the construct clear: was a theory, conceptual framework or disease model used or clear rationale provided to define the construct to be measured?	Yes, see section 6.4.1 of this thesis.
3. Is a clear description provided of the target population for which the patient reported outcome measure (PROM) was developed?	Yes, see section 6.5.1 of this thesis.
4. Is a clear description provided of the context of use?	Yes, see section 6.6.1 of this thesis.
5. Was the PROM development study performed in a sample representing the target population for which the PROM was developed?	Yes, see section 6.5.1.2 of this thesis.
6. Was an appropriate qualitative data collection method used to identify relevant items for a new PROM?	Yes, focus groups were used.
7. Were skilled group moderators/ interviewers used?	Yes, the moderator attended focus group specific training to be prepared for the sessions.
8. Were the group meetings or interviews based on an appropriate topic or interview guide?	Yes, see topic guide (appendix 5.12)
9. Were the group meetings or interviews recorded and transcribed verbatim?	Yes, all focus groups were recorded and transcribed verbatim.
10. Was an appropriate approach used to analyse the data?	Yes, thematic analysis was used to analyse the data.
11. Was at least part of the data coded independently.	Only one researcher was involved in analysis due to this being a PhD project.

12. Was data collection continued until saturation was reached?	Sufficient data collected for tool development.
13. For quantitative studies (surveys): was the sample size appropriate?	N/A
Part 1b	
14. Was a cognitive interview study or other pilot test conducted?	Yes, this research involved conducting a pilot study, following participants completing the MoveMEY measurement tool.
15. Was the cognitive interview study or other pilot test performed in a sample representing the target population?	Yes, parents of pre-school aged children (aged 3 or 4 years old) completed the pilot, as this is the target population. Recruitment also targeted areas classified as lower socioeconomic status, based on the index of multiple deprivation.
16. Were participants asked about the comprehensibility of the PROM?	Yes, see interview guide (appendix 6.10)
17. Were all items tested in their final form?	Yes, a public involvement and engagement group (consisting of parents of pre-school children) assessed the final version of the tool.
18. Was an appropriate qualitative method used to assess the comprehensibility of the PROM instructions, items, response options, and recall period?	Yes, interviews were conducted with participants following completion of piloting the measurement tool. Interviews are a widely recognised or qualitative method.
19. Was each item tested in an appropriate number of patients? (≥ 7 for qual studies is very good)	Yes, the tool was tested with 12 parents, and a public involvement and engagement group of 3 people.
20. Were skilled interviewers used?	Yes, interviewer had experience of qualitative data collection.
21. Were the interviews based on an appropriate interview guide?	Yes, Please see interview guide (appendix 6.10). Questions were guided by the standards outlined by COSMIN.
22. Were the interviews recorded and transcribed verbatim?	Yes, all interviews were recorded and transcribed verbatim.
23. Was an appropriate approach used to analyse the data?	Yes, thematic analysis was used to analyse the data.

24. Were at least two researchers involved in the analysis?	Only one researcher was involved in analysis due to this being a PhD project.
25. Were problems regarding the comprehensibility of the PROM instructions, items, response options, and recall period appropriately addressed by adapting the PROM?	Yes, although there were no issues with comprehensibility of the PROM, adaptations were made to the tool based on parental feedback.
26. Were patients asked about the comprehensiveness of the PROM?	Yes, see interview guide (see appendix 6.10).
27. Was the final set of items tested?	Yes, a public involvement and engagement group (consisting of parents of pre-school children) assessed the final version of the tool.
28. Was an appropriate method used for assessing the comprehensiveness of the PROM?	Yes, interviews were conducted with participants following completion of piloting the measurement tool. Interviews are a widely recognised or qualitative method.
29. Was each item tested in an appropriate number of patients? (≥ 7 for qual studies is very good)	Yes, the tool was tested with 12 parents, and a public involvement and engagement group of 3 people.
30. Were skilled interviewers used?	Yes, interviewer had experience of qualitative data collection.
31. Were the interviews based on an appropriate interview guide?	Yes, please see interview guide (appendix 6.10). Questions were guided by the standards outlined by COSMIN.
32. Were the interviews recorded and transcribed verbatim?	Yes, all interviews were recorded and transcribed verbatim.
33. Was an appropriate approach used to analyse the data?	Yes, thematic analysis was used to analyse the data.
34. Were at least two researchers involved in the analysis?	Only one researcher was involved in analysis due to this being a PhD project.
35. Were problems regarding the comprehensiveness of the PROM appropriately addressed by adapting the PROM?	See interview guide (appendix 6.10) for questions to ensure that the comprehensiveness of the PROM instructions, items, response options, and recall period are assessed.
Part 2	

1. Was an appropriate method used to ask patients whether each item is relevant for their experience with the condition?	Yes, interviews were conducted with participants following completion of piloting the measurement tool. Interviews are a widely recognised or qualitative method.
2. Was each item tested in an appropriate number of patients? (≥ 7 for qual studies is very good)	Yes, the tool was tested with 12 parents, and a public involvement and engagement group of 3 people.
3. Were skilled group moderators/interviewers used?	Yes, interviewer had experience of qualitative data collection.
4. Were the group meetings or interviews based on an appropriate topic or interview guide?	Yes, Please see interview guide (appendix 6.10)-. Questions were guided by the standards outlined by COSMIN.
5. Were the group meetings or interviews recorded and transcribed verbatim?	Yes, all interviews were recorded and transcribed verbatim.
6. Was an appropriate approach used to analyse the data?	Yes, thematic analysis was used to analyse the data.
7. Were at least two researchers involved in the analysis?	Only one researcher was involved in analysis due to this being a PhD project.
8. Was an appropriate method used for assessing the comprehensiveness of the PROM?	Yes, interviews were conducted with participants following completion of piloting the measurement tool. Interviews are a widely recognised or qualitative method.
9. Was each item tested in an appropriate number of patients? (≥ 7 for qual studies is very good)	Yes, the tool was tested with 12 parents, and a public involvement and engagement group of 3 people.
10. Were skilled group moderators/interviewers used?	Yes, interviewer had experience of qualitative data collection.
11. Were the group meetings or interviews based on an appropriate topic or interview guide?	Yes, Please see interview guide (appendix 6.10)-. Questions were guided by the standards outlined by COSMIN.
12. Were the group meetings or interviews recorded and transcribed verbatim?	Yes, all interviews were recorded and transcribed verbatim.
13. Was an appropriate approach used to analyse the data?	Yes, thematic analysis was used to analyse the data.
14. Were at least two researchers involved in the analysis?	Only one researcher was involved in analysis due to this being a PhD project.

15. Was an appropriate qualitative method used for assessing the comprehensibility of the PROM instructions, items, response options, and recall period?	Yes, interviews were conducted with participants following completion of piloting the measurement tool. Interviews are a widely recognised or qualitative method.
16. Was each item tested in an appropriate number of patients? (≥ 7 for qual studies is very good)	Yes, the tool was tested with 12 parents, and a public involvement and engagement group of 3 people.
17. Were skilled group moderators/interviewers used?	Yes, interviewer had experience of qualitative data collection.
18. Were the group meetings or interviews based on an appropriate topic or interview guide?	Yes, Please see interview guide (appendix 6.10). Questions were guided by the standards outlined by COSMIN.
19. Were the group meetings or interviews recorded and transcribed verbatim?	Yes, all interviews were recorded and transcribed verbatim.
20. Was an appropriate approach used to analyse the data?	Yes, thematic analysis was used to analyse the data.
21. Were at least two researchers involved in the analysis?	Only one researcher was involved in analysis due to this being a PhD project.
Professionals	
22. Was an appropriate method used to ask professionals whether each item is relevant for the construct of interest?	Yes, qualitative surveys were used and professionals were asked about the relevance of the items.
23. Were professionals from all relevant disciplines included?	Yes, topic expert researchers of movement behaviours of young children were included. This included early years expert committees for movement guidelines in the UK, experts involved in measurement tool development, and experts in child public health.
24. Was each item tested in an appropriate number of professionals? (≥ 7 for qual studies is very good)	Each item was tested by 6 topic expert researchers. Two primary supervisors also commented on the appropriateness of the tool.
25. Was an appropriate approach used to analyse the data?	Yes, thematic analysis was used to analyse the data.
26. Were at least two researchers involved in the analysis?	Only one researcher was involved in analysis due to this being a PhD project.

27. Was an appropriate method used for assessing the comprehensiveness of the PROM?	Yes, qualitative surveys were used and professionals were asked about the relevance of the items
28. Were professionals from all relevant disciplines included?	Yes, topic expert researchers of movement behaviours of young children were included. This included early years expert committees for movement guidelines in the UK, experts involved in measurement tool development, and experts in child public health.
29. Was each item tested in an appropriate number of professionals? (≥ 7 for qual studies is very good)	Each item was tested by 6 topic expert researchers. Two primary supervisors also commented on the appropriateness of the tool.
30. Was an appropriate approach used to analyse the data?	Yes, thematic analysis was used to analyse the data.
31. Were at least two researchers involved in the analysis?	Only one researcher was involved in analysis due to this being a PhD project.

Appendix 6.2 Studies examining movement behaviours of pre-school children at scale

Source <i>Study name if applicable</i>	Measure(s)	Information about measure	Sample (Country, age, n)
Survey based (n=4)			
Hesketh et al. (2017) <i>Melbourne Infant Feeding, Activity and Nutrition Trial Program</i>	Questionnaire (written format – no specific name)	Mothers reported the number of hours and minutes spent engaged in the behaviour (matched to guidelines- tummy time; television viewing; restriction of movement – bouncer or swing, stroller or pram, car seat or capsule, high chair or other chair, playpen, baby carrier or sling; average sleep duration during the night and day) on an average day.	Australia; Mean age = 3.6 months; N=455
Health Survey for England (2015) <i>Health Survey for England</i>	Health Survey for England- Survey (oral)	Questions about physical activity and sedentary behaviour over past 7 days.	England, UK; 2-4 years
Australia (2011/12) <i>Australian Health Survey</i>	Australian Health Survey 2011/12 (Physical activity questions in 2014/15 survey are for individuals 15+ only – no reports for younger age groups)	Average physical activity /Screen time over 7 days prior to interview: - Outdoor physical activity or active play included throwing a ball, walking, running around, free play in playgrounds or sandpits, helping in the garden, or any other outdoor activity that involved movement. - Indoor physical activity or active play included dancing, jumping, 'rough and tumble', tidying up, helping to set the table, or any other indoor activity that involved movement. Answers in hours : 0, 0-1, 1-2, 2-3, 3-4, 4-5, 5-6, 6-7... 12+	Australia; 2-4 years

		- Average daily duration of sedentary screen based activity 7 days prior to interview Answers in hours : 0, 0-1, 1-2, 2-3, 3-4, 4-5, 5-6, 6+	
Chia et al. (2020) N/A	Parents completed an online and anonymous SMALLQ® (Surveillance of digital Media hAbits in early childhood Questionnaire) on children's physical activity, sedentary behaviour and sleep.	1. Screen media environment at home: Information about the type of technological devices that children accessed at home was solicited. These included fixed screens—television, desktop computers and game consoles, and mobile screens—smartphones, tablets and other hand-held devices and technological toys 2. Parent screen media use: Information about parental screen media habits on a typical weekend day was solicited. The choices were segregated as use for work, entertainment, social networking and for personal development. Screen media use meant time spent accessing content transmitted via the internet or any computer or TV network. 3. Screen media use of the child outside of pre-school: Parents were asked to provide details of children's screen media habits outside of the pre-school/kindergarten, on a typical weekend day. This included the frequency, duration and purpose of screen media use (e.g. for learning, for entertainment, for communication) on fixed screens, and on mobile screens. 4. Parent perception of screen media: Parents were asked to indicate their perceived importance of screen media in relation to them and to children using a 5-point Likert scale from 'not important' to 'very important' 5. Parental concern of children screen media use: Parents reported perceived concerns on screen media use of children on a 5-point Likert scale from 'not concerned' to 'seriously concerned'. Options for parents included 'Poor sleep', 'Poor eyesight', 'Lack of physical exercise or play', 'Exposure to inappropriate content', 'Addiction', and 'Lack of parent-child interaction' 6. Parental awareness and enforcement of children screen media use guidelines: Parents indicated whether they were aware of guidelines on screen media use by children and if they enforced those guidelines. The guidelines are (i) limit screen time to 60 min per day (ii) introduce high-quality educational programmes (iii) co-watch or co-play with child when engaged in screen media.	Singapore; 2-4 years; N=2384

		7. Non screen habits: Parents reported on children's physical activity and play on a typical weekend day outside of the preschool. Information on the duration spent on various indoor activities (nonscreen reading and drawing), duration of indoor play and outdoor play, and an estimated apportion of time the child spent on moderate-to-vigorous activities, on a typical weekend day, was obtained. 8. Demographics: The smallQ® also included demographic questions about the parent (e.g. relationship to the child, monthly household income, highest educational attainment, and ethnicity) and other information on the child (e.g., gender, age, height and weight, sleep duration, and sleep quality).	
Accelerometers (n=1)			
Taylor et al. (2018) <i>Prevention of Overweight in Infancy (POI)</i>	Actical (Waist, 24hr, 7 days)- physical activity, sedentary behaviour, sleep	Actical accelerometer- Physical activity, sedentary behaviour, sleep As Actical cutoffs for physical activity intensity categories do not exist that span this full age range (1–5years), we defined sedentary time as 0– 25 counts (per 15s), light physical activity (LPA) as 26–697 counts, and moderate-to-vigorous physical activity (MVPA) as ≥698 counts.	New Zealand; 12-35.9 months; N=568
Accelerometers and Surveys Combined (n=13)			
Accelerometer= Actigraph (n=9)			
Lee et al. (2017) <i>Parents' Role in Establishing healthy Physical activity and Sedentary behaviour habits (PREPS) project</i>	Actigraph Wgt3x-BT (waist, waking hours, 7 days)- physical activity/sedentary behaviour PREPS Questionnaire- sedentary behaviour/ sleep	Actigraph –physical activity/sedentary behaviour PREPS Questionnaire- sedentary behaviour/sleep: Screen time: Parents reported how many hours and minutes their toddlers typically: (1) watched television, videos, or DVDs on a television, computer, or portable device; and (2) played video/computer games on electronic devices (e.g., a learning laptop, LeapFrog Leapster, computer, laptop, tablet, cellphone, the internet, PlayStation, or Xbox) per day during weekdays and weekend days. Time spent restrained: Parents reported the number of days per typical week that they limit the time their child spends being restrained (e.g., stroller, high chair, or car seat) to <1 h at a time.	Canada; 12-23 months; N=151

		<i>Response options ranged from 0 (never) to 7 (daily).</i> Sleep time: Parents reported the time their toddlers usually spent sleeping during the daytime (i.e., nap) and nighttime. <i>The total sleep time in hours was calculated by adding time spent in day- and nighttime sleep.</i>	
Santos et al. (2017) <i>Get Up! Study.</i>	Actigraph GT3X (Waist, 24hr, 7 days)- physical activity/sedentary behaviour/sleep Activity monitor log (for accel) Questionnaire-screen time (specific to study)	Actigraph- physical activity/sedentary behaviour /sleep Questionnaire: Screen time: “For how long does your child use screen entertainment on a typical weekday?” “For how long does your child use screen entertainment on a typical weekend day?” <i>Time reported in both questions was summed as follows: (screen time weekdays * 5 + screen time weekend days *2) / 7 = average screen time per day.</i>	Australia; 12-28 months; N=202
Cliff et al. (2017) <i>Preschool Activity, Technology, Health, Adiposity, Behaviour and Cognition (PATH-ABC) Study.</i>	Actigraph GT3X+ (waist, 24hr 7 day)- physical activity Proxy report-sedentary behaviour/Sleep (Sedentary behaviour questions from	Actigraph- physical activity Proxy report: Screen time: Total amount of time their child spent in separate screen-based behaviours on weekdays (Monday to Friday) and on weekends (Saturday and Sunday). Original items were modified to include mobile devices as follows: television programs/movies/internet clips on traditional devices (e.g., TV/DVD), television programs/ movies/internet clips on other devices (e.g., tablet, DVD in car, computer, laptop, mobile phone), games/apps on portable handheld devices (e.g., tablet, mobile phone, handheld game system), console games (non-active) (e.g., PlayStation, Xbox), and console games (active) (e.g., Wii, Xbox Kinect).	Australia; 3-4 year olds; N=248

	Hinkley et al., 2012 and sleep q's from Children's Sleep Habits Questionnaire: Goodlin-Jones et al., 2008)	Sleep: Parents reported children's usual daily sleep duration.	
De Craemer et al. (2018) <i>The Toybox Study</i>	Actigraph GT1M, GT3X, GT3X+ (right hip, 24hr, 6 days (including 2 weekend days))- physical activity Proxy report - screen time/sleep (Screen time questions from main Toybox study questionnaire – Gonzalez-Gil et al., 2014)	Actigraph - physical activity Proxy report: Screen time: For television viewing, the question was formulated separately for weekdays and weekend days as follows: "About how many hours a day does your child usually watch television (including DVDs and videos) in his/her free time?". <i>Answer possibilities were "never", "less than 30 min/day", "30 min to <1 h/day", "1–2 h/day", "3–4 h/day", "5–6 h/day", "7–8 h/day", "8 h per day", "more than 8 h/day", and "I don't know".</i> For computer use, the question was also formulated separately for weekdays and weekend days as follows: "About how many hours a day does your child use the computer for activities like playing games on a computer, game consoles (e.g., PlayStation, Xbox, GameCube) during leisure time?". <i>Answer possibilities were identical to the television viewing questions. Answer possibilities were recoded into minutes of television viewing and computer playing per day by using the midpoint method</i> <i>-These were then added up to reflect total screen time.</i> Sleep: "How many hours of sleep does your child usually have during the night?".	Belgium; 3-4 year olds; N=595

		Answer possibilities were “less than 6 h”, “6–7 h”, “8–9 h”, “10–11 h”, “12–13 h”, “14 h”, “more than 14 h”, and “I don’t know”.	
Berglind et al. (2018) <i>PRIMROSE Study</i>	Actigraph GT3X+ (right hip, waking hours, 7 days)- physical activity Questionnaire: Screen time/sleep (specific to study)	Actigraph- Physical activity Questionnaire: Screen time: -Separate questions on how much time on weekdays and weekend days children usually spent playing computer/video games and watching TV/DVDs. <i>The questions each had 14 response options, ranging from 0 to 6 or more hours per day.</i> Sleep: Sleeping hours on weekdays and on weekends. <i>14 response options ranging from ‘5 h or less’ to ‘11 h or more’ per night</i>	Sweden; 4-5 years; N=830
Hinkley et al. (2020) <i>Healthy Active Preschool and Primary Years (HAPPY) Study</i>	Actigraph GT1M (right hip, 8 days)- Physical activity Proxy report- Screen time/Sleep (q’s from Hinkley et al., 2012b)	Actigraph – Physical activity Proxy report: Screen time: Parents reported their child’s week (Monday-Friday) and weekend (Saturday-Sunday) sedentary screen-time in TV/DVD/video viewing, internet/computer use and electronic games Sleep: Parents reported the duration of their child’s usual night- and day-time sleep	Australia; 3-5 years; N=471
Leppänen et al. (2019) <i>DAGIS study-increased health and</i>	Actigraph Wgt3x-BT (Waist 24hr, 7 days) - Physical activity	Actigraph- Physical activity <i>The time spent in different physical activity intensities was calculated using Butte’s cut-points. Overall physical activity was calculated as: [(mean physical activity on weekdays × 5) + (mean physical activity on weekend days × 2)] / 7. To calculate the proportion of children complying with the physical activity guideline separately for overall, weekdays,</i>	Finland; 3-6 years; N=778 (of 864)

<i>wellbeing in pre-schoolers</i>	(Diary for non-wear time) Diary (7 days) – sedentary behaviour and asleep (based on Wen et al., 2010)	<i>and weekend days, physical activity was dichotomized into 0 (< 180 min/day of total physical activity, of which at least 60 min was moderate to vigorous physical activity) and 1 (≥180 min/day of total physical activity, of which at least 60 min was moderate to vigorous physical activity).</i> Diary: Parents were asked to assess the frequency and time (hours/minutes) that their child spent each day: (1) watching TV, (2) watching DVDs or videos, (3) using tablets or smartphones, and (4) using computers or playing computer games. Screen time is a composition variable of all of the above-mentioned types of activities. <i>Overall screen time was calculated as follows: [(mean screen time on weekdays × 5) + (mean screen time on weekend days × 2)] / 7</i> Diary: Guardians were asked to report their child's bedtime and wake-up time. <i>Sleep duration was calculated by subtracting bedtime from wake-up time. Overall sleep duration was calculated as follows: [(mean sleep duration on weekdays × 5) + (mean sleep duration on weekend days × 2)] / 7. To calculate the proportion of children complying with the sleep guideline separately for overall, weekdays, and weekend days, sleep was dichotomized into 0 (< 10 or > 13 h of sleep/day) and 1 (10–13 h of sleep/ day)</i>	
Guan et al. (2020) N/A	Actigraph GT9X (right hip, 24hr, 3 days)- physical activity, sedentary behaviour, sleep (Activity log – naps, bed time, wake time)	Actigraph- Physical activity, sedentary behaviour, sleep Proxy report, screen time: How long was your child's screen time (e.g., watching television, using iPad or phone, etc) on a typical weekday, a typical Saturday and a typical Sunday, respectively? <i>Child's average daily screen time was calculated (average daily screen time = (weekday screen time *5 + Saturday screen time + Sunday screen time)/7)</i> Sleep:	China; 4.1-6.3 years N=254

	Proxy report- Screen time (specific to study)	The nap offset/wake-up time was firstly located when a switch in the output from inclinometer lying or off to inclinometer sitting or standing was detected, which corresponded with the nap offset registered in the activity log. The offset was then identified as the first minute followed by at least 10 consecutive minutes with a vector magnitude > 0 in the output. <i>Nap duration was calculated as the period between the nap onset time and the offset time. Similarly, nighttime sleep duration was calculated as the period between bedtime and waketime. Total sleep duration was calculated as the sum of the nighttime sleep duration and nap(s) duration. For the calculation of sleep duration, wake after sleep onset was not included.</i>	
Chen et al. (2019) <i>Growing Up in Singapore towards Healthy Outcomes (GUSTO)</i>	Actigraph (wrist worn, 24hr, 7 days)- Physical activity/Sleep Proxy report – Screen time (Items from PRE-PAQ, Dwyer et al., 2011).	Actigraph- physical activity and sleep Proxy report (adapted from PRE-PAQ): Screen time: Which of the following activities did your child do <u>yesterday</u> and <u>last weekend</u>? (If yes, record the times for each activity) a) a. Sat or lay still while watching television/DVD/VCR/Blu-ray disc player or other videos on a television screen b) Sat or lay still while watching videos or playing games on a computer, such as a desktop, laptop, or ultrabook c) Sat or lay still while watching videos or playing games on mobile screen devices, such as a smartphone or tablet d) Sat or lay still while playing games on video game console, such as Xbox, Nintendo Wii, Nintendo DS, PSP	Singapore; 5.5 years; N=547
Accelerometer = Actical (n=4)			
Chaput et al. (2017) <i>Canadian Health</i>	Actical (Right hip 7 days)- physical activity	Actical - physical activity Sedentary behaviour: On average, how many hours a day the child spends 1) watching TV or videos or playing video games, and 2) on a computer.	Canada; 3-4 years; N=803

<i>Measures Survey (Cycles 2 and 3 included 3-6 year olds; cycle 1 only included 6 + year olds)</i>	Proxy report-Screen time/Sleep(specific to study)	<i>The response categories differed between cycle 2 (none, <1, 1 to 2, 3 to 4, 5 to 6, ≥7) and cycles 3–4 (none, <1, 1 to <3, 3 to <5, 5 to <7, ≥7). Results of the two screen time questions were summed to determine average daily screen time. Participants were categorized as meeting the screen time recommendation if that sum was ≤1 h/day.</i> Sleep: “How many hours does your child usually spend sleeping in a 24-hour period, excluding time spent resting?” <i>Responses were rounded to the closest half-hour by the interviewer. Participants were categorized as meeting the sleep duration recommendation if parents reported that their child obtained between 10 and 13 h of sleep a day.</i>	
Carson et al. (2017a) <i>Canadian Health Measures Survey</i>	Actical (right ip, 7 days)- physical activity /sedentary behaviour Proxy report: Sleep (specific to study)	Actical – sedentary behaviour, light physical activity, moderate to vigorous physical activity Proxy report: Sleep: ‘How many hours does your child usually spend sleeping in a 24 hour period, excluding time spent resting?’ <i>Responses recorded to the nearest half hour and mins of sleep were calculated.</i>	Canada; N=552; 3-4 years
Meredith Jones et al. (2019) <i>Prevention of Overweight in Infancy (POI)</i>	Actical accelerometers (Waist, 24hr, 5-7 days)- physical activity/sleep Proxy report (interview administered-specific to study): screen time (and	Actical for physical activity and sleep Questions: • Parents asked to indicate how many times per usual week and for how many minutes each time, their child watched television, videos or DVDs. • When the children were 5 years of age, the number of media types was expanded with parents indicating the frequency (per week) and duration (each time) their child spent watching/using each of television (including free to air, cable, and online streaming), DVDs or videos, computers for games (not schoolwork), passive games consoles (including handheld), and mobile phones or tablets for playing games.	New Zealand; Children at ages 1, 2, 5 years; N=547

	restraint for 1/2 year olds only)	At 1 and 2 years of age, parents indicated the number of times over the past week, and the length of time each session, the infant was placed in a car seat or stroller	
Kuzik et al. (2016) <i>Supporting Active Living Behaviors in Alberta Child Care Settings study</i>	Actual (right hip, 5 days)- physical activity, sedentary behaviour, sleep Proxy report- Sedentary behaviour, moderate to vigorous physical activity, sleep (Moderate to vigorous physical activity and sedentary behaviour questions adopted from Statistic Canada's Canadian Health Measures Survey.)	Actual – physical activity, sedentary behaviour, sleep Proxy report: Moderate to vigorous physical activity: “About how many hours a week does your child usually take part in physical activity (that makes him/her out of breath or warmer than usual) outside of child care while participating in...” both “...organized activities (e.g., swimming lessons, skating lessons, gymnastics)?” and “...non-organized activities (e.g., going for a walk, drop-in skating, playing at a splash pad or wading pool, bike or tricycle ride, playing at the park or in the yard)?” <i>For both questions, a 5-point scale ranging from “never” to “7+ hours/week” was used.</i> Sedentary behaviour: Screen time: “On average, how much time per day outside of child care does your child...” both “...watch television, videos or DVDs on a television, computer or portable device?” and “...play video/computer games on devices such as a learning laptop, leapfrog leapster, computer, laptop, tablet, cell phone, the internet, Playstation, XBOX?”. Non-screen time sedentary behaviour was assessed by asking parents two questions: “On average, how much time per day outside of child care does your child spend...” both “...in a motor vehicle (e.g., car, LRT [light rail transit], bus)?” and “...being safely restrained in a high chair, stroller, etc. (do not include when they are in a motor vehicle)?” <i>For all questions, a 7-point scale ranging from “none” to “3 h or more” were used for weekdays and weekend days separately.</i> Sleep (total night time sleep): “How long does your child usually sleep per night at the moment?” <i>Responses for hours and minutes were used to calculate an hours/night variable.</i>	Canada; 19-60 months; N=100

Roberts et al. (2017) <i>Canadian Health Measures Survey</i>	Actical (right hip, 7 days during waking hours) – Moderate to vigorous physical activity. Proxy report-screen time and sleep	Actical – number of minutes of moderate to vigorous physical activity on each valid day was summed, daily totals were averaged. Sleep: Sleep duration was parent-reported (ages 5 to 11)- average daily sleep time was determined from the question: “How many hours does your child usually spend sleeping in a 24-hour period, excluding time spent resting?” Answers were rounded to the closest half hour by the interviewer. Screen time: Parent reported: How many hours per day the child spends: 1) watching TV or videos or playing video games; and 2) on a computer. The response categories differed between cycle 2 (none, less than 1, 1 to 2, 3 to 4, 5 to 6, 7 or more) and cycle 3 (none, less than 1, 1 to less than 3, 3 to less than 5, 5 to less than 7, 7 or more). For cycle 2, average daily screen time was derived using the midpoint of the response category (0, 0.5, 1.5, 3.5, 5.5, 7). To maintain consistency, the same values were assigned to each of the respective categories for cycle 3.	Canada; 5-11 year olds; N=1,985
Carson et al. (2017b) <i>Canadian Health Measures Survey</i>	Actical (waist worn) - Moderate to vigorous physical activity. Proxy report-Sedentary behaviour and sleep	Actical accelerometer to determine moderate to vigorous physical activity, required to have more than 4 valid days with more than 10 hr wear time. Screen time: On average, how many hours a day does your child 1) watch TV or videos or play video games? 2) spend on a computer? Sleep: ‘How many hours do you (your child) usually spend sleeping in a 24 hour period, excluding time spent resting?’	Canada; 6-17 years; N=4157

*New Zealand: New Zealand Health Survey 2016/17. Q’s (around screen time) for children aged 2-14 years:
<https://www.health.govt.nz/publication/content-guide-2016-17-new-zealand-health-survey>

*Wales: National Survey for Wales: Proxy report for children aged 3-7 years on duration of MVPA in last 7 days; *Scotland: Scotland Health Survey: Q’s for children aged 2-15 years (7 day recall of activities); * N Ireland: Continuous Household Survey: Q’s for children aged 4-19 years (proxy reported for younger children): includes single item on average time spent in sport/exercise activities per day (Strain et al., 2019)

References

Australian Health Survey (2012).

<https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4364.0.55.004Chapter3002011-12>

Berglind, D., Ljung, R., Tynelius, P., & Brooke, H. L. (2018). Cross-sectional and prospective associations of meeting 24-h movement guidelines with overweight and obesity in preschool children. *Pediatric obesity*, 13(7), 442-449.

Carson, V., Tremblay, M. S., & Chastin, S. F. (2017a). Cross-sectional associations between sleep duration, sedentary time, physical activity, and adiposity indicators among Canadian preschool-aged children using compositional analyses. *BMC Public Health*, 17(5), 848

Carson, V., Chaput, J. P., Janssen, I., & Tremblay, M. S. (2017b). Health associations with meeting new 24-hour movement guidelines for Canadian children and youth. *Preventive medicine*, 95, 7-13.

Chaput, J. P., Colley, R. C., Aubert, S., Carson, V., Janssen, I., Roberts, K. C., & Tremblay, M. S. (2017). Proportion of preschool-aged children meeting the Canadian 24-Hour Movement Guidelines and associations with adiposity: results from the Canadian Health Measures Survey. *BMC Public Health*, 17(5), 829.

Chen, B., Bernard, J. Y., Padmapriya, N., Yao, J., Goh, C., Tan, K. H., ... & Chan, S. Y. (2019). Socio-demographic and maternal predictors of adherence to 24-hour movement guidelines in Singaporean children. *International Journal of Behavioral Nutrition and Physical Activity*, 16(1), 1-11.

Chia, M. Y. H., Tay, L. Y., & Chua, T. B. K. (2019). Quality of Life and Meeting 24-h WHO Guidelines Among Preschool Children in Singapore. *Early Childhood Education Journal*, 1-11.

Cliff, D. P., McNeill, J., Vella, S. A., Howard, S. J., Santos, R., Batterham, M., ... & de Rosnay, M. (2017). Adherence to 24-Hour Movement Guidelines for the Early Years and associations with social-cognitive development among Australian preschool children. *BMC Public Health*, 17(5), 857.

De Craemer, M., McGregor, D., Androutsos, O., Manios, Y., & Cardon, G. (2018). Compliance with 24-h movement behaviour guidelines among Belgian pre-school children: the ToyBox-study. *International journal of environmental research and public health*, 15(10), 2171

Goodlin-Jones, B. L., Sitnick, S. L., Tang, K., Liu, J., & Anders, T. F. (2008). The Children's Sleep Habits Questionnaire in toddlers and preschool children. *Journal of Developmental & Behavioral Pediatrics*, 29(2), 82-88.

Guan, H., Zhang, Z., Wang, B., Okely, A. D., Tong, M., Wu, J., & Zhang, T. (2020). Proportion of kindergarten children meeting the WHO guidelines on physical activity, sedentary behaviour and sleep and associations with adiposity in urban Beijing. *BMC pediatrics*, 20(1), 1-9.

Health Survey for England. (2015).

Hesketh, K. D., Downing, K. L., Campbell, K., Crawford, D., Salmon, J., & Hnatiuk, J. A. (2017). Proportion of infants meeting the Australian 24-hour Movement Guidelines for the Early Years: data from the Melbourne InFANT Program. *BMC public health*, 17(5), 856.

Hinkley, T., Timperio, A., Watson, A., Duckham, R. L., Okely, A. D., Cliff, D., ... & Hesketh, K. D. (2020). Prospective associations with physiological, psychosocial and educational outcomes of meeting

Australian 24-Hour Movement Guidelines for the Early Years. *International Journal of Behavioral Nutrition and Physical Activity*, 17(1), 1-12.

Hinkley, T., Salmon, J. O., Okely, A. D., Crawford, D., & Hesketh, K. (2012). Preschoolers' physical activity, screen time, and compliance with recommendations. *Medicine & Science in Sports & Exercise*, 44(3), 458-465.

Hinkley, T., Salmon, J., Okely, A. D., Crawford, D., & Hesketh, K. (2012). The HAPPY study: development and reliability of a parent survey to assess correlates of preschool children's physical activity. *Journal of science and medicine in sport*, 15(5), 407-417.

Kuzik, N., & Carson, V. (2016). The association between physical activity, sedentary behavior, sleep, and body mass index z-scores in different settings among toddlers and preschoolers. *BMC pediatrics*, 16(1), 100.

Lee E-Y, Hesketh KD, Hunter S, Kuzik N, Rhodes RE, Rinaldi CM, et al. Meeting new Canadian 24-hour movement guidelines for the early years and associations with adiposity among toddlers living in Edmonton, Canada

Leppänen, M. H., Ray, C., Wennman, H., Alexandrou, C., Sääksjärvi, K., Koivusilta, L., ... & Roos, E. (2019). Compliance with the 24-h movement guidelines and the relationship with anthropometry in Finnish preschoolers: the DAGIS study. *BMC Public Health*, 19(1), 1618.

Roberts, K. C., Yao, X., Carson, V., Chaput, J. P., Janssen, I., & Tremblay, M. S. (2017). Meeting the Canadian 24-hour movement guidelines for children and youth. *Health Rep*, 28(10), 3-7.

Santos, R., Zhang, Z., Pereira, J. R., Sousa-Sá, E., Cliff, D. P., & Okely, A. D. (2017). Compliance with the Australian 24-hour movement guidelines for the early years: associations with weight status. *BMC Public Health*, 17(5), 867.

Taylor RW, Haszard JJ, Meredith-jones KA, Galland BC, Heath AM, Lawrence J, et al. 24-h movement behaviors from infancy to preschool: Cross-sectional and longitudinal relationships with body composition and bone health; 2018

Meredith-Jones, K., Galland, B., Haszard, J., Gray, A., Sayers, R., Hanna, M., ... & Taylor, R. (2019). Do young children consistently meet 24-h sleep and activity guidelines? A longitudinal analysis using actigraphy. *International Journal of Obesity*, 43(12), 2555-2564.

Strain, T., Milton, K., Dall, P., Standage, M., & Mutrie, N. (2019). How are we measuring physical activity and sedentary behaviour in the four home nations of the UK? A narrative review of current surveillance measures and future directions. *British journal of sports medicine*, bjsports-2018.

Appendix 6.3: Current activity surveillance systems

Country	Name of Surveillance/Survey	Methods used	Relevant references
UK	Health Survey for England	Questions include reporting their child's physical activity (frequency and duration) during the past week, in line with the following topics: • Active travel: walk or cycle to/from school • Informal activities done outside school hours (e.g. cycling, walking, playing on trampoline, dancing, skipping rope) • Formal activities done outside school hours (e.g. sports, swimming) • Any other activities not listed • Activities conducted during school time/ in lessons – including which days and how long (may not always be possible for 2-4 year olds as these are proxy reported). For sedentary behaviour, questions include reporting child's average sedentary behaviour (frequency and duration) for a typical weekday and weekend day, in line with the following topics: • Television viewing • Other sedentary behaviour (e.g. reading, doing homework, using a computer or playing video games).	Basterfield, L., Adamson, A. J., Parkinson, K. N., Maute, U., Li, P. X., Reilly, J. J., & Gateshead Millennium Study Core Team. (2008). Surveillance of physical activity in the UK is flawed: validation of the Health Survey for England Physical Activity Questionnaire. <i>Archives of disease in childhood</i>, 93(12), 1054-1058. NHS Digital. (2016). <i>Health Survey for England 2015 Physical Activity in Children</i>. Available at: hse2015-child-phy-act.pdf (digital.nhs.uk). Strain, T., Milton, K., Dall, P., Standage, M., & Mutrie, N. (2019). How are we measuring physical activity and sedentary behaviour in the four home nations of the UK? A narrative review of current surveillance measures and future directions. <i>British journal of sports medicine</i>, bjsports-2018.
Northern Ireland	Continuous Household Survey	• Asked of children aged 4-19 years, parental proxy for younger children • Single item on time spent in other sport and exercise activities per day • Unanchored recall period 'on average'	Strain, T., Milton, K., Dall, P., Standage, M., & Mutrie, N. (2019). How are we measuring physical activity and sedentary behaviour in the four home nations of the UK? A narrative review of current

			surveillance measures and future directions. <i>British journal of sports medicine</i> , bjsports-2018.
Scotland	Scotland Health Survey	• Asked of all children aged 2-15 years • BUT there is no reporting on the under 5s recommendation • 7-day recall period, activities reported on specific days • Domain specific, including activity at school	Strain, T., Milton, K., Dall, P., Standage, M., & Mutrie, N. (2019). How are we measuring physical activity and sedentary behaviour in the four home nations of the UK? A narrative review of current surveillance measures and future directions. <i>British journal of sports medicine</i>, bjsports-2018. Scottish Government. (2019). <i>Scottish Health Survey 2019 - volume 1: main report</i>. Available at: Chapter 7: Physical Activity - Scottish Health Survey 2019 - volume 1: main report - gov.scot (www.gov.scot).
Wales	National Survey for Wales	• Proxy report for all children aged 3-7 years • Duration of any moderate to vigorous physical activity reported for each day in last 7 days	Strain, T., Milton, K., Dall, P., Standage, M., & Mutrie, N. (2019). How are we measuring physical activity and sedentary behaviour in the four home nations of the UK? A narrative review of current surveillance measures and future directions. <i>British journal of sports medicine</i>, bjsports-2018. Welsh Government. (2018). <i>National Survey for Wales 2017-18 questionnaire</i>. Available at: https://gov.wales/sites/default/files/statistics-and-research/2020-01/national-survey-wales-questionnaire-2017-18.pdf.
Canada *Only surveillance system that is able to capture movement behaviour guidelines	Canadian Health Measures Survey	• Children aged 3-4 years • Physical activity accelerometer derived (Actical accelerometers for 7 days) • Parents/guardians were asked, on average, how many hours a day the child spends 1) watching TV or videos or playing video games, and 2) on a computer. The response categories differed between cycle 2 (none, <1,	Chaput, J. P., Colley, R. C., Aubert, S., Carson, V., Janssen, I., Roberts, K. C., & Tremblay, M. S. (2017). Proportion of preschool-aged children meeting the Canadian 24-Hour Movement Guidelines and associations with adiposity: results from the Canadian Health Measures Survey. <i>BMC Public Health</i>, 17(5), 829.

		1 to 2, 3 to 4, 5 to 6, ≥ 7 hours) and cycles 3–4 (none, <1, 1 to <3, 3 to <5, 5 to <7, ≥ 7 hours). Results of the two screen time questions were summed to determine average daily screen time. Participants were categorized as meeting the screen time recommendation if that sum was ≤ 1 h/day. Parents/guardians were asked: “How many hours does your child usually spend sleeping in a 24-hour period, excluding time spent resting?” Responses were rounded to the closest half-hour by the interviewer. Participants were categorized as meeting the sleep duration recommendation if parents reported that their child obtained between 10 and 13 h of sleep a day.	
Australia	Australian Health Survey 2011/12	- Children aged 2–4 years (sporadically asked for this age group, 2011/12 is the most recent survey that included this age group) - Average physical activity over 7 days prior to interview - Outdoor physical activity or active play included throwing a ball, walking, running around, free play in playgrounds or sandpits, helping in the garden, or any other outdoor activity that involved movement. - Indoor physical activity or active play included dancing, jumping, 'rough and tumble', tidying up, helping to set the table, or any other indoor activity that involved movement.	Australian Bureau of Statistics. (2013). <i>Australian Health Survey: Physical Activity</i>. Available at: https://www.abs.gov.au/statistics/health/health-conditions-and-risks/australian-health-survey-physical-activity/latest-release. Evidence that 2014/15 survey does not include questions for PA for children under 15 years old: Australian Bureau of Statistics. (2016). <i>National Health Survey: First Results, 2014–15</i>. Available at: https://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/EFC93ECE5ABE3894CA257F1400133CE1/\$File/national%20health%20survey%202014-15%20questionnaire.pdf

		2. Average daily duration of sedentary screen based activity 7 days prior to interview Response options: number of hours	Australian Bureau of Statistics. (2018). <i>National Health Survey: First Results</i>. Available at: https://www.abs.gov.au/statistics/health/health-conditions-and-risks/national-health-survey-first-results/latest-release#key-statistics.
South Africa	No surveillance of guidelines as of yet, cannot find any source of information to say how this will be done. But these guidelines were very recently implemented.		
New Zealand	New Zealand Health Survey 2016/17	Children aged 2-14 years, questions for pre-school aged children in relation to screen time, but PA questions are restricted to children aged 5 years and older only 1. What is the average amount of time child spends watching TV <u>each weekday</u>? This could be anywhere, not just in your home, and includes DVDs / videos but does not include games. Responses rounded to nearest hour, and could include television programmes, videos and movies watched on any device. Include those accessed online (e.g. via YouTube) and on-demand (e.g. via Netflix). Response options: _____ hours (range 0–24) 2. What is the average amount of time child spends <u>each weekday</u> looking at a screen doing things other than watching TV or videos? For example, playing video games or browsing the Internet. This does <u>not</u> include time spent at school or on homework. Responses rounded to nearest hour, and could	Ministry of Health, New Zealand. (2017a). <i>Content Guide 2016/17: New Zealand Health Survey</i>. Wellington: Ministry of Health. Available at: Content Guide 2016/17: New Zealand Health Survey Ministry of Health NZ.

		include activities on a tablet, computer, electronic gaming device or other hand-held electronic device such as a smart phone. Include texting, emailing and using social media. Don't count time reported in the previous question. Response options: _____ hours (range 0–24) 3. What is the average amount of time child spends watching TV <u>in the weekend</u>? Again, this could be anywhere, not just in your home and includes DVDs / videos but does not include games. Responses rounded to nearest hour. Record total hours over <u>both</u> Saturday and Sunday. Include television programmes, videos and movies watched on any device. Include those accessed online (e.g. via YouTube) and on-demand (e.g. via Netflix). Response options: _____ hours (range 0–48) 4. What is the average amount of time child spends <u>in the weekend</u> looking at a screen doing things other than watching TV or videos? For example, playing video games or browsing the Internet. This does <u>not</u> include time spent at school or on homework. Responses rounded to nearest hour. Record total hours over <u>both</u> Saturday and Sunday. Include activities on a tablet, computer, electronic gaming device or other hand-held electronic device such as a smart phone.	

		Include texting, emailing and using social media. Don't count time reported in the previous question. Response options: _____ hours (range 0–48)	
World Health Organisation	Guidelines only released this year, but the accompanying report has no suggestion on surveillance of these. -World Health Survey that has been used previously only explored health of adults (ages 18+)		https://apps.who.int/iris/bitstream/handle/10665/311664/9789241550536-eng.pdf
National Sleep Foundation	National Sleep Foundation Questionnaire, 2004	1. Think about your child's sleep schedule in the past two weeks. On a typical day in the past two weeks, what was the usual time that your child woke up in the morning? (Range of responses). 2. On a typical night in the past two weeks, what was the usual time that your child went to sleep for the night? (Range of responses). 3. How often does your child go to sleep at this time? I. Every night or almost every night II. A few nights a week III. About once a week IV. Less than once a week V. Refused don't know 4. On a typical night in the past 2 weeks, how many minutes did it take your child to fall asleep from lights out to being asleep? 5. On a typical night in the past 2 weeks, how many hours did your child actually sleep at night before the hours of 6pm and 8am? 6. On a typical night in the past 2 weeks, how many hours did your child actually sleep in the daytime between 8am and 6pm?	National Sleep Foundation. (2002). <i>National Sleep Foundation 2003 Sleep In America Poll Screening Questionnaire</i>. Available at: https://www.thensf.org/wp-content/uploads/2021/03/2003-SIA-Questionnaire.pdf.

		7. How much sleep do you think your child needs in a 24 hour period? (Hours and minutes). 8. On most nights, do you believe that our child gets... I. Too little sleep II. Too much sleep III. Right amount of sleep 9. How does your child's sleep pattern differ on the weekends compared to weekdays? Do they... I. Get less sleep on weekends II. More sleep on weekends III. Same amount of sleep 10. Does your child have a usual bedtime routine, that is on most nights, do the same activities occur? 11a. What are the 3 most common activities that occur before your child goes to sleep? (list of assorted activities) 11. Does Your child currently take a nap? 12. On days when your child takes a nap, how many times each day do they take a nap? 13. Does your child have their own bedroom? 14. Who does your child share a room with? 15. Does your child have their own bed? 16. Does your child have any of the following in their room? (television, computer, telephone, nightlight) 17. Who usually puts your child to bed? 18. How often is a parent or other adult present in the room when your child falls asleep ? (every night, few nights a week, about once a week, rarely, never)	

		19. Does your child regularly spend the night at another home during the week or weekends, such as at another parents home or grandparents house? 20. How many nights per week does your child spend at the other home? 21. During a typical night in the past 2 weeks, how many times did your child wake up and need your help or attention? 22. What was the approx. number of minutes your child was awake during the night? 23. How often do you usually do the following when your child wakes up during the night? (list of things – allow them to fall back to sleep, briefly see them, stay with them, read a story etc.) 24. How often does your child return to sleep without help? 25. Who usually goes to your child when they wake up? 26. Thinking about your child's sleep do you think they have any sleep problems? 27. During the past 2 weeks, how often does your child ..? (have trouble breathing, snore, stall about going to bed etc) 28. Is there was one thing that you could change about your child sleep, what would it be? 29. On a typical night in the past 2 weeks, where did your child fall asleep? 30. On a typical night in the past 2 weeks, where did your child usually wake up in the morning? 31. During the past 2 weeks, how often during the course of the day does your child engage in	

		following behs? (cling to adult, whine/complain, smile/be happy, fuss when they do not get their own way). -Additional questions for pre-schooler (whom fall in 0-5 age range): 32. Approx how many hours per week is your child usually cared for by someone other than a parent or primary caregiver, including attending a preschool program?	

Additional UK Surveillance of Health Behaviours in the Early Years

Country	Name of Surveillance/Survey	Methods used	Relevant references
UK	Early Years Foundation Stage	- Children aged 2/3 years - Progress check in the form of a written summary. One aspect to be covered is that of 'physical development' (moving and handling; health and self care – one aspect of which is: Children should know the importance of physical activity and a healthy diet, and why this is important for good health of physical exercise) Second profile completed when child is age 5 years. - This involves an assessment of the child's development in line with the early learning goals (for physical development, this includes 1) Moving and handling and 2) Health and self care).	https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/596629/EYFS_STATUTORY_FRAMEWORK_2017.pdf
UK	Personal Child Health Record (PCHR)	- Review when child is age 2-2.5 years - Physical development outlined as a topic to discuss, but there are no direct questions/reports of physical activity, sedentary behaviour, or sleep at this age (sleep is included for baby and 1 year old review).	https://www.rcpch.ac.uk/resources/personal-child-health-record-pchr
UK	Healthy Child Programme (HCP) Health and Development Review	For physical health the following items can be asked and recorded. However, there are no questions directly relating to physical activity and sedentary behaviours of children, despite this section exploring physical health. Height and weight: Height and weight measurement of child using approved height measure and weighing scales. General health: Parent or carer is asked 'In general would you say that your child's health is excellent, very good, good, fair or poor?' Sleep: 'Is your child's sleep a problem for you?' – None, mild, moderate, large.	https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/377800/dh_108329.pdf

Appendix 6.4: MoveMEY Scoring Sheet (Provisional)

Provisional Scoring System for MoveMEY Questionnaire



Behaviour/Construct	Questions used for scoring	Additional Comments
Physical Activity		
Total Daily Physical Activity (PA)	Q1a + Q2a+ Q3a	- Daily PA for each question calculated by: duration of activity each day (e.g. 16 mins) multiplied by the frequency that activity occurred (e.g. 5 days), this will be divided by 7 to estimate daily minutes of PA. - The 'daily PA' time for all 3 individual questions will be summed to provide a total daily PA. - For the calculation of PA, minimum reported durations will be used (e.g. if the box 16-30 minutes was ticked, 16 mins would be used), unless free text

		responses (exceeding 60 min/per) are used.
Total Daily moderate to vigorous physical activity (MVPA)	Q1c + Q2c + Q3c	Total of the 3 questions/7
Sedentary Behaviour		
Total Daily Sedentary Behaviour (SB)	Q4 + Q5+ Q6	- Daily SB for each question calculated by: duration of activity each day (e.g. 45 mins) multiplied by the frequency that activity occurred (e.g. 5 days), this will be divided by 7 to estimate daily minutes of SB. - The 'daily SB for all 3 individual questions will be summed to provide a total daily SB. - For the calculation of SB, highest reported durations will be used (e.g. 45 mins), unless free text responses (exceeding 60 min/per) are used.

Total Daily Screen Time	Q5	Duration of activity each day (e.g. 45 mins) multiplied by the frequency that activity occurred (e.g. 5 days), this will be divided by 7 to estimate daily screen time.
Not being restrained for more than 1 hour	Q4	Duration of activity each day (e.g. 45 mins) multiplied by the frequency that activity occurred (e.g. 5 days), this will be divided by 7 to estimate daily time restrained.
When sedentary engaging in pursuits such as storytelling/reading	Q6	Duration of activity each day (e.g. 45 mins) multiplied by the frequency that activity occurred (e.g. 5 days), this will be divided by 7 to estimate daily sedentary time engaged in non-screen based activities.
Sleep		
10-13 hrs good quality sleep (may include a nap)	Q7a - Q8 - Q9 + Q10	Total mins from time in bed (Q7a) – total mins to fall

		asleep – total mins awake during night + total nap duration. • As night wakings and nap duration were in the same format as PA and SB questions, duration of nap/night waking each day (e.g. 45 mins) multiplied by the frequency that activity occurred (e.g. 3 days), and divided by 7 to estimate nap/night wakings. • For the calculation of night wakings, highest reported durations will be used (e.g. if the box 16-30 minutes was ticked, 30 mins would be used), unless free text responses (exceeding 60 min/per) are used. • For the calculation of nap duration, lowest reported

		durations will be used (e.g if the box 16-30 minutes was ticked, 16 mins would be used), unless free text responses (exceeding 60 min/per) are used.
Good quality sleep	Time in bed= Q7a Total sleep time=Q7a- Q8 - Q9	• Based on sleep efficiency – ratio of total sleep time to time in bed. • $\geq 85\%$ = good sleep quality • Total sleep time (total time in bed – sleep latency – night time wakings) divided by time in bed (bed time and wake time) x 100
Consistent wake and bed time	Q7b	≥ 5 days

Appendix 6.5: Face and content validity qualitative survey

Face/Content Validity of MoveMEY Questionnaire

Name of Expert Reviewer: _____

Date: ____/____/____

1. Do the items of the questionnaire look as though they are an adequate reflection of the physical activity, sedentary behaviour and sleep of pre-school children?

	Yes	No	Comments
Physical activity			
Sedentary behaviour			
Sleep			

2. Do the items of the questionnaire look as though they are an adequate reflection of the pre-school physical activity and movement guidelines (both UK and WHO)?

	Yes	No	Comments

UK			
WHO			

3. Are all of the questions relevant to at least one of the domains being measured (physical activity, sedentary behaviour or sleep)?

Yes	No	Comments:

4. Do the questions cover are all relevant activities in sufficient detail?

Yes	No	Comments:

5. Do the items on the questionnaire present information in a way that will evoke accurate answers?

Yes	No	Comments:

6. Are the items of the questionnaire comprehensible?

Yes	No	Comments:

7. Is the questionnaire comprehensible?

Yes	No	Comments:

8. Are there any important questions missing?

Yes	No	Comments:

9. Are there any unsuitable questions included?

Yes	No	Comments (including how these might be changed):

Scoring of the Questionnaire

10. Are the response options used to score the questions suitable?

Yes	No	Comments:

11. Do the scores of the questionnaire look as though they would adequately identify whether pre-school children are meeting the physical activity and movement guidelines (both UK and WHO)?

	Yes	No	Comments
UK			
WHO			

Please provide any further comments, suggestions or potential adjustments to the questionnaire:

Thank you very much for taking the time to review this questionnaire.

Appendix 6.6: Recruitment poster

Are you a parent or carer of a child aged 3 or 4 years old?



Take part in an exciting research study (from home!) and receive a **£20 Love2Shop voucher**

What would taking part look like?

Provide some information about you and your child

➔

Complete a diary (named MoveMEY) about your child's movement behaviour

➔

Chat to a researcher (phone or video call) about your experiences of completing MoveMEY

All information will be non-identifiable and will be kept fully confidential.

For more information or to take part, please contact me- Sophie Phillips (researcher) on email: sophie.m.phillips@durham.ac.uk or call on **07989143478**





School for Public Health Research



This project is funded and supported by the National Institute for Health Research (NIHR) School for Public Health Research (SPHR), Grant Reference Number PD-SPH-2015. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Appendix 6.7: Information sheet

Information for Parents and Carers**Movement Measurement in the Early Years (MoveMEY): Pilot Study**

We would like to invite you to take part in this research study. Before you decide if you would like to take part, you need to understand, why the research is being done and what it would involve. Please take time to read the following information carefully.

Please ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is this study about?

We have previously done some research with parents and carers of pre-school aged children that involved developing a measurement tool that can be used to look at physical activity, sedentary behaviour (behaviours when your child is sitting, reclining or lying) and sleep. Now we are doing some research that will involve testing the tool to see how it works in practice.

What will be involved if I decide to take part?**Can I take part?**

You are eligible to take part if you have a child aged 3 or 4 years old and are able to speak English (as the study materials are written in English).

Do I have to take part?

No. Your participation is voluntary and you do not have to agree to take part. Once you have read this information sheet, if you decide to take part then you will be asked to sign a consent form to agree to take part. If you change your mind, you can withdraw at any time, without giving a reason by contacting the Sophie (lead researcher) and providing your unique participant number.

What personal information will be asked from me and my child?

We will ask you to complete a questionnaire to provide us with some information about you and your child. We will also ask you to complete the MoveMEY diary regarding your child's movement behaviours, and we will have an interview/chat at the end of the week to discuss your experiences of completing this tool. All of this information will be kept confidential, and at no point will this

of completing this tool. All of this information will be kept confidential, and at no point will this information be linked to you or your child's name. You can leave any questions out if you do not feel comfortable answering them.

Are there any risks in participating?

There are no foreseen risks associated with taking part in this study.

What will happen to the results of the study?

The results will be analysed by the research team. The results will help to identify if the newly developed measurement tool works in practice. The results will be presented in a scientific journal, at conferences, and in the lead researcher's PhD thesis. You can ask the researchers for copies of these publications. The findings will be entirely anonymous and you and your child will not be able to be identified.

Will mine and my child's participation in this study be kept confidential?

All information collected during the study will be kept confidential. All files will be saved on a password protected research laptop in encrypted files. The raw data will only be seen by the lead researcher (Sophie Phillips) and her supervisory team for purposes of analysing the data.

What will we get for taking part?

On returning of the study materials, you will receive a £20 Love2Shop voucher. In addition, at the beginning of the study, we will send your child some 'MoveMEY' stickers for being part of the study.

How is this research being funded and approved?

This project is funded and supported by the National Institute for Health Research (NIHR) School for Public Health Research (SPHR), Grant Reference Number PD-SPH-2015. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

I have some more questions, who should I contact?

We hope that this information sheet has been helpful. However, **you can contact us at any time (now and during the study) using the details below.**

Sophie Phillips (Lead Researcher (PhD Researcher), Department of Sport and Exercise Science, Durham University) (sophie.m.phillips@durham.ac.uk) or call (07989143478).

Supervised by Professor Carolyn Summerbell (Professor Department of Sport and Exercise Science, Durham University) (carolyn.summerbell@durham.ac.uk)

Appendix 6.8: Privacy notice

Privacy Notice

**PART 1 – GENERIC PRIVACY NOTICE**

Durham University has a responsibility under data protection legislation to provide individuals with information about how we process their personal data. We do this in a number of ways, one of which is the publication of privacy notices. Organisations variously call them a privacy statement, a fair processing notice or a privacy policy.

To ensure that we process your personal data fairly and lawfully we are required to inform you:

- Why we collect your data
- How it will be used
- Who it will be shared with

We will also explain what rights you have to control how we use your information and how to inform us about your wishes. Durham University will make the Privacy Notice available via the website and at the point we request personal data.

Our privacy notices comprise two parts – a generic part (ie common to all of our privacy notices) and a part tailored to the specific processing activity being undertaken.

Data Controller

The Data Controller is Durham University. If you would like more information about how the University uses your personal data, please see the University's [Information Governance webpages](#) or contact Information Governance Unit:

Telephone: (0191 33) 46246 or 46103

E-mail: information.governance@durham.ac.uk

Information Governance Unit also coordinate response to individuals asserting their rights under the legislation. Please contact the Unit in the first instance.

Data Protection Officer

The Data Protection Officer is responsible for advising the University on compliance with Data Protection legislation and monitoring its performance against it. If you have any concerns regarding the way in which the University is processing your personal data, please contact the Data Protection Officer:

Jennifer Sewel

University Secretary

Telephone: (0191 33) 46144

E-mail: university.secretary@durham.ac.uk

Your rights in relation to your personal data

Privacy notices and/or consent

You have the right to be provided with information about how and why we process your personal data. Where you have the choice to determine how your personal data will be used, we will ask you for consent. Where you do not have a choice (for example, where we have a legal obligation to process the personal data), we will provide you with a privacy notice. A privacy notice is a verbal or written statement that explains how we use personal data.

Whenever you give your consent for the processing of your personal data, you receive the right to withdraw that consent at any time. Where withdrawal of consent will have an impact on the services we are able to provide, this will be explained to you, so that you can determine whether it is the right decision for you.

Accessing your personal data

You have the right to be told whether we are processing your personal data and, if so, to be given a copy of it. This is known as the right of subject access. You can find out more about this right on the University's [Subject Access Requests webpage](#).

Right to rectification

If you believe that personal data we hold about you is inaccurate, please contact us and we will investigate. You can also request that we complete any incomplete data.

Once we have determined what we are going to do, we will contact you to let you know.

Right to erasure

You can ask us to erase your personal data in any of the following circumstances:

- We no longer need the personal data for the purpose it was originally collected
- You withdraw your consent and there is no other legal basis for the processing
- You object to the processing and there are no overriding legitimate grounds for the processing
- The personal data have been unlawfully processed
- The personal data have to be erased for compliance with a legal obligation
- The personal data have been collected in relation to the offer of information society services (information society services are online services such as banking or social media sites).

Once we have determined whether we will erase the personal data, we will contact you to let you know.

Right to restriction of processing

You can ask us to restrict the processing of your personal data in the following circumstances:

- You believe that the data is inaccurate and you want us to restrict processing until we determine whether it is indeed inaccurate
- The processing is unlawful and you want us to restrict processing rather than erase it
- We no longer need the data for the purpose we originally collected it but you need it in order to establish, exercise or defend a legal claim and
- You have objected to the processing and you want us to restrict processing until we determine whether our legitimate interests in processing the data override your objection.

Once we have determined how we propose to restrict processing of the data, we will contact you to discuss and, where possible, agree this with you.

Retention

The University keeps personal data for as long as it is needed for the purpose for which it was originally collected. Most of these time periods are set out in the [University Records Retention Schedule](#).

Making a complaint

If you are unsatisfied with the way in which we process your personal data, we ask that you let us know so that we can try and put things right. If we are not able to resolve issues to your satisfaction, you can refer the matter to the Information Commissioner's Office (ICO). The ICO can be contacted at:

Information Commissioner's Office Wycliffe House Water Lane Wilmslow Cheshire SK9 5AF

Telephone: 0303 123 1113

Website: [Information Commissioner's Office](#)

PART 2 – TAILORED PRIVACY NOTICE

This section of the Privacy Notice provides you with the privacy information that you need to know before you provide personal data to the University for the particular purpose(s) stated below.

Project Title:

A pilot study of the Movement Measurement in the Early Years (MoveMEY) measurement tool.

Type(s) of personal data collected and held by the researcher and method of collection:

Personal data will be collected through a consent form, demographic questionnaire, 7 day movement diary, and interview (which will be recorded).

- The consent form will include the name of the participant and contact details (including email address, phone number, home address (only for sending study materials and compensation for time vouchers)).
- The demographic questionnaire will include questions about your child, including: child's age, sex, ethnicity, religion, disabilities, and questions about you, including: age, sex, ethnicity, religion, relationship to the child of which you are completing questionnaire/diary (e.g mother, grandfather), parent or carers level of education, employment status, household income and home postcode.
- The interview discussion will be recorded and include views and experiences of completing the measurement tool. If the interview takes place via video call (Zoom), the participant will choose whether they wish to keep their video on or switch this off. Participants will be reminded that the whole screen will be recorded, which will include their picture if they wish

to keep their camera on. However, only the audio recording will be accessed following the interview for purposes of transcription.

Lawful Basis

The present research is taking place under the 'University's public task', whereby collection and use of personal data is carried out under the University's public task, which includes teaching, learning and research.

- For further information see

<https://durham.ac.uk/research.innovation/governance/ethics/governance/dp/legalbasis/>

How personal data is stored:

All personal data will be held securely and strictly confidential to the research team. All data (consent forms, information from the demographic questionnaire and diary, and the recording and transcript from the interviews) will be inputted and stored electronically on to a Durham University password protected research laptop in password protected files - this electronically stored data will only refer to participants by their participant number. The interview recording and paper diaries will be deleted and shredded once this information is stored on the laptop.

How personal data is processed:

The demographic questionnaire information will be entered into a database, to contextualise the responses of the study. The recorded interviews will be transcribed, and personal information will be anonymised. All information inputted for analysis will refer to participants by their participant (and pseudonym when writing up), and never their real name.

Withdrawal of data

Participants will be able to withdraw participation at any point throughout the study, by contacting the lead researcher and providing their participant number.

Who the researcher shares personal data with:

Data will only be available to the lead researcher and her supervisory team for purposes of analysis of this project.

How long personal data is held by the researcher:

All data will be kept for 10 years. After 10 years the data will be discarded appropriately (deletion of electronic files), unless used for a publishable journal, in which case, the data will be discarded 10 years from publishing.

How to object to the processing of your personal data for this project:

If you have any concerns regarding the processing of your personal data, or you wish to withdraw your data from the project, contact: **Sophie Phillips** (Lead Researcher & PhD Student, Department of Sport and Exercise Science, Durham University) on sophie.m.phillips@durham.ac.uk or her supervisor: **Professor Carolyn Summerbell** (Professor, Department of Sport and Exercise Science, Durham University) on carolyn.summerbell@durham.ac.uk.

Appendix 6.9: Consent form

Participant Number: _____

CONSENT FORM
'Movement Measurement in the Early Years
(MoveMEY)': Pilot Study



This form is to confirm that you understand the purposes of the project, what is involved and that you are happy to take part. Please initial each box to indicate your agreement, and send this back to the lead researcher on: sophie.m.phillips@durham.ac.uk

I understand that this research is being conducted as part of a PhD project at Durham University. I confirm that the task has been explained to my fully, I have read and understand the information sheet and privacy notice for the above project. I have had the opportunity to consider the information, ask questions and have had these answered to my satisfaction.	
I understand that my participation is voluntary, that I have the right to not answer any questions I do not feel comfortable with and I am free to withdraw at any time.	
I understand that all data will be kept confidential and will be securely stored.	
I understand that the data obtained from this study, including my words from the interview, may be quoted in publications, reports, and other research outputs, but that my name will not be used and my identity will be kept anonymous at all times.	
I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project.	
I give my permission for the interview to be recorded and transcribed.	

Having read the information sheet, privacy notice and consent form, I confirm that I understand what is involved for me and my child in this research study and that I am happy to take part.

Your name (participant):	
Your signature (participant):	
Your contact details (Phone number and email address):	
Address for sending study materials and vouchers on completion:	
Researcher's Signature:	
Date:	

Contact Details: Researcher: Sophie Phillips (sophie.m.phillips@durham.ac.uk) and primary supervisor: Professor Carolyn Summerbell (carolyn.summerbell@durham.ac.uk).

Department of Sport and Exercise Science, Durham University

Appendix 6.10: Interview discussion guide

Interview Guide: MoveMEY pilot study

Welcome and thank you for volunteering to take part in this research study. I am a PhD student at Durham University interested in improving child health, I am specifically looking at physical activity, sedentary behaviours (any behavior when sat or lying down whilst awake) and sleep of young children. You have been asked to participate as you are the experts where your children are concerned, and I am interested in your opinions and point of view on the newly developed measurement tool that you have just spent the week completing. I realise you are busy and I really appreciate your time.

Introduction: This interview/chat is designed to explore your opinions on the 'MoveMEY' measurement tool that aims to assess the physical activity, sedentary (sitting) behaviours and sleep of your pre-school children, to explore your experiences of using the tool and to see how understandable, relevant, and usable the tool is. The interview will take no more than 1 hour, and the whole discussion will be recorded which will subsequently be transcribed. Double check that participant is happy for the discussion to be audio recorded.

Anonymity: Despite being audio recorded, I would like to assure you that the discussion will be anonymous. The files will be stored in an encrypted folder on a password protected research laptop. If there are any questions that you do not wish to answer you do not have to do so. You are able to withdraw at any time.

Initial warm up questions

1. How have you found the week completing the MoveMEY measurement tool?
2. What are the initial thoughts that come to mind when I ask you of your 'experience' completing the tool?

Guiding questions***Comprehensible***

3. Were the questions asked in the diary appropriately worded/understandable?
4. How did you find the instructions to the tool?
5. Did you find the response options appropriate?
6. Do response options match the question?
7. Were there any questions that you did not put a response to? If yes, which questions and why.
8. What sort of activities came to mind for each of the questions?

9. Did you find it helpful having space to expand on responses by providing additional comments?

Comprehensive

10. Does the tool cover all aspects of your child's movement?
11. Were any behaviours accounted for twice?
12. Did the instructions cover everything that you needed?
13. Do the questions on the tool cover all of your child's behaviours? Were there any behaviours missing that your child engaged in but you did not feel there was space to report?
14. Were the response options sufficient? Were other response options needed?

Relevance

15. Are the included items relevant for measuring the constructs of interest-physical activity, sedentary behavior and sleep? If not, what would make them more relevant?
16. Are the items relevant for measurement of pre-school aged children's behaviours?
17. Are the included items relevant for the context of use ?
18. How did you find reporting daily? Any better alternatives?

Usability

19. How would you describe your overall experience with the tool?
20. What, if anything, surprised you about the experience?
21. What did you like the most about using this tool?
22. What did you like the least? Was there anything you didn't like or that caused frustration?
23. Were there any specific problems with the use of the tool for measuring the behaviours of your pre-school child?

Additional questions...

24. Are there any issues with the readability or clarity of the tool?
25. Do you think the questions are presented in a suitable sequence?
26. Is it unclear what behaviours are being targeted with any of the questions?
27. Are there any suggestions for improvements that could be made to the tool that would make it a more applicable measure?
28. What was it like fitting the tool into existing life?
29. Were there any contexts (i.e if going away for the weekend, staying with grandparent, or a different routines) that impacted on how suitable the tool was?

30. How practical, acceptable and suitable the method may be for assessing the behaviours of their pre-school child?

Concluding questions

31. Is there anything else you wish to say about the tool that you haven't had chance to say?
32. Do you have any questions for me?

Conclusion

- Thank you for participating and for all your time.
- Please use the pre-paid envelope to return the completed MoveMEY diary. I will send the vouchers using the address provided.
- I will email over the debrief sheet. If there is anything you are unhappy with or if you have any questions, please contact me using the email address provided on the debrief sheet. I would like to remind you that any comments featuring in this report will be completely anonymous.

Appendix 6.11: Demographic information questionnaire

Participant Number: _____

Demographic Questionnaire
Movement Measurement in the Early
Years (MoveMEY): Pilot Study



Please answer the following questions in the spaces provided, or tick the most appropriate options. You may also either **highlight** or **change the colour of text for your answers**. Once completed, please send this back to the lead researcher on: sophie.m.phillips@durham.ac.uk

Part 1: Questions about your pre-school child

1. Please provide the date of birth of your child: ____ / ____ ____
2. Please provide the sex of your child.
☐ Male ☐ Female ☐ Other
3. Which of the following best describes your child's ethnicity?
 - ☐ White British
 - ☐ White Irish
 - ☐ Any other White background (Please specify: _____)
 - ☐ Mixed White and Black: Caribbean
 - ☐ Mixed White and Black: African
 - ☐ Mixed White and Asian
 - ☐ Any other Mixed background (Please specify: _____)
 - ☐ Asian or Asian British: Indian
 - ☐ Asian or Asian British: Pakistani
 - ☐ Asian or Asian British: Bangladeshi
 - ☐ Any other Asian background (Please specify: _____)
 - ☐ Black or Black British: Caribbean
 - ☐ Black or Black British: African
 - ☐ Any other Black background (Please specify: _____)
 - ☐ Chinese
 - ☐ Any other ethnic group (Please specify: _____)
 - ☐ Prefer not to say
 - ☐ Don't know

Participant Number: _____

4. Which of the following best describes your child's religion?

- ☐ No religion
- ☐ Christian (including Church of England, Catholic, Protestant and all other Christian denominations)
- ☐ Buddhist
- ☐ Hindu
- ☐ Jewish
- ☐ Muslim
- ☐ Sikh
- ☐ Other religion (please specify _____)
- ☐ Prefer not to say

5. Does your child have any physical or medical condition that affects their ability to play and be physically active?

- ☐ No ☐ Yes (Please state nature of condition _____)

Part 2: Questions about you**6. Please state your relationship to the pre-school child who you are reporting on:**

- ☐ Mother
- ☐ Father
- ☐ Grandmother
- ☐ Grandfather
- Other (please state relationship to child) _____

7. Please provide your age in years _____**8. Please provide your sex**

- ☐ Male ☐ Female ☐ Other

9. Which of the following best describes your ethnicity?

- ☐ White British
- ☐ White Irish
- ☐ Any other White background (Please specify: _____))
- ☐ Mixed White and Black: Caribbean

Participant Number: _____

- ☐ Mixed White and Black: African
- ☐ Mixed White and Asian
- ☐ Any other Mixed background (Please specify: _____))
- ☐ Asian or Asian British: Indian
- ☐ Asian or Asian British: Pakistani
- ☐ Asian or Asian British: Bangladeshi
- ☐ Any other Asian background (Please specify: _____))
- ☐ Black or Black British: Caribbean
- ☐ Black or Black British: African
- ☐ Any other Black background (Please specify: _____))
- ☐ Chinese
- ☐ Any other ethnic group (Please specify: _____))
- ☐ Prefer not to say
- ☐ Don't know

10. Which of the following best describes your religion?

- ☐ No religion
- ☐ Christian (including Church of England, Catholic, Protestant and all other Christian denominations)
- ☐ Buddhist
- ☐ Hindu
- ☐ Jewish
- ☐ Muslim
- ☐ Sikh
- ☐ Other religion (please specify _____)
- ☐ Prefer not to say

Participant Number: _____

11. What is the highest degree or level of school you have completed? If currently enrolled, please state highest degree received.

- ☐ No formal qualifications
- ☐ GCSEs/ O-levels/ CSEs
- ☐ Vocational qualifications such as Apprenticeships or City and Guilds (= NVQ1+2)
- ☐ A-level, Scottish Higher or equivalent (=NVQ3)
- ☐ Diplomas in higher education, HNC/HND/BTEC Higher or equivalent
- ☐ Bachelor Degree or equivalent (=NVQ4-6)
- ☐ Masters/ PhD or equivalent (=NVQ7+8)
- ☐ Other (Please specify _____)
- ☐ Don't know
- ☐ Prefer not to say

12. Which of the following questions best describes your current employment status?

- ☐ Not working, but actively looking for paid work
- ☐ Not working and not looking for paid work
- ☐ Working part-time including self-employment (up to 15 hours)
- ☐ Working part-time including self-employment (16-30 hours)
- ☐ Working full-time including self-employment (30+ hours)
- ☐ Retired early and claiming benefits (ESA or JSA)
- ☐ Retired early and not claiming benefits
- ☐ Retired and receiving state pension
- ☐ Student
- ☐ Unable to work
- ☐ Looking after the home
- ☐ Other (Please specify _____)
- ☐ Prefer not to say

Participant Number: _____

13. Which of the following groups best represents your household's total take-home income, after tax and other deductions?

	Per week	Per month	Per year
☐	Under £48	Under £208	Under £2,500
☐	£48 - £96	£208 - £416	£2,500 - £4,999
☐	£97 - £192	£417 - £833	£5,000 - £9,999
☐	£193 - £388	£834 - £1,289	£10,000 - £14,999
☐	£389 - £585	£1,290 - £1,667	£15,000 - £19,999
☐	£586 - £781	£1,668 - £2,083	£20,000 - £24,999
☐	£782 - £977	£2,084 - £2,500	£25,000 - £29,999
☐	£978 - £1,173	£2,501 - £2,917	£30,000 - £34,999
☐	£1,174 - £1,369	£2,918 - £3,333	£35,000 - £39,999
☐	£1,370 - £1,565	£3,334 - £3,750	£40,000 - £44,999
☐	£1,566 - £1,761	£3,751 - £4,167	£45,000 - £49,999
☐	£1,762 - £1,957	£4,168 - £4,583	£50,000 - £54,999
☐	£1,958 - £2,153	£4,584 - £5,000	£55,000 - £59,999
☐	£2,154 - £2,349	£5,001 - £5,417	£60,000 - £64,999
☐	£2,350 - £2,545	£5,418 - £5,833	£65,000 - £69,999
☐	£2,546 - £2,741	£5,834 - £6,250	£70,000 - £74,999
☐	£2,742 - £2,937	£6,251 - £6,667	£75,000 - £79,999
☐	£2,938 - £3,133	£6,668 - £7,083	£80,000 - £84,999
☐	£3,134 - £3,329	£7,084 - £7,500	£85,000 - £89,999
☐	£3,330 - £3,525	£7,501 - £7,917	£90,000 - £94,999
☐	£3,526 - £3,721	£7,918 - £8,333	£95,000 - £99,999
☐	£3,722 - £3,917	£8,334 - £8,750	£100,000 or more
☐	£3,918 - £4,113	£8,751 - £9,167	
☐	£4,114 - £4,309	£9,168 - £9,583	
☐	£4,310 - £4,505	£9,584 - £10,000	
☐	£4,506 - £4,701	£10,001 - £10,417	
☐	£4,702 - £4,897	£10,418 - £10,833	
☐	£4,898 - £5,093	£10,834 - £11,250	
☐	£5,094 - £5,289	£11,251 - £11,667	
☐	£5,290 - £5,485	£11,668 - £12,083	
☐	£5,486 - £5,681	£12,084 - £12,500	
☐	£5,682 - £5,877	£12,501 - £12,917	
☐	£5,878 - £6,073	£12,918 - £13,333	
☐	£6,074 - £6,269	£13,334 - £13,750	
☐	£6,270 - £6,465	£13,751 - £14,167	
☐	£6,466 - £6,661	£14,168 - £14,583	
☐	£6,662 - £6,857	£14,584 - £15,000	
☐	£6,858 - £7,053	£15,001 - £15,417	
☐	£7,054 - £7,249	£15,418 - £15,833	
☐	£7,250 - £7,445	£15,834 - £16,250	
☐	£7,446 - £7,641	£16,251 - £16,667	
☐	£7,642 - £7,837	£16,668 - £17,083	
☐	£7,838 - £8,033	£17,084 - £17,500	
☐	£8,034 - £8,229	£17,501 - £17,917	
☐	£8,230 - £8,425	£17,918 - £18,333	
☐	£8,426 - £8,621	£18,334 - £18,750	
☐	£8,622 - £8,817	£18,751 - £19,167	
☐	£8,818 - £9,013	£19,168 - £19,583	
☐	£9,014 - £9,209	£19,584 - £20,000	
☐	£9,210 - £9,405	£20,001 - £20,417	
☐	£9,406 - £9,601	£20,418 - £20,833	
☐	£9,602 - £9,797	£20,834 - £21,250	
☐	£9,798 - £9,993	£21,251 - £21,667	
☐	£10,000 or more	£21,668 or more	
☐	Nothing		
☐	Don't know		
☐	Prefer not to say		

14. Please state your current home postcode: _____



Thank you for taking the time to complete this questionnaire.

Appendix 6.12: Movement Measurement in the Early Years measurement tool following co-development

Participant Number _____

Movement Measurement in the Early Years



Diary for the caregivers of a child aged 3 or 4 years old. This diary asks about the movement behaviours that make up the 24 hours of your child's day.

There are four parts to this diary:

Part 1: **General questions** – Your child's health and if they have any disabilities.

Part 2: **Physical activity**- Any bodily movement, including when your child is engaged in active play or is actively travelling.

Part 3: **Sedentary behaviour**- Activities when your child is sitting, reclining or lying whilst awake.

Part 4: **Sleep**- The time your child spends asleep, during the day and at night.

Please answer all questions as accurately as possible and try not to count the same activity more than once. Put a 0 if the activity did not happen.
There are no right or wrong answers and all responses are confidential.



This project is funded and supported by the National Institute for Health Research (NIHR) School for Public Health Research (SPHR), Grant Reference Number PD-SPH-2015. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Participant Number _____

Part 1: General Questions

1. Does your child have any physical or medical condition that affects their ability to play and be physically active?

Yes	No (Please proceed to Q2)

If yes, please state the physical or medical condition _____

2. Does your child have any medical sleep problems, such as night terrors?

Yes	No (Please proceed to Q3)

If yes, please state the sleep problem: _____

3. Is your child currently suffering from an illness that may affect their normal behaviours, including being active, movement, sitting or sleep?

Yes	No (Please proceed to Q4a)

If yes, please state the illness: _____

Participant Number _____

Part 2- Physical Activity

Part 2 of the diary is looking at the physical activity of your child, this is separated into different sections:

- Activity whilst outdoors
- Activity used as a method of transport
- Activity whilst indoors



Note: Some questions ask for activity at 'home' and 'school', by this we mean:

- **Home** – Any time when your child is not at school e.g. under parental/primary caregiver's care, at a friend or relatives house
- **School**- When your child is at pre-school, nursery, playgroup, in childcare, childminders etc.

Participant Number _____

4a. Please state how many hours and minutes your child spends actively playing **outdoors** in each of the following (activities may include: running around, jumping on a trampoline, climbing, skipping, throw/catch).

	Monday		Tuesday		Wednesday		Thursday		Friday		Saturday		Sunday	
	Home	School	Home	School	Home	School	Home	School	Home	School	Home	School	Home	School
	Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes	
Outdoor play at school														
Playing in the garden/yard														
Playing in the street														
Playing in a park / playground														
Playing in open outdoor spaces (e.g. beach, forest)														
Sport sessions (e.g. football)														
Out for a walk/cycle (not for travel- see Q5a)														

Participant Number _____

Other activity (please specify):														
Other activity (please specify):														

4b. Did any of these activities make your child 'huff and puff' or breathe harder? (Please circle)

Yes	No (Please proceed to Q5)

4c. If yes, please state how many hours/minutes of this activity made your child 'huff and puff' or breathe harder.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Total hours/minutes							

Participant Number _____

5a. Please state how many hours and minutes your child spends **actively travelling**, which could include travelling for leisure (e.g. to/from school, the shops, the park) each day.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes
Walking							
Cycling							
Scooter							
Other means of active transport (please specify):							

5b. Did any of these activities make your child 'huff and puff' or breathe harder? (Please circle)

Yes	No (Please proceed to Q6)

5c. If yes, please state how many hours/minutes of this activity made your child 'huff and puff' or breathe harder.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Total hours/minutes							

Participant Number _____

6a. Please state how many hours and minutes your child spends actively playing **indoors** (activities may include: dancing, running around, rough and tumble play, sit and ride push toys).

	Monday		Tuesday		Wednesday		Thursday		Friday		Saturday		Sunday	
	Home	School	Home	School	Home	School	Home	School	Home	School	Home	School	Home	School
	Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes	
Playing actively in the home (including physically active video games e.g. Wii Sports)														
Playing actively indoors at school														
Indoor play areas (e.g. soft play)														
Sport sessions (e.g. swimming, dance, motor skill classes)														
Other activity (please specify):														
Other activity (please specify):														

Participant Number _____

6b. Did any of these activities make your child ‘huff and puff’ or breathe harder? (Please circle)

Yes	No (Please proceed to Q7)

6c. If yes, please state how many hours/minutes of this activity made your child ‘huff and puff’ or breathe harder.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Total hours/minutes							

Any additional comments for this section on physical activity:

Participant Number _____

Part 3- Sedentary Behaviour

Part 3 of the diary is looking at the sedentary behaviour (activities when your child is sitting, lying or reclining whilst awake) of your child, this is separated into different sections:

- Screen based activities
- Other sedentary activities
- Sedentary when travelling

Note: Some questions ask for activity at 'home' and 'school', by this we mean:

- **Home** – Any time when your child is not at school e.g. under parental/primary caregiver's care, at a friend or relatives house
- **School**- When your child is at preschool, nursery, playgroup, in childcare, childminders etc.



Participant Number _____

7. Please state how long your child spends in **screen based** activities whilst in a sitting, reclining or lying position.

	Monday		Tuesday		Wednesday		Thursday		Friday		Saturday		Sunday	
	Home	School	Home	School	Home	School	Home	School	Home	School	Home	School	Home	School
	Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes	
Watching TV														
Using a computer/laptop														
Using a tablet/mobile phone														
Playing on a games console (e.g. PlayStation- not physically active games- see Q6a)														
Other (please specify):														

8. Please state at what time your child last uses a screen before going to bed (e.g. if child watches a film before bed).

	<i>Example</i>	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Time	7.15pm							

Participant Number _____

9. Please state how long your child spends **playing and in other activities** whilst sitting, reclining or lying, including **quiet or carpet time**.

	Monday		Tuesday		Wednesday		Thursday		Friday		Saturday		Sunday	
	Home	School	Home	School	Home	School	Home	School	Home	School	Home	School	Home	School
	Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes	
Playing with toys (including jigsaws/puzzles)														
Crafts (including colouring, painting, Play-Doh)														
Reading, storytelling, phonics														
Bath time														
Sat on the toilet/potty														
Sitting whilst eating														
Other (please specify):														
Other (please specify):														

Participant Number _____

10. Please state how long your child spends **seated whilst travelling**.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes
In the pushchair							
Being carried							
Sat in a car							
Sat on the bus/train/tram							
Other (please specify)							

Any additional comments for this section on sedentary behaviour:

Participant Number _____

Part 4- Sleep

Part 4 of the diary is looking at the sleep of your child, this is separated into different sections:

- Bed time and wake time
- Night time waking
- Day time sleep



Participant Number _____

11. Please write your child's usual bed time and wake up time.

11a. Bed time: _____:_____am/pm

11b. Wake up time: _____:_____am/pm

12. On which days of the week is this the case? (Please circle and provide a reason in the box if different)

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Difference and reason for difference							

13. From the time that your child goes to bed, how long does it take them to fall asleep?

Minutes: _____

14. From the time that your child wakes up, how long does it take them to get out of bed?

Minutes: _____

Participant Number _____

15. Please state how many times and for how long each time, that your child wakes up during their night time sleep.

	<i>Example</i>	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
How many times	2							
How long each time	20 minutes							

16. Please state how many times and for how long each time that your child naps during the day.

	<i>Example</i>	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
How many times	1							
How long each time	45 minutes							

Any additional comments for this section on sleep:



Thank you for completing this diary!



Appendix 6.13: Study pack letter

Study Pack**Movement Measurement in the Early Years (MoveMEY): Pilot Study**

Thank you for taking part in this study! We have received your consent form via email and you have completed the questionnaire providing information about you and your child.

Please find in this envelope:

1. A copy of MoveMEY – a 7 day diary used to look at the physical activity, sedentary behaviour (activities when your child is sitting, reclining or lying) and sleep of your pre-school aged child. **Please complete this over the next 7 days, we will then have a conversation about your experiences of completing the diary via video or phone call at the end of the week.**
2. Some pens.
3. Some stickers for your child.
4. A pre-paid envelope to return the diary following our video/phone call.

Following the interview/conversation at the end of the week, please send me the completed MoveMEY diary using the pre-paid envelope in this study pack. On completion of the study, I will send you £20 in Love2Shop vouchers using the address provided.

If you have any questions, have any difficulty completing the MoveMEY diary, or wish to withdraw at any point then please do not hesitate to contact me on:

Sophie Phillips (sophie.m.phillips@durham.ac.uk or call on 07989143478).

Appendix 6.14: Debrief form



Debriefing Sheet

Movement Measurement in the Early Years (MoveMEY): Pilot Study



Thank you very much for taking part in this study. This study is part of developing and testing a new measurement tool to assess the physical activity, sedentary behaviour (behaviours when your child is sitting, reclining or lying) and sleep of pre-school children. This will help further research in this area, so that associations between these behaviours and health and development can be better understood.

We will be analysing the interviews to determine if the tool is relevant, easy to understand, is representative of your child's behaviours, and to help determine if the tool is easy to use. The information you provided about you and your child (including age, ethnicity) will only be used to provide context to the information.

If you would like further information about the study or would like to know the findings when all the data has been collected and analysed, then please contact me using the contact details provided. All data will be securely stored, confidential and anonymised.

I will send you a £20 voucher to the address you provided, for your time taking part in this research.

If taking part in this study has raised any specific concerns or you have any further questions then please do feel free to contact me.

Contact Details:

Lead researcher: Sophie Phillips (sophie.m.phillips@durham.ac.uk)

Primary Supervisor's name: Professor Carolyn Summerbell
(carolyn.summerbell@durham.ac.uk)

Department of Sport and Exercise Science, Durham University.

Appendix 6.15: MoveMEY Scoring Sheet (Tested in Pilot)

Scoring System for MoveMEY Measurement Tool



Behaviour/Construct	Questions used for scoring	Additional Comments
General questions	1, 2, 3	General questions to be used as screening questions. If the answer is 'yes' to any of these questions, then child may be excluded or reported separately.
Physical Activity		
Total Daily Physical Activity (PA)	Q4a + Q5a+ Q6a	- Daily PA for each question calculated by sum of all hours and minutes of activity recorded each day. - The 'daily PA' time for all 3 individual questions will be summed to provide a total daily PA. - From this, will be able to determine how many days children are active and their average activity on those days (to establish if they meet the recommendation).
Total Daily Moderate to Vigorous Physical Activity (MVPA)	Q4c + Q5c + Q6c	- Daily MVPA calculated by the sum of the 3 questions. - From this, we can determine how many days children engage in MVPA and their average MVPA activity on those days (to establish if they meet the recommendation).

Sedentary Behaviour		
Total Daily Sedentary Behaviour (SB)	Q7 + Q9+ Q10	- Daily SB for each question calculated by sum of hours and minutes spent in each of the activities each day. - The 'daily SB for all 3 individual questions will be summed to provide a total daily SB. - From this, will be able to determine average SB of each day.
Total Daily Screen Time	Q7	- Sum of duration of hours and minutes spent in each activity, each day. - From this, can determine how many days children engage in screen time and average screen time use per day.
Screen time before bed	Q8 (time), Q11 (bed time)	- Hours/minutes difference between the time at which child last uses a screen (Q8), and the time child goes to bed (Q11). - From this can determine if/when children use a screen in the 1 hour leading up to bed time.
Not being restrained for more than 1 hour	Q10 (in the pushchair, sat in a car, 'other' where child would be restrained- please specify what activity this is).	Sum of duration of hours and minutes recorded for child being: 1) in the pushchair; 2) sat in a car; 3) other (please specify)- if

		this includes an activity where child would usually be restrained (please specify what activity this is), each day. From this, we can determine when children are restrained and average time per day.
When sedentary engaging in pursuits such as storytelling/reading (daily SB engaged in non-screen based activities)	Q9	- Sum of duration of hours and minutes spent in each activity, each day. - From this, we will be able to determine how many days children spend engaged in non screen based sedentary pursuits, and the average time spent on these per day.
Sleep		
10-13 hrs good quality sleep (may include a nap)	Q11a, Q11b – Q13 – Q15 + Q16	- Total hours/mins from bed time to wake time (Q11a, Q11b) – total mins to fall asleep (Q13) – total mins awake during night (Q15) + total nap duration (Q16). - From this, we will be able to determine the amount of time a child spends a sleep in a 24 hr day (and from this, whether they meet the recommendation).

Good quality sleep	Total time in bed= Q11a, Q11b +Q14 Total sleep time = Q11a, Q11b – Q13- Q15	• Total time in bed = total hours/minutes between bed time and wake time + mins to get out of bed • Total sleep time = total hours/mins between bed time and wake time – sleep latency – night time wakings. • Calculated by: total sleep time/ total time in bed x 100 • 'Good sleep quality' will be based on sleep efficiency – ratio of total sleep time to time in bed, $\geq 85\%$ = good sleep quality.
Consistent wake and bed time	Q12	≥ 5 days

Appendix 6.16: Movement Measurement in the Early Years measurement tool (Final version following development and pilot)

Participant Number _____

Movement Measurement in the Early Years



Diary for the caregivers of a child aged 3 or 4 years old. This diary asks about the movement behaviours that make up the 24 hours of your child's day.

There are four parts to this diary:

Part 1: **General questions** – Your child's health and if they have any disabilities.

Part 2: **Physical activity**- Any bodily movement, including when your child is engaged in active play or is actively travelling.

Part 3: **Sedentary behaviour**- Activities when your child is sitting, reclining or lying whilst awake.

Part 4: **Sleep**- The time your child spends asleep, during the day and at night.

Please complete this diary over the next 7 days. Please answer all questions as accurately as possible and try not to count the same activity more than once. Please put a 0 if the activity did not happen. There are no right or wrong answers and all responses are confidential.



This project is funded and supported by the National Institute for Health Research (NIHR) School for Public Health Research (SPHR), Grant Reference Number PD-SPH-2015. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Participant Number _____

Part 1: General Questions

1. Does your child have any physical, neurodevelopmental, or medical condition or disability that affects their ability to play and be physically active?

Yes	No (Please proceed to Q2)

If yes, please state the condition or disability _____

2. Does your child have any medical sleep problems, such as night terrors?

Yes	No (Please proceed to Q3)

If yes, please state the sleep problem: _____

3. Is your child currently suffering from an illness, unwell or poorly (short or long term) that may affect their normal behaviours, including being active, movement, sitting or sleep?

Yes	No (Please proceed to Q4a)

If yes, please state the illness: _____

Participant Number _____

Part 2- Physical Activity

Part 2 of the diary is looking at the physical activity of your child, this is separated into different sections:

- Activity whilst outdoors
- Activity used as a method of transport
- Activity whilst indoors



Note: Some questions ask for activity at 'home' and 'nursery', by this we mean:

- **Home** – Any time when your child is not at nursery e.g. under parental/primary caregiver's care, at a friend or relatives house
- **Nursery**- When your child is at pre-school, nursery, playgroup, in childcare, childminders etc.

Participant Number _____

4a. Please state how many hours and minutes your child spends actively playing **outdoors** in each of the following (activities may include: running around, jumping on a trampoline, climbing, skipping, throw/catch).

	Monday		Tuesday		Wednesday		Thursday		Friday		Saturday		Sunday	
	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery
	Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes	
<i>Example</i>	1 hour, 15 minutes	2 hours	1 hour	1 hour 30 minutes	1 hour, 40 minutes	0	50 minutes	2 hours 20 minutes	2 hours 20 minutes	0	3 hours 15 minutes	0	1 hour	0
Outdoor play at nursery														
Playing in the garden/yard														
Playing in the street														
Playing in a park / playground														
Playing in open outdoor spaces (e.g. beach, forest)														
Sport sessions (e.g. football)														

Participant Number _____

Out for a walk/cycle (not for travel- see Q5a)														
Other activity (please specify):														
Other activity (please specify):														

4b. Did any of these activities make your child ‘huff and puff’, breathe harder, breathless, hot and sweaty, or need a drink or rest (Please circle)

Yes	No (Please proceed to Q5)

4c. If yes, please state how many hours/minutes of this activity made your child ‘huff and puff’, breathe harder, breathless, hot and sweaty, or need a drink or rest.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Total hours/minutes							

Participant Number _____

5a. Please state how many hours and minutes your child spends **actively travelling**, which could include travelling for leisure (e.g. to/from school, the shops, the park) each day.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes
Walking							
Cycling							
Scooter							
Other means of active transport (please specify):							

5b. Did any of these activities make your child 'huff and puff', breathe harder, breathless, hot and sweaty, or need a drink or rest.

Yes	No (Please proceed to Q6)

5c. If yes, please state how many hours/minutes of this activity made your child 'huff and puff', breathe harder, breathless, hot and sweaty, or need a drink or rest.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Total hours/minutes							

Participant Number _____

6a. Please state how many hours and minutes your child spends actively playing **indoors** (activities may include: dancing, running around, rough and tumble play, sit and ride push toys).

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes
Playing actively in the home (including physically active video games e.g. Wii Sports or online exercise videos e.g. yoga, Zumba)							
Playing actively indoors at nursery							
Indoor play areas (e.g. soft play)							
Sport sessions (e.g. swimming, dance, motor skill classes)							
Other activity (please specify):							
Other activity (please specify):							

Participant Number _____

6b. Did any of these activities make your child ‘huff and puff’, breathe harder, breathless, hot and sweaty, or need a drink or rest. E.g. at soft play (Please circle)

Yes	No (Please proceed to Q7)

6c. If yes, please state how many hours/minutes of this activity made your child ‘huff and puff’, breathe harder, breathless, hot and sweaty, or need a drink or rest.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Total hours/minutes							

Any additional comments for this section on physical activity:

Participant Number _____

Part 3- Sedentary Behaviour

Part 3 of the diary is looking at the sedentary behaviour (activities when your child is sitting, lying or reclining whilst awake) of your child. This is separated into different sections:

- Screen based activities
- Other sedentary activities
- Sedentary when travelling



Note: Some questions ask for activity at 'home' and 'nursery', by this we mean:

- **Home** – Any time when your child is not at school e.g. under parental/primary caregiver's care, at a friend or relatives house
- **Nursery**- When your child is at preschool, nursery, playgroup, in childcare, childminders etc.

Participant Number _____

7. Please state how long your child spends in **screen based** activities whilst in a sitting, reclining or lying position.

	Monday		Tuesday		Wednesday		Thursday		Friday		Saturday		Sunday	
	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery
	Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes	
Watching TV														
Using a computer/laptop														
Using a tablet/mobile phone														
Playing on a games console (e.g. PlayStation- not physically active games- see Q6a)														
Other (please specify):														

8. Please state at what time your child last uses a screen before going to bed (e.g. if child watches a film before bed).

	<i>Example</i>	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Time	7.15pm							

Participant Number _____

9. Please state how long your child spends **playing and in other activities** whilst sitting, reclining or lying, including **quiet or carpet time**.

	Monday		Tuesday		Wednesday		Thursday		Friday		Saturday		Sunday	
	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery	Home	Nursery
	Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes		Total hours/minutes	
Playing with toys (including jigsaws/puzzles)														
Crafts (including colouring, painting, Play-Doh)														
Reading, storytelling, phonics														
Bath time														
Sat on the toilet/potty														
Sitting whilst eating														
Other (please specify):														
Other (please specify):														

Participant Number _____

10. Please state how long your child spends **seated whilst travelling**.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes	Total hours/minutes
In the pushchair							
Being carried							
Sat in a car							
Sat on the bus/train/tram							
Other (please specify)							

Any additional comments for this section on sedentary behaviour:

Participant Number _____

Part 4- Sleep

Part 4 of the diary is looking at the sleep of your child, this is separated into different sections:

- Bed time, sleep time, wake time, out of bed time
- Night time waking
- Day time sleep



Participant Number _____

11. Please write the time your child goes to bed, the time they fall asleep, the time they wake up, and the time they get out of bed each day.

	<i>Example</i>	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Bed time (Time your child goes to bed)	<i>7.00pm</i>							
Sleep time (Time your child falls asleep)	<i>7.15pm</i>							
Morning wake time (Time your child wakes up)	<i>7.00am</i>							
Out of bed time (Time your child gets out of bed after their wake up time)	<i>7.05am</i>							

Participant Number _____

12. Please state how many times and for how long each time, that your child wakes up during their night time sleep.

	<i>Example</i>	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
How many times	2							
How long each time	20 minutes							

13. Please state how many times and for how long each time that your child naps during the day.

	<i>Example</i>	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
How many times	1							
How long each time	45 minutes							

Any additional comments for this section on sleep:



Thank you for completing this diary!



Appendix 6.17: MoveMEY Scoring Sheet (Final)

Scoring System for MoveMEY Measurement Tool



Behaviour/Construct	Questions used for scoring	Additional Comments
General questions	1, 2, 3	General questions to be used as screening questions. If the answer is 'yes' to any of these questions, then child may be excluded or reported separately.
Physical Activity		
Total Daily Physical Activity (PA)	Q4a + Q5a+ Q6a	- Daily PA for each question calculated by sum of all hours and minutes of activity recorded each day. - The 'daily PA' time for all 3 individual questions will be summed to provide a total daily PA. - From this, will be able to determine how many days children are physically active and their average physical activity on those days (to establish if they meet the recommendation).
Total Daily Moderate to Vigorous Physical Activity (MVPA)	Q4c + Q5c + Q6c	- Daily MVPA calculated by the sum of the 3 questions. - From this, we can determine how many days children engage in MVPA and their average MVPA activity on those days (to establish if they meet the recommendation).
Sedentary Behaviour		

Total Daily Sedentary Behaviour (SB)	Q7 + Q9+ Q10	- Daily SB for each question calculated by sum of hours and minutes spent in each of the activities each day. - The 'daily SB' for all 3 individual questions will be summed to provide a total daily SB. - From this, will be able to determine average SB of each day.
Total Daily Screen Time	Q7	- Sum of duration of hours and minutes spent in each activity, each day. - From this, can determine how many days children engage in screen time and average screen time use per day.
Screen time before bed	Q8 (time), Q11 (bedtime)	- Hours/minutes difference between the time at which child last uses a screen (Q8), and the time child goes to bed (Q11). - From this can determine if/when children use a screen in the 1 hour leading up to bedtime.
Not being restrained for more than 1 hour	Q10 (in the pushchair, sat in a car, 'other' where child would be restrained- please specify what activity this is).	- Sum of duration of hours and minutes recorded for child being: 1) in the pushchair; 2) sat in a car; 3) other (please specify)- if this includes an activity where child would usually be restrained (please specify what activity this is), each day. - From this, we can determine when children are restrained and average time per day.

When sedentary engaging in pursuits such as storytelling/reading (daily ST engaged in non-screen based activities)	Q9	- Sum of duration of hours and minutes spent in each activity, each day. - From this, we will be able to determine how many days children spend engaged in non-screen based sedentary pursuits, and the average time spent on these per day.
Sleep		
10-13 hrs good quality sleep (may include a nap)	Q11 - Q12 + Q13	- Total hours/mins from sleep time to wake time – total mins awake during night (Q12) + total nap duration (Q13). - From this, we will be able to determine the amount of time a child spends asleep in a 24 hr day (and from this, whether they meet the recommendation).
Good quality sleep	Total time in bed= Q11 bedtime to out of bedtime Total sleep time = Q11 sleep time to wake time Q11 –Q12+Q13	- Total time in bed = total hours/minutes between bedtime and out of bedtime - Total sleep time = total hours/mins between sleep time and wake time – night-time waking + nap time - Calculated by total sleep time divided by total time in bed x 100 'Good sleep quality' will be based on sleep efficiency – ratio of total sleep time to time in bed, $\geq 85\%$ = good sleep quality.

Consistent sleep and wake time	Q11	≥5 days for sleep time and morning wake time. Consistent classified as within 30 minutes.

References

A Better Start. (2021). *Better Start Bradford*. Available at:

<https://www.betterstartbradford.org.uk/>. (Accessed: 25th October 2021).

Acebo, C., Sadeh, A., Seifer, R., Tzischinsky, O., Hafer, A., & Carskadon, M. A. (2005).

Sleep/wake patterns derived from activity monitoring and maternal report for healthy 1-to 5-year-old children. *Sleep*, 28(12), 1568-1577.

Acebo, C., Sadeh, A., Seifer, R., Tzischinsky, O., Wolfson, A. R., Hafer, A., & Carskadon, M. A.

(1999). Estimating sleep patterns with activity monitoring in children and adolescents: how many nights are necessary for reliable measures?. *Sleep*, 22(1), 95-103.

Actigraph Manual. (2013). *GT3X+ and wGT3X+ Device Manual*. Available at:

<http://s3.amazonaws.com/actigraphcorp.com/wp-content/uploads/2018/02/22094126/GT3X-wGT3X-Device-Manual-110315.pdf>
(Accessed: 10th February 2022).

Adamo, K. B., Prince, S. A., Tricco, A. C., Connor-Gorber, S., & Tremblay, M. (2009). A

comparison of indirect versus direct measures for assessing physical activity in the pediatric population: a systematic review. *International Journal of Pediatric Obesity*, 4(1), 2-27.

Adams, J. (2010). Prevalence and socio-demographic correlates of “active transport” in the

UK: analysis of the UK time use survey 2005. *Preventive Medicine*, 50(4), 199-203.

Adams, S. A., Matthews, C. E., Ebbeling, C. B., Moore, C. G., Cunningham, J. E., Fulton, J., &

Hebert, J. R. (2005). The effect of social desirability and social approval on self-reports of physical activity. *American Journal of Epidemiology*, 161(4), 389-398.

- Addy, C. L., Trilk, J. L., Dowda, M., Byun, W., & Pate, R. R. (2014). Assessing preschool children's physical activity: how many days of accelerometry measurement. *Pediatric Exercise Science*, 26(1), 103-109.
- Adolph, A. L., Puyau, M. R., Vohra, F. A., Nicklas, T. A., Zakeri, I. F., & Butte, N. F. (2012). Validation of uniaxial and triaxial accelerometers for the assessment of physical activity in preschool children. *Journal of Physical Activity & Health*, 9(7), 944-953
- Aguilar-Farias, N., Toledo-Vargas, M., Miranda-Marquez, S., Cortinez-O'Ryan, A., Cristi-Montero, C., Rodriguez-Rodriguez, F., ... & del Pozo Cruz, B. (2020). Sociodemographic predictors of changes in physical activity, screen time, and sleep among toddlers and preschoolers in Chile during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, 18(1), 176.
- Aguilar-Farias, N., Toledo-Vargas, M., Miranda-Marquez, S., Cortinez-O'Ryan, A., Martino-Fuentealba, P., Cristi-Montero, C., ... & Okely, A. D. (2021). Associations between movement behaviors and emotional changes in toddlers and preschoolers during early stages of the COVID-19 pandemic in Chile. *Frontiers in Pediatrics*, 9:667362.
- Ainsworth, B., Cahalin, L., Buman, M., & Ross, R. (2015). The current state of physical activity assessment tools. *Progress in Cardiovascular Diseases*, 57(4), 387-395.
- Alghaeed, Z., Reilly, J. J., Chastin, S. F., Martin, A., Davies, G., & Paton, J. Y. (2013). The influence of minimum sitting period of the ActivPAL™ on the measurement of breaks in sitting in young children. *PloS one*, 8(8), e71854.
- Alhassan, S., Sirard, J. R., Kurdziel, L. B., Merrigan, S., Greever, C., & Spencer, R. M. (2017). Cross-validation of two accelerometers for assessment of physical activity and sedentary time in preschool children. *Pediatric Exercise Science*, 29(2), 268-277.

- Alonso-Martínez, A. M., Ramírez-Vélez, R., García-Alonso, Y., Izquierdo, M., & García-Hermoso, A. (2021). Physical activity, sedentary behavior, sleep and self-regulation in Spanish preschoolers during the COVID-19 lockdown. *International Journal of Environmental Research and Public Health*, 18(2), 693.
- Althoff, T., Sosič, R., Hicks, J. L., King, A. C., Delp, S. L., & Leskovec, J. (2017). Large-scale physical activity data reveal worldwide activity inequality. *Nature*, 547(7663), 336-339.
- Atkin, A. J., Gorely, T., Clemes, S. A., Yates, T., Edwardson, C., Brage, S., ... & Biddle, S. J. (2012). Methods of measurement in epidemiology: sedentary behaviour. *International Journal of Epidemiology*, 41(5), 1460-1471.
- Arts, J., Gubbels, J. S., Verhoeff, A. P., Chinapaw, M., Lettink, A., & Altenburg, T. M. (2022). A systematic review of proxy-report questionnaires assessing physical activity, sedentary behavior and/or sleep in young children (aged 0–5 years). *International Journal of Behavioral Nutrition and Physical Activity*, 19, 18.
- Aubert, S., Brazo-Sayavera, J., González, S. A., Janssen, I., Manyanga, T., Oyeyemi, A. L., ... & Tremblay, M. S. (2021). Global prevalence of physical activity for children and adolescents; inconsistencies, research gaps, and recommendations: a narrative review. *International Journal of Behavioral Nutrition and Physical Activity*, 18, 81.
- Australian Bureau of Statistics. (2013). *Australian Health Survey: Physical Activity*. Available at: <https://www.abs.gov.au/statistics/health/health-conditions-and-risks/australian-health-survey-physical-activity/latest-release>. (Accessed: 2nd December 2021).
- Australian Bureau of Statistics. (2016). *National Health Survey: First Results, 2014-15*. Available at: <https://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/EFC93ECE5ABE3894CA2>

[57F1400133CE1/\\$File/national%20health%20survey%202014-15%20questionnaire.pdf](#) (Accessed: 2nd December 2021).

Australian Bureau of Statistics. (2018). *National Health Survey: First Results*. Available at: <https://www.abs.gov.au/statistics/health/health-conditions-and-risks/national-health-survey-first-results/latest-release#key-statistics>. (Accessed: 2nd December 2021).

Bacardi-Gascón, M., Reveles-Rojas, C., Woodward-Lopez, G., Crawford, P. and Jiménez-Cruz, A., 2012. Assessing the validity of a physical activity questionnaire developed for parents of preschool children in Mexico. *Journal of Health, Population, and Nutrition*, 30(4), p.439.

Bambra, C., Barr, B., & Milne, E. (2014). North and South: addressing the English health divide. *Journal of Public Health*, 36(2), 183-186.

Banda, J. A., Haydel, K. F., Davila, T., Desai, M., Bryson, S., Haskell, W. L., ... & Robinson, T. N. (2016). Effects of varying epoch lengths, wear time algorithms, and activity cut-points on estimates of child sedentary behavior and physical activity from accelerometer data. *PloS one*, 11(3), e0150534.

Bao, X., Qu, H., Zhang, R., & Hogan, T. P. (2020). Modeling reading ability gain in kindergarten children during COVID-19 school closures. *International Journal of Environmental Research and Public Health*, 17(17), 6371.

Barber, S. E., Akhtar, S., Jackson, C., Bingham, D. D., Hewitt, C., Routen, A., ... & Wright, J. (2015). Preschoolers in the playground: a pilot cluster randomised controlled trial of a physical activity intervention for children aged 18 months to 4 years. NIHR Journals Library, Southampton (UK); PMID: 25950067.

- Barber, S. E., Jackson, C., Hewitt, C., Ainsworth, H. R., Buckley, H., Akhtar, S., ... & Wright, J. (2016). Assessing the feasibility of evaluating and delivering a physical activity intervention for pre-school children: a pilot randomised controlled trial. *Pilot and Feasibility Studies*, 2(1), 12.
- Bar-Or, T., Bar-Or, O., Waters, H., Hirji, A., & Russell, S. (1996). Validity and social acceptability of the Polar Vantage XL for measuring heart rate in preschoolers. *Pediatric Exercise Science*, 8(2), 115-121.
- Basterfield, L., Adamson, A. J., Parkinson, K. N., Maute, U., Li, P. X., Reilly, J. J., & Gateshead Millennium Study Core Team. (2008). Surveillance of physical activity in the UK is flawed: validation of the Health Survey for England Physical Activity Questionnaire. *Archives of disease in childhood*, 93(12), 1054-1058.
- Bauer, K. M., & Blunden, S. (2008). How accurate is subjective reporting of childhood sleep patterns? A review of the literature and implications for practice. *Current Pediatric Reviews*, 4(2), 132-142.
- Bauman, A., Pedišić, Ž., & Bragg, K. (2016). Objective measurement in physical activity surveillance: present role and future potential. In *The objective monitoring of physical activity: Contributions of accelerometry to epidemiology, exercise science and rehabilitation*. Springer, Cham, 347-367.
- Bauman, A., Phongsavan, P., Schoeppe, S., & Owen, N. (2006). Physical activity measurement-a primer for health promotion. *Promotion & Education*, 13(2), 92-103.
- Bélanger, M. È., Bernier, A., Paquet, J., Simard, V., & Carrier, J. (2013). Validating actigraphy as a measure of sleep for preschool children. *Journal of Clinical Sleep Medicine*, 9(7), 701-706.

- Bennett, D. H., Lee, K., Cassady, D. L., Ritz, B., & Hertz-Picciotto, I. (2012). Feasibility of using web surveys to collect time–activity data. *Journal of Exposure Science & Environmental Epidemiology*, 22(2), 116-125.
- Bentley, G. F., Jago, R., & Turner, K. M. (2015). Mothers' perceptions of the UK physical activity and sedentary behaviour guidelines for the early years (Start Active, Stay Active): a qualitative study. *Bmj Open*, 5(9), e008383.
- Berglind, D., Ljung, R., Tynelius, P., & Brooke, H. L. (2018). Cross-sectional and prospective associations of meeting 24-h movement guidelines with overweight and obesity in preschool children. *Pediatric Obesity*, 13(7), 442-449.
- Biddle, S. J., Bennie, J. A., De Cocker, K., Dunstan, D., Gardiner, P. A., Healy, G. N., ... & Vincent, G. E. (2019). Controversies in the science of sedentary behaviour and health: insights, perspectives and future directions from the 2018 Queensland Sedentary Behaviour Think Tank. *International Journal of Environmental Research and Public Health*, 16(23), 4762.
- Biddle, S. J., Gorely, T., & Marshall, S. J. (2009). Is television viewing a suitable marker of sedentary behavior in young people?. *Annals of Behavioral Medicine*, 38(2), 147-153.
- Biddle, S. J., Gorely, T., Pearson, N., & Bull, F. C. (2011). An assessment of self-reported physical activity instruments in young people for population surveillance: Project ALPHA. *International Journal of Behavioral Nutrition and Physical Activity*, 8(1).
- Biddle, S. J., Pearson, N., Ross, G. M., & Braithwaite, R. (2010). Tracking of sedentary behaviours of young people: a systematic review. *Preventive medicine*, 51(5), 345-351.

- Bingham, D. D., Collings, P. J., Clemes, S. A., Costa, S., Santorelli, G., Griffiths, P., & Barber, S. E. (2016a). Reliability and validity of the early years physical activity questionnaire (EY-PAQ). *Sports*, 4(2), 30.
- Bingham, D. D., Costa, S., Clemes, S. A., Routen, A. C., Moore, H. J., & Barber, S. E. (2016b). Accelerometer data requirements for reliable estimation of habitual physical activity and sedentary time of children during the early years-a worked example following a stepped approach. *Journal of Sports Sciences*, 34(20), 2005-2010.
- Biswas, A., Oh, P. I., Faulkner, G. E., Bajaj, R. R., Silver, M. A., Mitchell, M. S., & Alter, D. A. (2015). Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Annals of Internal Medicine*, 162(2), 123-132.
- Bikchu, C.H., 2014. Convergent validity of the electronic pedometer with the TriTrac RT3 accelerometer for measuring structured play activities in preschool children. *Journal of Biological Education*, 11, 10.
- Boshuizen, H. C., Viet, A. L., Picavet, H. S. J., Botterweck, A., & Van Loon, A. J. M. (2006). Non-response in a survey of cardiovascular risk factors in the Dutch population: determinants and resulting biases. *Public Health*, 120(4), 297-308.
- Bouchard, C., Blair, S. N., & Haskell, W. L. (2012). *Physical activity and health*. Human Kinetics.
- Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., ... & Fernandez, M. (2009). How we design feasibility studies. *American Journal of Preventive Medicine*, 36(5), 452-457.
- Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. Sage.

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Braun, V., & Clarke, V. (2014). What can “thematic analysis” offer health and wellbeing researchers?. *International Journal of Qualitative Studies on Health and Well-being*, 9(1), 26152.
- Brodie, M. A., Pliner, E. M., Ho, A., Li, K., Chen, Z., Gandevia, S. C., & Lord, S. R. (2018). Big data vs accurate data in health research: large-scale physical activity monitoring, smartphones, wearable devices and risk of unconscious bias. *Medical Hypotheses*, 119, 32-36.
- Brønd, J. C., & Arvidsson, D. (2016). Sampling frequency affects the processing of Actigraph raw acceleration data to activity counts. *Journal of Applied Physiology*, 120(3), 362-369.
- Brown, T., Moore, T. H., Hooper, L., Gao, Y., Zayegh, A., Ijaz, S., ... & Summerbell, C. D. (2019). Interventions for preventing obesity in children. *Cochrane Database of Systematic Reviews*, (7).
- Brown, W. H., Pfeiffer, K. A., McIver, K. L., Dowda, M., Almeida, J. M., & Pate, R. R. (2006). Assessing preschool children's physical activity: the Observational System for Recording Physical Activity in children-preschool version. *Research Quarterly for Exercise and Sport*, 77(2), 167-176.
- Brownson, R. C., Jones, D. A., Pratt, M., Blanton, C., & Heath, G. W. (2000). Measuring physical activity with the behavioral risk factor surveillance system. *Medicine and Science in Sports and Exercise*, 32(11), 1913-1918.

- Bryant, M. J., Lucove, J. C., Evenson, K. R., & Marshall, S. (2007). Measurement of television viewing in children and adolescents: a systematic review. *Obesity Reviews*, 8(3), 197-209.
- Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., ... & Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, 54(24), 1451-1462.
- Buman, M. P., Winkler, E. A., Kurka, J. M., Hekler, E. B., Baldwin, C. M., Owen, N., ... & Gardiner, P. A. (2014). Reallocating time to sleep, sedentary behaviors, or active behaviors: associations with cardiovascular disease risk biomarkers, NHANES 2005–2006. *American journal of epidemiology*, 179(3), 323-334.
- Bunn, J. A., Navalta, J. W., Fountaine, C. J., & Reece, J. D. (2018). Current state of commercial wearable technology in physical activity monitoring 2015–2017. *International Journal of Exercise Science*, 11(7), 503-515.
- Buyse, D. J. (2014). Sleep health: can we define it? Does it matter?. *Sleep*, 37(1), 9-17.
- Byrne, R., Terranova, C. O., & Trost, S. G. (2021). Measurement of screen time among young children aged 0–6 years: A systematic review. *Obesity Reviews*, 22(8), e13260.
- Byun, W., Beets, M. W., & Pate, R. R. (2015). Sedentary behavior in preschoolers: how many days of accelerometer monitoring is needed?. *International Journal of Environmental Research and Public Health*, 12(10), 13148-13161.
- Byun, W., Kim, Y., & Brusseau, T. A. (2018a). The use of a fitbit device for assessing physical activity and sedentary behavior in preschoolers. *The Journal of Pediatrics*, 199, 35-40.
- Byun, W., Lee, J. M., Kim, Y., & Brusseau, T. A. (2018b). Classification accuracy of a wearable activity tracker for assessing sedentary behavior and physical activity in 3–5-year-old

- children. *International Journal of Environmental Research and Public Health*, 15(4), 594.
- Cain, K. L., Sallis, J. F., Conway, T. L., Van Dyck, D., & Calhoun, L. (2013). Using accelerometers in youth physical activity studies: a review of methods. *Journal of Physical Activity and Health*, 10(3), 437-450.
- Cardon, G., & De Bourdeaudhuij, I. (2007). Comparison of pedometer and accelerometer measures of physical activity in preschool children. *Pediatric exercise science*, 19(2), 205-214.
- Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L., & Morin, C. M. (2012). The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep*, 35(2), 287-302.
- Carskadon, M. A., & Dement, W. C. (2005). Normal human sleep: an overview. *Principles and practice of sleep medicine*, 4(1), 13-23.
- Carson, V., Ezeugwu, V. E., Tamana, S. K., Chikuma, J., Lefebvre, D. L., Azad, M. B., ... & Mandhane, P. J. (2019). Associations between meeting the Canadian 24-Hour Movement Guidelines for the Early Years and behavioral and emotional problems among 3-year-olds. *Journal of Science and Medicine in Sport*, 22(7), 797-802.
- Carson, V., Kuzik, N., Hunter, S., Wiebe, S. A., Spence, J. C., Friedman, A., ... & Hinkley, T. (2015). Systematic review of sedentary behavior and cognitive development in early childhood. *Preventive Medicine*, 78, 115-122.
- Carson, V., Lee, E. Y., Hewitt, L., Jennings, C., Hunter, S., Kuzik, N., ... & Tremblay, M. S. (2017). Systematic review of the relationships between physical activity and health indicators in the early years (0-4 years). *BMC Public Health*, 17(5), 33-63.

- Carson, V., Rahman, A. A., & Wiebe, S. A. (2017). Associations of subjectively and objectively measured sedentary behavior and physical activity with cognitive development in the early years. *Mental Health and Physical Activity, 13*, 1-8.
- Carson, V., Tremblay, M. S., & Chastin, S. F. (2017). Cross-sectional associations between sleep duration, sedentary time, physical activity, and adiposity indicators among Canadian preschool-aged children using compositional analyses. *BMC Public Health, 17*(5), 123-131.
- Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports, 100*(2), 126.
- Chaput, J. P., Colley, R. C., Aubert, S., Carson, V., Janssen, I., Roberts, K. C., & Tremblay, M. S. (2017a). Proportion of preschool-aged children meeting the Canadian 24-Hour Movement Guidelines and associations with adiposity: results from the Canadian Health Measures Survey. *BMC Public Health, 17*(5), 829.
- Chaput, J. P., Dutil, C., & Sampasa-Kanyinga, H. (2018). Sleeping hours: what is the ideal number and how does age impact this?. *Nature and Science of Sleep, 10*, 421-430.
- Chaput, J. P. (2019). The integration of pediatric sleep health into public health in Canada. *Sleep Medicine, 56*, 4-8.
- Chaput, J. P., Saunders, T. J., & Carson, V. (2017). Interactions between sleep, movement and other non-movement behaviours in the pathogenesis of childhood obesity. *Obesity Reviews, 18*, 7-14.
- Chaput, J. P., & Shiao, J. (2019). Routinely assessing patients' sleep health is time well spent. *Preventive Medicine Reports, 14*, 100851.

- Chaput, J. P., Carson, V., Gray, C. E., & Tremblay, M. S. (2014a). Importance of all movement behaviors in a 24 hour period for overall health. *International Journal of Environmental Research and Public Health*, 11(12), 12575-12581.
- Chaput, J. P., Gray, C. E., Poitras, V. J., Carson, V., Gruber, R., Birken, C. S., ... & Tremblay, M. S. (2017b). Systematic review of the relationships between sleep duration and health indicators in the early years (0–4 years). *BMC Public Health*, 17(5), 855.
- Chaput, J. P., Leduc, G., Boyer, C., Belanger, P., LeBlanc, A. G., Borghese, M. M., & Tremblay, M. S. (2014b). Objectively measured physical activity, sedentary time and sleep duration: independent and combined associations with adiposity in canadian children. *Nutrition & Diabetes*, 4(6), e117-e117.
- Chaput, J. P., Saunders, T. J., & Carson, V. (2017c). Interactions between sleep, movement and other non-movement behaviours in the pathogenesis of childhood obesity. *Obesity Reviews*, 18, 7-14.
- Chastin, S. F., Palarea-Albaladejo, J., Dontje, M. L., & Skelton, D. A. (2015). Combined effects of time spent in physical activity, sedentary behaviors and sleep on obesity and cardio-metabolic health markers: a novel compositional data analysis approach. *PloS One*, 10(10), e0139984.
- Chen, B., Bernard, J. Y., Padmapriya, N., Yao, J., Goh, C., Tan, K. H., ... & Müller-Riemenschneider, F. (2019). Socio-demographic and maternal predictors of adherence to 24-hour movement guidelines in Singaporean children. *International Journal of Behavioral Nutrition and Physical Activity*, 16(1), 70.
- Chen, S., Hong, J., Wang, G., & Liu, Y. (2021). Move More, Sit Less and Sleep Well: An analysis of WHO movement guidelines for children under 5 years of age. *Sports Medicine and Health Science*, 3(1), 54-57.

- Chen, X., Sekine, M., Hamanishi, S., Wang, H., Hayashikawa, Y., Yamagami, T., & Kagamimori, S. (2002). The validity of nursery teachers' report on the physical activity of young children. *Journal of Epidemiology*, 12(5), 367-374.
- Chia, M. Y. H., Tay, L. Y., & Chua, T. B. K. (2020). Quality of life and meeting 24-h WHO guidelines among preschool children in Singapore. *Early Childhood Education Journal*, 48(3), 313-323.
- Chia, M., Tay, L. Y., & Chua, T. B. K. (2019). The development of an online surveillance of digital media use in early childhood questionnaire-SMALLQ™-for Singapore. *Montenegrin Journal of Sports Science and Medicine*, 8(2), 77-80.
- Chief Medical Officers. (2019). *UK Chief Medical Officers' Physical Activity Guidelines*. Available at: [UK Chief Medical Officers' Physical Activity Guidelines \(publishing.service.gov.uk\)](https://www.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/821117/uk-cmo-physical-activity-guidelines.pdf). (Accessed: 20th January 2021).
- Chinapaw, M. J., Mokkink, L. B., van Poppel, M. N., van Mechelen, W., & Terwee, C. B. (2010). Physical activity questionnaires for youth. *Sports Medicine*, 40(7), 539-563.
- Chinapaw, M. J. M., Proper, K. I., Brug, J., Van Mechelen, W., & Singh, A. S. (2011). Relationship between young peoples' sedentary behaviour and biomedical health indicators: a systematic review of prospective studies. *Obesity Reviews*, 12(7), e621-e632.
- Choi, L., Liu, Z., Matthews, C. E., & Buchowski, M. S. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and Science in Sports and Exercise*, 43(2), 357.
- Chow, B. C., & Yung Ching, A. (2009). Validation of Parent Proxy Report with Pedometry on Preschool Children's Physical activity. *Asian Journal of Physical Education & Recreation*, 15(2), 65-74.

- Chu, E. Y., McManus, A. M., & Yu, C. C. (2007). Calibration of the RT3 accelerometer for ambulation and nonambulation in children. *Medicine and Science in Sports and Exercise*, 39(11), 2085-2091.
- Clarke, J., Kipping, R., Chambers, S., Willis, K., Taylor, H., Brophy, R., ... & Langford, R. (2021). Impact of COVID-19 restrictions on preschool children's eating, activity and sleep behaviours: a qualitative study. *BMJ Open*, 11(10), e051497.
- Clarke, V. & Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology*, 12, 297-298.
- Clevenger, K. A., Pfeiffer, K. A., Mackintosh, K. A., McNarry, M. A., Brønd, J., Arvidsson, D., & Montoye, A. H. (2019). Effect of sampling rate on acceleration and counts of hip-and wrist-worn ActiGraph accelerometers in children. *Physiological Measurement*, 40(9), 095008.
- Cliff, D., Janssen, I., Kemp, B., Green, J., Draper, C., Silva, D.A.S., ... & Santos, R. (2021). *Early Years Accelerometry – The SADEY (Sleep and Activity Database for the Early Years) project*. Available at: <https://www.uow.edu.au/global-challenges/living-well-longer/early-years-accelerometry/>. (Accessed: 2nd July 2021).
- Cliff, D. P., McNeill, J., Vella, S. A., Howard, S. J., Santos, R., Batterham, M., ... & de Rosnay, M. (2017). Adherence to 24-hour movement guidelines for the early years and associations with social-cognitive development among Australian preschool children. *BMC Public Health*, 17(5), 857.
- Cliff, D. P., Reilly, J. J., & Okely, A. D. (2009). Methodological considerations in using accelerometers to assess habitual physical activity in children aged 0–5 years. *Journal of Science and Medicine in Sport*, 12(5), 557-567.

- Clifford Astbury, C., Foley, L., Penney, T. L., & Adams, J. (2020). How does time use differ between individuals who do more versus less foodwork? A compositional data analysis of time use in the United Kingdom time use survey 2014–2015. *Nutrients*, *12*(8), 2280.
- Cole, R. J., Kripke, D. F., Gruen, W., Mullaney, D. J., & Gillin, J. C. (1992). Automatic sleep/wake identification from wrist activity. *Sleep*, *15*(5), 461-469.
- Colley, R. C., Harvey, A., Grattan, K. P., & Adamo, K. B. (2014). Impact of accelerometer epoch length on physical activity and sedentary behaviour outcomes for preschool-aged children. *Health Reports*, *25*(1), 3-9.
- Corder, K., Ekelund, U., Steele, R. M., Wareham, N. J., & Brage, S. (2008). Assessment of physical activity in youth. *Journal of Applied Physiology*, *105*(3), 977-987.
- Corder, K., van Sluijs, E. M., Wright, A., Whincup, P., Wareham, N. J., & Ekelund, U. (2009). Is it possible to assess free-living physical activity and energy expenditure in young people by self-report?. *The American Journal of Clinical Nutrition*, *89*(3), 862-870.
- Costa, S., Barber, S. E., Griffiths, P. L., Cameron, N., & Cledes, S. A. (2013). Qualitative feasibility of using three accelerometers with 2–3-year-old children and both parents. *Research Quarterly for Exercise and Sport*, *84*(3), 295-304.
- Craigie, A. M., Lake, A. A., Kelly, S. A., Adamson, A. J., & Mathers, J. C. (2011). Tracking of obesity-related behaviours from childhood to adulthood: a systematic review. *Maturitas*, *70*(3), 266-284.
- Crouter, S. E., Schneider, P. L., Karabulut, M., & Bassett, D. R. (2003). Measuring steps, distance, and energy cost. *Medicine and Science in Sports and Exercise*, *35*(8), 1455-1460.

- Davies, A., Allman-Farinelli, M., Owen, K., Signal, L., Hosking, C., Wang, L., & Bauman, A. (2020). Feasibility study comparing physical activity classifications from accelerometers with wearable camera data. *International Journal of Environmental Research and Public Health*, 17(24), 9323.
- Davies, G., Reilly, J., McGowan, A., Dall, P., Granat, M., & Paton, J. (2012). Validity, practical utility, and reliability of the activPAL in preschool children. *Medicine and Science in Sports and Exercise*, 44(4), 761-768.
- Davies, G., Reilly, J. J., & Paton, J. Y. (2012). Objective measurement of posture and posture transitions in the pre-school child. *Physiological Measurement*, 33(11), 1913.
- De Craemer, M., De Decker, E., Santos-Lozano, A., Verloigne, M., De Bourdeaudhuij, I., Deforche, B., & Cardon, G. (2015). Validity of the Omron pedometer and the actigraph step count function in preschoolers. *Journal of Science and Medicine in Sport*, 18(3), 289-293.
- De Craemer, M., Decraene, M., Willems, I., Buysse, F., Van Driessche, E., & Verbestel, V. (2021). Objective Measurement of 24-Hour Movement Behaviors in Preschool Children Using Wrist-Worn and Thigh-Worn Accelerometers. *International Journal of Environmental Research and Public Health*, 18(18), 9482.
- De Craemer, M., McGregor, D., Androutsos, O., Manios, Y., & Cardon, G. (2018). Compliance with 24-h movement behaviour guidelines among Belgian pre-school children: the ToyBox-study. *International Journal of Environmental Research and Public Health*, 15(10), 2171.
- De Craemer, M., Verbestel, V., Cardon, G., Androutsos, O., Manios, Y., & Chastin, S. (2020). Correlates of meeting the physical activity, sedentary behavior, and sleep guidelines for the early years among belgian preschool children: The toybox-

- study. *International Journal of Environmental Research and Public Health*, 17(19), 7006.
- De Decker, E., De Craemer, M., Santos-Lozano, A., Van Cauwenberghe, E., De Bourdeaudhuij, I., & Cardon, G. (2013). Validity of the ActivPAL™ and the ActiGraph monitors in preschoolers. *Medicine and Science in Sports and Exercise*, 45(10), 2002-2011.
- De Heer, W., & De Leeuw, E. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. *Survey Nonresponse*, 41, 41-54.
- de Lucena Martins, C. M., Lemos, L. F. G. B. P., de Souza Filho, A. N., Bezerra, T. A., Soares, I. A. A., Mota, J. G., ... & Duncan, M. J. (2021). Adherence to 24-hour movement guidelines in low-income Brazilian preschoolers and associations with demographic correlates. *American Journal of Human Biology*, 33(4), e23519.
- Department of Health and Social Care. (2011). *UK physical activity guidelines*. Available at: [\[Withdrawn\] UK physical activity guidelines - GOV.UK \(www.gov.uk\)](#). (Accessed: 20th January 2021).
- De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: a practical guide*. Cambridge university press.
- Deeks, J. J., Dinnes, J., D'Amico, R., Sowden, A. J., Sakarovich, C., Song, F., ... & Altman, D. G. (2003). Evaluating non-randomised intervention studies. *Health Technology Assessment*, 7(27), iii-173.
- Deeks, J. J., Higgins, J.P.T, Altman, D.G (editors). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (editors). *Cochrane Handbook for Systematic Reviews of*

- Interventions* version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook.
- Degroote, L., DeSmet, A., De Bourdeaudhuij, I., Van Dyck, D., & Crombez, G. (2020). Content validity and methodological considerations in ecological momentary assessment studies on physical activity and sedentary behaviour: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 17(1), 35.
- Delisle Nyström, C., Pomeroy, J., Henriksson, P., Forsum, E., Ortega, F. B., Maddison, R., ... & Löf, M. (2017). Evaluation of the wrist-worn ActiGraph wGT3x-BT for estimating activity energy expenditure in preschool children. *European Journal of Clinical Nutrition*, 71(10), 1212-1217.
- Demarest, S., Van der Heyden, J., Charafeddine, R., Tafforeau, J., Van Oyen, H., & Van Hal, G. (2013). Socio-economic differences in participation of households in a Belgian national health survey. *The European Journal of Public Health*, 23(6), 981-985.
- Dietz Jr, W. H., & Gortmaker, S. L. (1985). Do we fatten our children at the television set? Obesity and television viewing in children and adolescents. *Pediatrics*, 75(5), 807-812.
- Dishman, R. K., Washburn, R. A., & Schoeller, D. A. (2001). Measurement of physical activity. *Quest*, 53(3), 295-309.
- Djfarian, K., Speakman, J. R., Stewart, J., & Jackson, D. M. (2013). Comparison of activity levels measured by a wrist worn accelerometer and direct observation in young children. *Open Journal of Pediatrics*, 3, 422-427.
- Dobell, A. P., Eyre, E. L., Tallis, J., Chinapaw, M. J., Altenburg, T. M., & Duncan, M. J. (2019). Examining accelerometer validity for estimating physical activity in pre-schoolers

- during free-living activity. *Scandinavian Journal of Medicine & Science in Sports*, 29(10), 1618-1628.
- Dobbels, F., de Jong, C., Drost, E., Elberse, J., Feridou, C., Jacobs, L., ... & PROactive Consortium. (2014). The PROactive innovative conceptual framework on physical activity. *European Respiratory Journal*, 44(5), 1223-1233.
- Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M. H., ... & Wareham, N. J. (2017). Large scale population assessment of physical activity using wrist worn accelerometers: the UK biobank study. *PloS One*, 12(2), e0169649.
- Doherty, A. R., Hodges, S. E., King, A. C., Smeaton, A. F., Berry, E., Moulin, C. J., ... & Foster, C. (2013a). Wearable cameras in health: the state of the art and future possibilities. *American Journal of Preventive Medicine*, 44(3), 320-323.
- Doherty, A. R., Kelly, P., Kerr, J., Marshall, S., Oliver, M., Badland, H., ... & Foster, C. (2013b). Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. *International Journal of Behavioral Nutrition and Physical Activity*, 10(1), 22.
- Dössegger, A., Ruch, N., Jimmy, G., Braun-Fahrländer, C., Mäder, U., Hänggi, J., ... & Bringolf-Isler, B. (2014). Reactivity to accelerometer measurement of children and adolescents. *Medicine and science in sports and exercise*, 46(6), 1140-1146.
- Dowd, K. P., Szeklicki, R., Minetto, M. A., Murphy, M. H., Polito, A., Ghigo, E., ... & Donnelly, A. E. (2018). A systematic literature review of reviews on techniques for physical activity measurement in adults: a DEDIPAC study. *International Journal of Behavioral Nutrition and Physical Activity*, 15(1), 15.

- Downing, K. L., Janssen, X., & Reilly, J. J. (2019). Feasibility of wearable cameras to assess screen time and time spent restrained in children aged 3 to 5 years: a study protocol. *BMJ Open*, 9(5), e028265.
- Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*, 52(6), 377-384.
- Draper, C. E., Tomaz, S. A., Biersteker, L., Cook, C. J., Couper, J., De Milander, M., ... & Okely, A. D. (2020a). The South African 24-hour movement guidelines for birth to 5 years: an integration of physical activity, sitting behavior, screen time, and sleep. *Journal of Physical Activity and Health*, 17(1), 109-119.
- Draper, C. E., Tomaz, S. A., Cook, C. J., Jugdav, S. S., Ramsammy, C., van Heerden, A., ... & Okely, A. D. (2020b). Understanding the influence of 24-hour movement behaviours on the health and development of preschool children from low-income South African settings: the SUNRISE pilot study. *South African Journal of Sports Medicine*, 32(1), 1-7.
- Dunton, G. F. (2017). Ecological momentary assessment in physical activity research. *Exercise and Sport Sciences Reviews*, 45(1), 48-54.
- Dunton, G. F., Kawabata, K., Intille, S., Wolch, J., & Pentz, M. A. (2012). Assessing the social and physical contexts of children's leisure-time physical activity: an ecological momentary assessment study. *American Journal of Health Promotion*, 26(3), 135-142.

- Dunton, G. F., Liao, Y., Intille, S. S., Spruijt-Metz, D., & Pentz, M. (2011). Investigating children's physical activity and sedentary behavior using ecological momentary assessment with mobile phones. *Obesity*, 19(6), 1205-1212.
- Duraccio, K. M., Carbine, K. A., Barnett, K. A., Stevens, K. S., & Jensen, C. D. (2018). The utility of the children's sleep habits questionnaire: associations between parental report and an objective measure of sleep behavior. *Children's Health Care*, 47(2), 119-135.
- Duraccio, K. M., & Jensen, C. D. (2017). Associations between physical and sedentary activity regularity and sleep in preschoolers and kindergartners. *Sleep Health*, 3(4), 263-268.
- Dwyer, G. M., Hardy, L. L., Peat, J. K., & Baur, L. A. (2011). The validity and reliability of a home environment preschool-age physical activity questionnaire (Pre-PAQ). *International Journal of Behavioral Nutrition and Physical Activity*, 8(1), 86.
- Department of Education. (2020). *School Admissions- School Starting Age*. Available at: <https://www.gov.uk/schools-admissions/school-starting-age>. (Accessed: 23rd March 2020).
- Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological methods*, 5(2), 155.
- Effective Public Health Practice Project. (2009). *Effective Public Health Practice Project (EPHPP) Quality assessment tool for quantitative studies*. Available at: [1.pdf \(springer.com\)](#). (Accessed: 24th July 2020).
- Ekblom, O., Nyberg, G., Bak, E. E., Ekelund, U., & Marcus, C. (2012). Validity and comparability of a wrist-worn accelerometer in children. *Journal of Physical Activity and Health*, 9(3), 389-393.

- Ekelund, U., Sjostrom, M., Yngve, A., Poortvliet, E., Nilsson, A., Froberg, K., Wedderkopp, N., & Westerterp, K. (2001). Physical activity assessed by activity monitor and doubly labeled water in children. *Medicine and Science in Sports and Exercise*, 33(2), 275-281.
- Ellison, R. C., Freedson, P. S., Zevallos, J. C., White, M. J., Marmor, J. K., Garrahe, E. J., & Moore, L. L. (1992). Feasibility and costs of monitoring physical activity in young children using the Caltrac accelerometer. *Pediatric Exercise Science*, 4(2), 136-141.
- Erwin, A. M., & Bashore, L. (2017). Subjective sleep measures in children: self-report. *Frontiers in Pediatrics*, 5, 22.
- Ettienne, R., Nigg, C. R., Li, F., Su, Y., McGlone, K., Luick, B., ... & Novotny, R. (2016). Validation of the Actical accelerometer in multiethnic preschoolers: the Children's healthy living (CHL) program. *Hawai'i Journal of Medicine & Public Health*, 75(4), 95-100.
- Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity*, 12(1), 159.
- Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of Sports Sciences*, 26(14), 1557-1565.
- Everson, B., Mackintosh, K. A., McNarry, M. A., Todd, C., & Stratton, G. (2019). Can wearable cameras be used to validate school-aged children's lifestyle behaviours?. *Children*, 6(2), 20.

- Fairclough, S. J., Noonan, R. J., Rowlands, A. V., Van Hees, V., Knowles, Z. R., & Boddy, L. M. (2016). Wear Compliance and Activity in Children Wearing Wrist and Hip-Mounted Accelerometers. *Medicine and Science in Sport and Exercise*, 48(2), 245-253.
- Fairweather, S. C., Reilly, J. J., Grant, S., Whittaker, A., & Paton, J. Y. (1999). Using the Computer Science and Applications (CSA) activity monitor in preschool children. *Pediatric Exercise Science*, 11(4), 413-420.
- Falck, R. S., Davis, J. C., Li, L., Stamatakis, E., & Liu-Ambrose, T. (2022). Preventing the '24-hour Babel': the need for a consensus on a consistent terminology scheme for physical activity, sedentary behaviour and sleep. *British Journal of Sports Medicine*, 56(7), 367-368.
- Fern, E. F., & Fern, E. E. (2001). *Advanced focus group research*. Sage.
- Feehan, L. M., Geldman, J., Sayre, E. C., Park, C., Ezzat, A. M., Yoo, J. Y., ... & Li, L. C. (2018). Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. *JMIR mHealth and uHealth*, 6(8), e10527.
- Finn, K. J., & Specker, B. (2000). Comparison of Actiwatch activity monitor and Children's Activity Rating Scale in children. *Medicine and Science in Sports and Exercise*, 32(10), 1794-1797.
- Fino, E., Plazzi, G., Filardi, M., Marzocchi, M., Pizza, F., Vandi, S., & Mazzetti, M. (2020). (Not so) Smart sleep tracking through the phone: Findings from a polysomnography study testing the reliability of four sleep applications. *Journal of Sleep Research*, 29(1), e12935.
- Fisher, R. S., Krietsch, K. N., & Janicke, D. (2019). Feasibility of a single-subject protocol to shift young children's sleep schedules. *Journal of Clinical Psychology in Medical Settings*, 26(4), 541-549.

- Flesch Kincaid Calculator. (2022). *Flesch Kincaid Calculator*. Available at: [Flesch Kincaid Calculator | Good Calculators](#). (Accessed: 23rd March 2022).
- Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., ... & Carrero, J. J. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, 388(10053), 1659-1724.
- Forouhi, N., Brage, S., & Wareham, N. (2017). *Diet, anthropometry and physical activity (DAPA) measurement toolkit*. Available at: <https://dapa-toolkit.mrc.ac.uk/>. (Accessed: 6th January 2020).
- Fotini, V., Antonis, K., & Dimitra, G. M. (2015). The validity of two Omron pedometers in preschool children under different conditions. *Sylwan*, 159, 60-89.
- Freedson, P., Pober, D., & Janz, K. F. (2005). Calibration of accelerometer output for children. *Medicine and Science in Sports and Exercise*, 37(11), S523.
- Fuller, D., Colwell, E., Low, J., Orychock, K., Tobin, M. A., Simango, B., ... & Taylor, N. G. (2020). Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. *JMIR mHealth and uHealth*, 8(9), e18694.
- Fulton, J. E., Burgeson, C. R., Perry, G. R., Sherry, B., Galuska, D. A., Alexander, M. P., ... & Caspersen, C. J. (2001). Assessment of physical activity and sedentary behavior in preschool-age children: priorities for research. *Pediatric Exercise Science*, 13(2), 113-126.
- Galea, S., & Tracy, M. (2007). Participation rates in epidemiologic studies. *Annals of Epidemiology*, 17(9), 643-653.

- Galland, B., Meredith-Jones, K., Gray, A., Sayers, R., Lawrence, J., Taylor, B., & Taylor, R. (2016). Criteria for nap identification in infants and young children using 24-h actigraphy and agreement with parental diary. *Sleep Medicine, 19*, 85-92.
- Galland, B., Meredith-Jones, K., Terrill, P., & Taylor, R. (2014). Challenges and emerging technologies within the field of pediatric actigraphy. *Frontiers in Psychiatry, 5*, 99.
- Ganann, R., Ciliska, D., & Thomas, H. (2010). Expediting systematic reviews: methods and implications of rapid reviews. *Implementation Science, 5*(1), 1-10.
- Garriguet, D., Carson, V., Colley, R. C., Janssen, I., Timmons, B. W., & Tremblay, M. S. (2016). Physical activity and sedentary behaviour of Canadian children aged 3 to 5. *Health Reports, 27*(9), 14.
- Gimeno-Santos, E., Frei, A., Dobbels, F., Rüdell, K., Puhan, M. A., & Garcia-Aymerich, J. (2011). Validity of instruments to measure physical activity may be questionable due to a lack of conceptual frameworks: a systematic review. *Health and Quality of Life Outcomes, 9*(1), 86.
- González-Gil, E. M., Mouratidou, T., Cardon, G., Androutsos, O., De Bourdeaudhuij, I., Gózdź, M., ... & ToyBox-study group. (2014). Reliability of primary caregivers reports on lifestyle behaviours of European pre-school children: the Toy Box-study. *Obesity Reviews, 15*, 61-66.
- Goodman, A., & Gatward, R. (2008). Who are we missing? Area deprivation and survey participation. *European Journal of Epidemiology, 23*(6), 379-387.
- Graham, C., Reither, E. N., Ciciurkaite, G., Dev, D. A., & Fargo, J. (2020). Does context matter? A multilevel analysis of neighborhood disadvantage and children's sleep health. *Sleep Health, 6*(5), 578-586.

- Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91-108.
- Griffin, N., Phillips, S. M., Hillier-Brown, F., Wistow, J., Fairbrother, H., Holding, E., ... & Summerbell, C. (2021). A critique of the English national policy from a social determinants of health perspective using a realist and problem representation approach: the 'Childhood Obesity: a plan for action'(2016, 2018, 2019). *BMC Public Health*, 21(1), 2284.
- Grigg-Damberger, M., Gozal, D., Marcus, C. L., Quan, S. F., Rosen, C. L., Chervin, R. D., ... & Iber, C. (2007). The visual scoring of sleep and arousal in infants and children. *Journal of Clinical Sleep Medicine*, 3(02), 201-240.
- Guan, H., Okely, A. D., Aguilar-Farias, N., del Pozo Cruz, B., Draper, C. E., El Hamdouchi, A., ... & Veldman, S. L. (2020a). Promoting healthy movement behaviours among children during the COVID-19 pandemic. *The Lancet Child & Adolescent Health*, 4(6), 416-418.
- Guan, H., Zhang, Z., Wang, B., Okely, A. D., Tong, M., Wu, J., & Zhang, T. (2020b). Proportion of kindergarten children meeting the WHO guidelines on physical activity, sedentary behaviour and sleep and associations with adiposity in urban Beijing. *BMC Pediatrics*, 20(1), 70.
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj*, 336(7650), 924-926.
- Haghighyegh, S., Khoshnevis, S., Smolensky, M. H., Diller, K. R., & Castriotta, R. J. (2019). Accuracy of wristband Fitbit models in assessing sleep: systematic review and meta-analysis. *Journal of Medical Internet Research*, 21(11), e16273.

- Hale, L., Berger, L. M., LeBourgeois, M. K., & Brooks-Gunn, J. (2009). Social and demographic predictors of preschoolers' bedtime routines. *Journal of Developmental and Behavioral Pediatrics: Journal of Developmental and Behavioral Pediatrics*, 30(5), 394-402.
- Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., Ekelund, U., & Lancet Physical Activity Series Working Group. (2012). Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet*, 380(9838), 247-257.
- Hamel, C., Michaud, A., Thuku, M., Skidmore, B., Stevens, A., Nussbaumer-Streit, B., & Garritty, C. (2021). Defining rapid reviews: a systematic scoping review and thematic analysis of definitions and defining characteristics of rapid reviews. *Journal of Clinical Epidemiology*, 129, 74-85.
- Hands, B., & Larkin, D. (2006). Physical activity measurement methods for young children: A comparative study. *Measurement in Physical Education and Exercise Science*, 10(3), 203-214.
- Harald, K., Salomaa, V., Jousilahti, P., Koskinen, S., & Vartiainen, E. (2007). Non-participation and mortality in different socioeconomic groups: the FINRISK population surveys in 1972–92. *Journal of Epidemiology & Community Health*, 61(5), 449-454.
- Hardman, A. E., & Stensel, D. J. (2009). *Physical activity and health: the evidence explained*. Routledge.
- Hatzinger, M., Brand, S., Perren, S., Von Wyl, A., Stadelmann, S., von Klitzing, K., & Holsboer-Trachsler, E. (2013). In pre-school children, sleep objectively assessed via sleep-EEGs remains stable over 12 months and is related to psychological functioning, but not to cortisol secretion. *Journal of Psychiatric Research*, 47(11), 1809-1814.

- Hefferon, C., Taylor, C., Bennett, D., Falconer, C., Campbell, M., Williams, J. G., ... & Taylor-Robinson, D. (2021). Priorities for the child public health response to the COVID-19 pandemic recovery in England. *Archives of Disease in Childhood*, 106(6), 533-538.
- Helmerhorst, H. H. J., Brage, S., Warren, J., Besson, H., & Ekelund, U. (2012). A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *International Journal of Behavioral Nutrition and Physical Activity*, 9(1), 103.
- HENRY. (2020). *About Henry*. Available at: <https://www.henry.org.uk/about>. (Accessed: 1st November 2021).
- Hesketh, K. R., Lakshman, R., & van Sluijs, E. M. (2017). Barriers and facilitators to young children's physical activity and sedentary behaviour: a systematic review and synthesis of qualitative literature. *Obesity Reviews*, 18(9), 987-1017.
- Hesketh, K. R., McMinn, A. M., Ekelund, U., Sharp, S. J., Collings, P. J., Harvey, N. C., ... & van Sluijs, E. M. (2014). Objectively measured physical activity in four-year-old British children: a cross-sectional analysis of activity patterns segmented across the day. *International Journal of Behavioral Nutrition and Physical Activity*, 11(1), 1.
- Hidding, L. M., Altenburg, T. M., Mokkink, L. B., Terwee, C. B., & Chinapaw, M. J. (2017). Systematic review of childhood sedentary behavior questionnaires: what do we know and what is next?. *Sports Medicine*, 47(4), 677-699.
- Hidding, L. M., Chinapaw, M.J.M., van Poppel, M. N., Mokkink, L. B., & Altenburg, T. M. (2018). An updated systematic review of childhood physical activity questionnaires. *Sports Medicine*, 48(12), 2797-2842.
- Hidding, L. M., Chinapaw, M. J., Belmon, L. S., & Altenburg, T. M. (2020). Co-creating a 24-hour movement behavior tool together with 9–12-year-old children using mixed-

- methods: MyDailyMoves. *International Journal of Behavioral Nutrition and Physical Activity*, 17(1), 63.
- Hillier-Brown, F. C., Bambra, C. L., Cairns, J. M., Kasim, A., Moore, H. J., & Summerbell, C. D. (2014). A systematic review of the effectiveness of individual, community and societal level interventions at reducing socioeconomic inequalities in obesity amongst children. *BMC Public Health*, 14(1), 834.
- Hillier, F., Pedley, C., & Summerbell, C. (2011). Evidence base for primary prevention of obesity in children and adolescents. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz*, 54(3), 259-264.
- Hinkley, T., O'Connell, E., Okely, A. D., Crawford, D., Hesketh, K., & Salmon, J. (2012a). Assessing volume of accelerometry data for reliability in preschool children. *Medicine and Science in Sports and Exercise*, 44(12), 2436-2441.
- Hinkley, T., Salmon, J., Okely, A. D., Crawford, D., & Hesketh, K. (2012b). The HAPPY study: development and reliability of a parent survey to assess correlates of preschool children's physical activity. *Journal of science and medicine in sport*, 15(5), 407-417.
- Hinkley, T., Timperio, A., Watson, A., Duckham, R. L., Okely, A. D., Cliff, D., ... & Hesketh, K. D. (2020). Prospective associations with physiological, psychosocial and educational outcomes of meeting Australian 24-Hour Movement Guidelines for the Early Years. *International Journal of Behavioral Nutrition and Physical Activity*, 17(1), 36.
- Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., ... & Ware, J. C. (2015a). National Sleep Foundation's updated sleep duration recommendations. *Sleep Health*, 1(4), 233-243.

- Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., ... & Hillard, P. J. A. (2015b). National Sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health, 1*(1), 40-43.
- Hislop, J. F., Bulley, C., Mercer, T. H., & Reilly, J. J. (2012a). Comparison of accelerometry cut points for physical activity and sedentary behavior in preschool children: a validation study. *Pediatric Exercise Science, 24*(4), 563-576.
- Hislop, J. F., Bulley, C., Mercer, T. H., & Reilly, J. J. (2012b). Comparison of epoch and uniaxial versus triaxial accelerometers in the measurement of physical activity in preschool children: a validation study. *Pediatric Exercise Science, 24*(3), 450-460.
- Hislop, J., Law, J., Rush, R., Grainger, A., Bulley, C., Reilly, J. J., & Mercer, T. (2014). An investigation into the minimum accelerometry wear time for reliable estimates of habitual physical activity and definition of a standard measurement day in pre-school children. *Physiological Measurement, 35*(11), 2213.
- Hislop, J., Palmer, N., Anand, P., & Aldin, T. (2016). Validity of wrist worn accelerometers and comparability between hip and wrist placement sites in estimating physical activity behaviour in preschool children. *Physiological Measurement, 37*(10), 1701-1714.
- Hjorth, M. F., Chaput, J. P., Damsgaard, C. T., Dalskov, S. M., Michaelsen, K. F., Tetens, I., & Sjödin, A. (2012). Measure of sleep and physical activity by a single accelerometer: can a waist-worn Actigraph adequately measure sleep in children?. *Sleep and Biological Rhythms, 10*(4), 328-335.
- Hoffmann, B., Kobel, S., Wartha, O., Kettner, S., Dreyhaupt, J., & Steinacker, J. M. (2019). High sedentary time in children is not only due to screen media use: A cross-sectional study. *BMC Pediatrics, 19*(1), 154.

- Hossain, M. S., Deeba, I. M., Hasan, M., Kariippanon, K. E., Chong, K. H., Cross, P. L., ... & Okely, A. D. (2021). International study of 24-h movement behaviors of early years (SUNRISE): a pilot study from Bangladesh. *Pilot and Feasibility Studies*, 7(1), 176
- Hounslow, J. L., Brewster, L. R., Lear, K. O., Guttridge, T. L., Daly, R., Whitney, N. M., & Gleiss, A. C. (2019). Assessing the effects of sampling frequency on behavioural classification of accelerometer data. *Journal of Experimental Marine Biology and Ecology*, 512, 22-30.
- Huberty, J., Ehlers, D. K., Kurka, J., Ainsworth, B., & Buman, M. (2015). Feasibility of three wearable sensors for 24 hour monitoring in middle-aged women. *BMC Women's Health*, 15(1), 55.
- Ibáñez, V., Silva, J., & Cauli, O. (2018). A survey on sleep assessment methods. *PeerJ*, 6, e4849.
- Iber, C., Ancoli-Israel, S., Chesson, A.L., & Quan, S.F. (2007). The AASM manual for the scoring of sleep and associated events: Rules. *Terminology and Technical Specification*. Westchester, IL: American Academy of Sleep Medicine.
- Ipsos MORI. (2021). *Active Lives Survey 2019/2020, Full-Year 5 Technical Note*. Available at: [Normal dot \(Rev02 January 2009\) \(sportengland-production-files.s3.eu-west-2.amazonaws.com\)](https://www.sportengland-production-files.s3.eu-west-2.amazonaws.com/Normal%20dot%20(R%20Rev02%20January%202009)%20(sportengland-production-files.s3.eu-west-2.amazonaws.com).). (Accessed: 31st August 2021).
- Ishihara, K., Doi, Y., & Uchiyama, M. (2014). The reliability and validity of the Japanese version of the Children's ChronoType Questionnaire (CCTQ) in preschool children. *Chronobiology International*, 31(9), 947-953.
- Iwasaki, M., Iwata, S., Iemura, A., Yamashita, N., Tomino, Y., Anme, T., ... & Matsuishi, T. (2010). Utility of subjective sleep assessment tools for healthy preschool children: a

- comparative study between sleep logs, questionnaires, and actigraphy. *Journal of Epidemiology*, 1002030129-1002030129.
- Iwata, S., Iwata, O., Iemura, A., Iwasaki, M., & Matsuishi, T. (2012). Sleep architecture in healthy 5-year-old preschool children: associations between sleep schedule and quality variables. *Acta Paediatrica*, 101(3), e110-e114.
- Janssen, X., & Cliff, D. P. (2015). Issues related to measuring and interpreting objectively measured sedentary behavior data. *Measurement in Physical Education and Exercise Science*, 19(3), 116-124.
- Janssen, X., Cliff, D., Okely, A. D., Jones, R. A., Batterham, M., Ekelund, U., ... & Reilly, J. J. (2013c). Practical utility and reliability of whole-room calorimetry in young children. *British Journal of Nutrition*, 109(10), 1917-1922.
- Janssen, X., Cliff, D., Reilly, J., Hinkley, T., Jones, R., Batterham, M., ... & Okely, T. (2015). Evaluation of Actical equations and thresholds to predict physical activity intensity in young children. *Journal of Sports Sciences*, 33(5), 498-506.
- Janssen, X., Cliff, D. P., Reilly, J. J., Hinkley, T., Jones, R. A., Batterham, M., ... & Okely, A. D. (2013a). Predictive validity and classification accuracy of ActiGraph energy expenditure equations and cut-points in young children. *PloS one*, 8(11), e79124.
- Janssen, X., Cliff, D. P., Reilly, J. J., Hinkley, T., Jones, R. A., Batterham, M., ... & Okely, A. D. (2013b). Validation of activPAL™ defined sedentary time and breaks in sedentary time in 4-to 6-year olds. *Pediatric Exercise Science*, 26 (1), 110-117.
- Janssen, X., Cliff, D. P., Reilly, J. J., Hinkley, T., Jones, R. A., Batterham, M., ... & Okely, A. D. (2014). Validation and calibration of the activPAL™ for estimating METs and physical activity in 4–6 year olds. *Journal of Science and Medicine in Sport*, 17(6), 602-606.

- Janssen, X., Martin, A., Hughes, A. R., Hill, C. M., Kotronoulas, G., & Hesketh, K. R. (2020). Associations of screen time, sedentary time and physical activity with sleep in under 5s: A systematic review and meta-analysis. *Sleep Medicine Reviews, 49*, 101226.
- Janz, K. F., Broffitt, B., & Levy, S. M. (2005). Validation evidence for the Netherlands physical activity questionnaire for young children: the Iowa bone development study. *Research Quarterly for Exercise and Sport, 76*(3), 363-369.
- Johansson, E., Ekelund, U., Nero, H., Marcus, C., & Hagströmer, M. (2015). Calibration and cross-validation of a wrist-worn Actigraph in young preschoolers. *Pediatric Obesity, 10*(1), 1-6.
- Jones, C. H., & Ball, H. (2014). Exploring socioeconomic differences in bedtime behaviours and sleep duration in English preschool children. *Infant and Child Development, 23*(5), 518-531.
- Jones, D., Innerd, A., Giles, E. L., & Azevedo, L. B. (2021). The Association between Physical Activity, Motor Skills and School Readiness in 4–5-Year-Old Children in the Northeast of England. *International Journal of Environmental Research and Public Health, 18*(22), 11931.
- Jones, R. A., Hinkley, T., Okely, A. D., & Salmon, J. (2013). Tracking physical activity and sedentary behavior in childhood: a systematic review. *American Journal of Preventive Medicine, 44*(6), 651-658.
- Jutai, J. W., Strong, J. G., & Russell-Minda, E. (2009). Effectiveness of assistive technologies for low vision rehabilitation: A systematic review. *Journal of Visual Impairment & Blindness, 103*(4), 210-222.

- Kahan, D., Nicaise, V., & Reuben, K. (2013). Convergent validity of four accelerometer cutpoints with direct observation of preschool children's outdoor physical activity. *Research quarterly for exercise and sport*, 84(1), 59-67.
- Katzmarzyk, P. T., Friedenreich, C., Shiroma, E. J., & Lee, I. M. (2022). Physical inactivity and non-communicable disease burden in low-income, middle-income and high-income countries. *British Journal of Sports Medicine*, 56(2), 101-106.
- Keadle, S. K., Shiroma, E. J., Freedson, P. S., & Lee, I. M. (2014). Impact of accelerometer data processing decisions on the sample size, wear time and physical activity level of a large cohort study. *BMC Public Health*, 14(1), 14.
- Kelly, L. A., Reilly, J. J., Fairweather, S. C., Barrie, S., Grant, S., & Paton, J. Y. (2004). Comparison of two accelerometers for assessment of physical activity in preschool children. *Pediatric Exercise Science*, 16(4), 324-333.
- Kelly, L. A., Reilly, J. J., Fisher, A., Montgomery, C., Williamson, A., McColl, J. H., ... & Grant, S. (2006). Effect of socioeconomic status on objectively measured physical activity. *Archives of Disease in Childhood*, 91(1), 35-38.
- Kelly, P., Doherty, A. R., Hamilton, A., Matthews, A., Batterham, A. M., Nelson, M., ... & Cowburn, G. (2012). Evaluating the feasibility of measuring travel to school using a wearable camera. *American Journal of Preventive Medicine*, 43(5), 546-550.
- Kelly, P., Fitzsimons, C., & Baker, G. (2016). Should we reframe how we think about physical activity and sedentary behaviour measurement? Validity and reliability reconsidered. *International Journal of Behavioral Nutrition and Physical Activity*, 13(1), 32.

- Kelly, P., Marshall, S. J., Badland, H., Kerr, J., Oliver, M., Doherty, A. R., & Foster, C. (2013).
An ethical framework for automated, wearable cameras in health behavior
research. *American Journal of Preventive Medicine*, 44(3), 314-319.
- Kelly, S. E., Moher, D., & Clifford, T. J. (2016). Quality of conduct and reporting in rapid
reviews: an exploration of compliance with PRISMA and AMSTAR
guidelines. *Systematic reviews*, 5(1), 79.
- Khanom, A., Evans, B. A., Lynch, R., Marchant, E., Hill, R. A., Morgan, K., ... & Brophy, S.
(2020). Parent recommendations to support physical activity for families with young
children: Results of interviews in deprived and affluent communities in South Wales
(United Kingdom). *Health Expectations*, 23(2), 284-295.
- Kim, H., Ma, J., Harada, K., Lee, S., & Gu, Y. (2020). Associations between adherence to
combinations of 24-h movement guidelines and overweight and obesity in Japanese
preschool children. *International Journal of Environmental Research and Public
Health*, 17(24), 9320.
- Kim, T. V., Pham, T. N., Nguyen, C. L., Nguyen, T. T., Okely, A. D., & Tang, H. K. (2022).
Prevalence of Physical Activity, Screen Time, and Sleep, and Associations with
Adiposity and Motor Development among Preschool-Age Children in Vietnam: The
SUNRISE Vietnam Pilot Study. *Indian Journal of Pediatrics*, 89(2), 148-153.
- Kim, Y., Beets, M. W., Pate, R. R., & Blair, S. N. (2013). The effect of reintegrating Actigraph
accelerometer counts in preschool children: Comparison using different epoch
lengths. *Journal of Science and Medicine in Sport*, 16(2), 129-134.
- Kinder, J. R., Lee, K. A., Thompson, H., Hicks, K., Topp, K., & Madsen, K. A. (2012). Validation
of a hip-worn accelerometer in measuring sleep time in children. *Journal of Pediatric
Nursing*, 27(2), 127-133.

- Klesges, R. C. (1984). The FATS: An observational system for assessing physical activity in children and associated parent behavior. *Behavioral Assessment*, 6(4), 333–345.
- Klesges, R. C., Woolfrey, J., & Vollmer, J. (1985). An evaluation of the reliability of time sampling versus continuous observation data collection. *Journal of Behavior Therapy and Experimental Psychiatry*, 16(4), 303-307.
- Knaeps, S., Lefevre, J., Wijtzes, A., Charlier, R., Mertens, E., & Bourgois, J. G. (2016). Independent associations between sedentary time, moderate-to-vigorous physical activity, cardiorespiratory fitness and cardio-metabolic health: a cross-sectional study. *PloS One*, 11(7), e0160166.
- Kocevska, D., Muetzel, R. L., Luik, A. I., Luijk, M. P., Jaddoe, V. W., Verhulst, F. C., ... & Tiemeier, H. (2017). The developmental course of sleep disturbances across childhood relates to brain morphology at age 7: the Generation R Study. *Sleep*, 40(1).
- Kohl III, H. W., Fulton, J. E., & Caspersen, C. J. (2000). Assessment of physical activity among children and adolescents: a review and synthesis. *Preventive Medicine*, 31(2), S54-S76.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163.
- Korakakis, V., Whiteley, R., Tzavara, A., & Malliaropoulos, N. (2018). The effectiveness of extracorporeal shockwave therapy in common lower limb conditions: a systematic review including quantification of patient-rated pain reduction. *British Journal of Sports Medicine*, 52(6), 387-407.
- Korkeila, K., Suominen, S., Ahvenainen, J., Ojanlatva, A., Rautava, P., Helenius, H., & Koskenvuo, M. (2001). Non-response and related factors in a nation-wide health survey. *European Journal of Epidemiology*, 17(11), 991-999.

- Kowalski, K., Rhodes, R., Naylor, P. J., Tuokko, H., & MacDonald, S. (2012). Direct and indirect measurement of physical activity in older adults: a systematic review of the literature. *International Journal of Behavioral Nutrition and Physical Activity*, 9(1), 148.
- Kracht, C. L., Katzmarzyk, P. T., & Staiano, A. E. (2021). Household chaos, family routines, and young child movement behaviors in the US during the COVID-19 outbreak: A cross-sectional study. *BMC Public Health*, 21(1), 860.
- Kracht, C. L., Webster, E. K., & Staiano, A. E. (2020). Relationship between the 24-Hour Movement Guidelines and fundamental motor skills in preschoolers. *Journal of Science and Medicine in Sport*, 23(12), 1185-1190.
- Kracht, C. L., Webster, E. K., & Staiano, A. E. (2019). Sociodemographic differences in young children meeting 24-hour movement guidelines. *Journal of Physical Activity and Health*, 16(10), 908-915.
- Kuntsche, E., & Labhart, F. (2013). Using personal cell phones for ecological momentary assessment. *European Psychologist*, 18, 3-11.
- Kushnir, J., & Sadeh, A. (2013). Correspondence between reported and actigraphic sleep measures in preschool children: the role of a clinical context. *Journal of Clinical Sleep Medicine*, 9(11), 1147-1151.
- Kuzik, N., & Carson, V. (2016). The association between physical activity, sedentary behavior, sleep, and body mass index z-scores in different settings among toddlers and preschoolers. *BMC Pediatrics*, 16(1), 100.
- Kuzik, N., Naylor, P. J., Spence, J. C., & Carson, V. (2020). Movement behaviours and physical, cognitive, and social-emotional development in preschool-aged children:

- Cross-sectional associations using compositional analyses. *PLoS One*, 15(8), e0237945.
- Kuzik, N., Poitras, V. J., Tremblay, M. S., Lee, E. Y., Hunter, S., & Carson, V. (2017). Systematic review of the relationships between combinations of movement behaviours and health indicators in the early years (0-4 years). *BMC Public Health*, 17(5), 849.
- Kyeremateng, R., Oguda, L., & Asemota, O. (2022). COVID-19 pandemic: health inequities in children and youth. *Archives of Disease in Childhood*, 107(3), 297-299.
- Lam, J. C., Mahone, E. M., Mason, T. B., & Scharf, S. M. (2011). The effects of napping on cognitive function in preschoolers. *Journal of Developmental and Behavioral Pediatrics*, 32(2), 90-97.
- Lancaster, G. A., Dodd, S., & Williamson, P. R. (2004). Design and analysis of pilot studies: recommendations for good practice. *Journal of Evaluation in Clinical Practice*, 10(2), 307-312.
- Larcher, V., & Brierley, J. (2020). Children of COVID-19: pawns, pathfinders or partners?. *Journal of Medical Ethics*, 46(8), 508-509.
- Larson, T. A., Normand, M. P., & Hustyi, K. M. (2011). Preliminary evaluation of an observation system for recording physical activity in children. *Behavioral Interventions*, 26(3), 193-203.
- Lasserson, T.J., Thomas, J., & Higgins, J.P.T. (2019). Chapter 1: Starting a review. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (editors). *Cochrane Handbook for Systematic Reviews of Interventions. 2nd Edition*. Chichester (UK): John Wiley & Sons, 3–12.
- LeBlanc, A. G., Spence, J. C., Carson, V., Connor Gorber, S., Dillman, C., Janssen, I., ... & Tremblay, M. S. (2012). Systematic review of sedentary behaviour and health

- indicators in the early years (aged 0–4 years). *Applied Physiology, Nutrition, and Metabolism*, 37(4), 753-772.
- LeBourgeois, M.K., Harsh, J.R. (2016). Development and psychometric evaluation of the Children's Sleep-Wake Scale. *Sleep Health*, 2(3), 198-204.
- Lee, I. M., & Shiroma, E. J. (2014). Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. *British Journal of Sports Medicine*, 48(3), 197-201.
- Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., Katzmarzyk, P. T., & Lancet Physical Activity Series Working Group. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet*, 380 (9838), 219-229.
- Lee, M., Cho, J., Lee, H., Park, C. H., Oh, J., & Choi, M. C. (2014). Validity evidence of objective physical activity measures for early childhood. *Research Quarterly for Exercise and Sport*, 85(S1), A77.
- Lee, S. T., Wong, J. E., Chan, G. K., & Poh, B. K. (2021). Association between compliance with movement behavior guidelines and obesity among Malaysian preschoolers. *International Journal of Environmental Research and Public Health*, 18(9), 4611.
- Lemos, L., Clark, C., Brand, C., Pessoa, M. L., Gaya, A., Mota, J., ... & Martins, C. (2021). 24-hour movement behaviors and fitness in preschoolers: A compositional and isotemporal reallocation analysis. *Scandinavian Journal of Medicine & Science in Sports*, 31(6), 1371-1379.
- Leppänen, M. H., Ray, C., Wennman, H., Alexandrou, C., Sääksjärvi, K., Koivusilta, L., ... & Roos, E. (2019). Compliance with the 24-h movement guidelines and the relationship

- with anthropometry in Finnish preschoolers: The DAGIS study. *BMC Public Health*, 19(1), 1618.
- Lewandowski, A. S., Toliver-Sokol, M., & Palermo, T. M. (2011). Evidence-based review of subjective pediatric sleep measures. *Journal of Pediatric Psychology*, 36(7), 780-793.
- Li, L., Sheehan, C. M., Valiente, C., Eisenberg, N., Doane, L. D., Spinrad, T. L., ... & Southworth, J. (2021). Similarities and differences between actigraphy and parent-reported sleep in a hispanic and non-hispanic white sample. *Sleep Medicine*, 83, 160-167.
- Liao, Y., Intille, S., Wolch, J., Pentz, M. A., & Dunton, G. F. (2014). Understanding the physical and social contexts of children's nonschool sedentary behavior: an ecological momentary assessment study. *Journal of Physical Activity and Health*, 11(3), 588-595.
- Liggett, L., Gray, A., Parnell, W., McGee, R., & McKenzie, Y. (2012). Validation and reliability of the New Lifestyles NL-1000 accelerometer in New Zealand preschoolers. *Journal of Physical Activity and Health*, 9(2), 295-299.
- Lima, R. A., Barros, S. S. H., Cardoso Júnior, C. G., Silva, G., Farias Júnior, J. C. D., Andersen, L. B., & de Barros, M. V. G. (2014). Influence of number of days and valid hours using accelerometry on the estimates of physical activity level in preschool children from Recife, Pernambuco, Brazil. *Revista Brasileira de Cineantropometria & Desempenho Humano*, 16, 171-181.
- Liu, Z., Tang, H., Jin, Q., Wang, G., Yang, Z., Chen, H., ... & Owens, J. (2021). Sleep of preschoolers during the coronavirus disease 2019 (COVID-19) outbreak. *Journal of Sleep Research*, 30(1), e13142.

- Lohr, K. N. (2002). Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of life Research*, 11(3), 193-205.
- López-Bueno, R., López-Sánchez, G. F., Casajús, J. A., Calatayud, J., Gil-Salmerón, A., Grabovac, I., ... & Smith, L. (2020). Health-related behaviors among school-aged children and adolescents during the Spanish Covid-19 confinement. *Frontiers in Pediatrics*, 573.
- López-Bueno, R., López-Sánchez, G. F., Casajús, J. A., Calatayud, J., Tully, M. A., & Smith, L. (2021). Potential health-related behaviors for pre-school and school-aged children during COVID-19 lockdown: A narrative review. *Preventive Medicine*, 143, 106349.
- Lopez-Alarcon, M., Merrifield, J., Fields, D. A., Hilario-Hailey, T., Franklin, F. A., Shewchuk, R. M., ... & Gower, B. A. (2004). Ability of the actiwatch accelerometer to predict free-living energy expenditure in young children. *Obesity Research*, 12(11), 1859-1865.
- Loprinzi, P. D., & Cardinal, B. J. (2011). Measuring children's physical activity and sedentary behaviors. *Journal of Exercise Science & Fitness*, 9(1), 15-23.
- Louie, L. & Chan, L. (2003). The use of pedometry to evaluate the physical activity levels among preschool children in Hong Kong. *Early Child Development and Care*, 173, 97-107.
- Love, R. E., Adams, J. & Van Sluijs, E. M. F. (2017). Equity effects of children's physical activity interventions: a systematic scoping review. *International Journal of Behavioral Nutrition and Physical Activity*, 14, 134.
- Lubans, D. R., Hesketh, K., Cliff, D. P., Barnett, L. M., Salmon, J., Dollman, J., ... & Hardy, L. L. (2011). A systematic review of the validity and reliability of sedentary behaviour measures used with children and adolescents. *Obesity reviews*, 12(10), 781-799.

- Lucas-de la Cruz, L., Martínez-Vizcaino, V., Álvarez-Bueno, C., Arias-Palencia, N., Sánchez-López, M., & Notario-Pacheco, B. (2016). Reliability and validity of the Spanish version of the Children's Sleep Habits Questionnaire (CSHQ-SP) in school-age children. *Child: Care, Health and Development*, 42(5), 675-682.
- Lucena Martins, C. M. D., Clark, C. C. T., Ribeiro Bandeira, P. F., Mota, J., & Duncan, M. J. (2020). Association between compliance with the 24-hour movement guidelines and fundamental movement skills in preschoolers: A network perspective. *International Journal of Environmental Research and Public Health*, 17(15), 5443.
- Mahase, E. (2020). Covid-19: schools set to close across UK except for children of health and social care workers, 368, m1140.
- Maidment, D. W., Barker, A. B., Xia, J., & Ferguson, M. A. (2018). A systematic review and meta-analysis assessing the effectiveness of alternative listening devices to conventional hearing aids in adults with hearing loss. *International Journal of Audiology*, 57(10), 721-729.
- Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). *Growth, maturation, and physical activity*. Human kinetics.
- Manios, Y., Kafatos, A., & Markakis, G. (1998). Physical activity of 6-year-old children: Validation of two proxy reports. *Pediatric Exercise Science*, 10(2), 176-188.
- Marmot, M., Allen, J., Boyce, T., Goldblatt, P., & Morrison, J. (2020a). Health Equity in England: The Marmot Review 10 Years On. London: Institute of Health Equity.
- Marmot, M., Allen, J., Goldblatt, P., Boyce, T., McNeish, D., Grady, M., & Geddes, I. (2010). Fair society. *Healthy Lives: The Marmot Review*.

- Marmot, M., Allen, J., Goldblatt, P., Herd, E., & Morrison, J. (2020b). Build back fairer: the COVID-19 Marmot review the pandemic, socioeconomic and health inequalities in England. London: Institute of Health Equity.
- Martin, A., McNeill, M., Penpraze, V., Dall, P., Granat, M., Paton, J. Y., & Reilly, J. J. (2011). Objective measurement of habitual sedentary behavior in pre-school children: comparison of activPAL With Actigraph monitors. *Pediatric Exercise Science*, 23(4), 468-476.
- Martin, R., Murphy, J., Molina-Soberanes, D., & Murtagh, E. M. (2022). The clustering of physical activity and screen time behaviours in early childhood and impact on future health-related behaviours: a longitudinal analysis of children aged 3 to 8 years. *BMC Public Health*, 22(1), 558.
- Mathyssek, C. M., Olino, T. M., Hartman, C. A., Ormel, J., Verhulst, F. C., & Van Oort, F. V. (2013). Does the Revised Child Anxiety and Depression Scale (RCADS) measure anxiety symptoms consistently across adolescence? The TRAILS study. *International Journal of Methods in Psychiatric Research*, 22(1), 27-35.
- Matricciani, L. (2013). Subjective reports of children's sleep duration: does the question matter? A literature review. *Sleep Medicine*, 14(4), 303-311.
- McDowall, P. S., Galland, B. C., Campbell, A. J., & Elder, D. E. (2017). Parent knowledge of children's sleep: A systematic review. *Sleep Medicine Reviews*, 31, 39-47.
- McHugh, M.L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276-282.
- McKee, D. P., Boreham, C. A., Murphy, M. H., & Nevill, A. M. (2005). Validation of the Digiwalker™ pedometer for measuring physical activity in young children. *Pediatric Exercise Science*, 17(4), 345-352.

- McNeill, J., Howard, S. J., Vella, S. A., & Cliff, D. P. (2020). Compliance with the 24-Hour movement guidelines for the early years: Cross-sectional and longitudinal associations with executive function and psychosocial health in preschool children. *Journal of Science and Medicine in Sport*, 23(9), 846-853.
- Ministry of Health, New Zealand. (2017a). *Content Guide 2016/17: New Zealand Health Survey*. Wellington: Ministry of Health. Available at: [Content Guide 2016/17: New Zealand Health Survey | Ministry of Health NZ](#). (Accessed: 10th February 2022).
- Ministry of Health, New Zealand. (2017b). *Sit Less, Move More, Sleep Well: Active play guidelines for under-fives*. Available at: [Active Play Guidelines for Under-fives \(health.govt.nz\)](#). (Accessed: 1st May 2019).
- Mekary, R. A., Lucas, M., Pan, A., Okereke, O. I., Willett, W. C., Hu, F. B., & Ding, E. L. (2013). Isotemporal substitution analysis for physical activity, television watching, and risk of depression. *American Journal of Epidemiology*, 178(3), 474-483.
- Meltzer, L. J., Hiruma, L. S., Avis, K., Montgomery-Downs, H., & Valentin, J. (2015). Comparison of a commercial accelerometer with polysomnography and actigraphy in children and adolescents. *Sleep*, 38(8), 1323-1330.
- Meltzer, L. J., Montgomery-Downs, H. E., Insana, S. P., & Walsh, C. M. (2012b). Use of actigraphy for assessment in pediatric sleep research. *Sleep Medicine Reviews*, 16(5), 463-475.
- Meltzer, L. J., Walsh, C. M., Traylor, J., & Westin, A. M. (2012a). Direct comparison of two new actigraphs and polysomnography in children and adolescents. *Sleep*, 35(1), 159-166.

- Mendoza, J. A., McLeod, J., Chen, T. A., Nicklas, T. A., & Baranowski, T. (2013). Convergent validity of preschool children's television viewing measures among low-income Latino families: a cross-sectional study. *Childhood Obesity*, 9(1), 29-34.
- Meredith-Jones, K., Galland, B., Haszard, J., Gray, A., Sayers, R., Hanna, M., ... & Taylor, R. (2019). Do young children consistently meet 24-h sleep and activity guidelines? A longitudinal analysis using actigraphy. *International Journal of Obesity*, 43(12), 2555-2564.
- Meredith-Jones, K., Williams, S., Galland, B., Kennedy, G., & Taylor, R. (2016). 24 h Accelerometry: impact of sleep-screening methods on estimates of sedentary behaviour and physical activity while awake. *Journal of Sports Sciences*, 34(7), 679-685.
- Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., ... & Ortega, F. B. (2017). Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports Medicine*, 47(9), 1821-1845.
- Migueles, J. H., Cadenas-Sanchez, C., Tudor-Locke, C., Löf, M., Esteban-Cornejo, I., Molina-Garcia, P., ... & Ortega, F. B. (2019a). Comparability of published cut-points for the assessment of physical activity: Implications for data harmonization. *Scandinavian Journal of Medicine & Science in Sports*, 29(4), 566-574.
- Migueles, J. H., Rowlands, A. V., Huber, F., Sabia, S., & van Hees, V. T. (2019b). GGIR: a research community-driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. *Journal for the Measurement of Physical Behaviour*, 2(3), 188-196.

- Mindell, J. A., Meltzer, L. J., Carskadon, M. A., & Chervin, R. D. (2009). Developmental aspects of sleep hygiene: findings from the 2004 National Sleep Foundation Sleep in America Poll. *Sleep Medicine*, 10(7), 771-779.
- Mindell, J. S., Coombs, N., & Stamatakis, E. (2014). Measuring physical activity in children and adolescents for dietary surveys: practicalities, problems and pitfalls. *Proceedings of the Nutrition Society*, 73(2), 218-225.
- Ministry of Housing, Communities and Local Government. (2015). *English indices of deprivation 2015*. Available at: <http://dclgappp.communities.gov.uk/imd/idmap.html>. (Accessed: 23rd April 2019).
- Ministry of Housing, Communities and Local Government. (2019). *English indices of deprivation 2019*. Available at: [English indices of deprivation 2019: Postcode Lookup \(opendatacommunities.org\)](https://www.opendatacommunities.org/england/indices-of-deprivation-2019). (Accessed: 15th July 2021).
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, 151(4), 264-269.
- Mokkink, L. B., Prinsen, C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., ... & Mokkink, L. (2018a). *COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). User manual*. Available at: [Microsoft Word - COSMIN syst review for PROMs manual version 1 feb 2018](#). (Accessed: 1st June 2021).
- Mokkink, L. B., De Vet, H. C., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018b). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1171-1179.

- Mokkink, L. B., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2019). COSMIN Study Design checklist for Patient-reported outcome measurement instruments. *Amsterdam, The Netherlands*.
- Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., ... & De Vet, H. C. (2010a). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC medical research methodology*, 10(1), 22.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & de Vet, H. C. (2010b). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), 737-745.
- Mölenberg, F. J., de Vries, C., Burdorf, A., & van Lenthe, F. J. (2021). A framework for exploring non-response patterns over time in health surveys. *BMC Medical Research Methodology*, 21(1), 37.
- Montoye, A. H. K., Nelson, M. B., Bock, J. M., Imboden, M. T., Kaminsky, L. A., Mackintosh, K. A., McNarry, M. A. & Pfeiffer, K. A. (2018). Raw and Count Data Comparability of Hip-Worn ActiGraph GT3X+ and Link Accelerometers. *Medicine and Science in Sports and Exercise*, 50(5), 1103-1112.
- Morgan, D. L. 2012. *Focus groups and social interaction*. In Gubrium, J.F., Holstein, J.A., Marvasti, A.B., and McKinney, K.D. (2nd Ed.) *The SAGE Handbook of Interview Research*. Thousand Oaks: SAGE.
- Mouthon, A. L., & Huber, R. (2015). Methods in pediatric sleep research and sleep medicine. *Neuropediatrics*, 46(03), 159-170.

- Murray, M.E. (2009). *Validity and reliability of using MVP 4 Function Walk4Life digital pedometers to assess physical activity levels among preschool-aged Head Start children*. Doctoral dissertation, The University of Texas School of Public Health.
- Nascimento-Ferreira, M. V., Collese, T. S., de Moraes, A. C. F., Rendo-Urteaga, T., Moreno, L. A., & Carvalho, H. B. (2016). Validity and reliability of sleep time questionnaires in children and adolescents: A systematic review and meta-analysis. *Sleep Medicine Reviews, 30*, 85-96.
- National Association for Sport and Physical Education (2011). *Active Start: A Statement of Physical Activity Guidelines for Children From Birth to Age 5, 2nd Edition*. Available at: [http://www.aahperd.org/naspe/standards/nationalGuidelines/Activ \(columbus.gov\)](http://www.aahperd.org/naspe/standards/nationalGuidelines/Activ (columbus.gov)) (Accessed: 23rd March 2020).
- National Center for Education Statistics. (2018). *Types of state and district requirements for kindergarten entrance and attendance, waivers and exemptions for kindergarten entrance, by state: 2018*. Available at: [Table 5.3. Types of state and district requirements for kindergarten entrance and attendance, waivers and exemptions for kindergarten entrance, by state: 2018 \(ed.gov\)](#). (Accessed: 26th March 2020).
- National Sleep Foundation. (2002). *National Sleep Foundation 2003 Sleep In America Poll Screening Questionnaire*. Available at: <https://www.thensf.org/wp-content/uploads/2021/03/2003-SIA-Questionnaire.pdf>. (Accessed: 14th March 2022).
- National Sleep Foundation. (2004). *Summary of findings*. Available at: [National Sleep Foundtion - Sleep in America Poll 2004 \(sleepfoundation.org\)](#). (Accessed:14th March 2022).

- Nelson, B. B., & Bonuck, K. (2015). Healthy sleep in young children: missed opportunity in early childhood programs and policies?. *Sleep Health*, 1(2), 86-87.
- NHS Digital. (2020). *Health Survey for England 2019*. Available at: [Health Survey for England - NHS Digital](#). (Accessed 2nd December 2021).
- NHS Digital. (2016). *Health Survey for England 2015 Physical Activity in Children*. Available at: [hse2015-child-phy-act.pdf \(digital.nhs.uk\)](#). (Accessed: 25th February 2021).
- NHS Digital. (2021). *National Child Measurement Programme England 2020/21 School Year*. Available at: [National Child Measurement Programme, England 2020/21 School Year - NHS Digital](#). (Accessed 23rd March 2022).
- Nishikido, N., Kashiwazaki, H., & Suzuki, T. (1982). Preschool children's daily activities: direct observation, pedometry or questionnaire. *Journal of Human Ergology*, 11(2), 214-218.
- Noland, M., Danner, F., Dewalt, K., McFadden, M., & Kotchen, J. M. (1990). The measurement of physical activity in young children. *Research Quarterly for Exercise and Sport*, 61(2), 146-153.
- Nyström, C. D., Alexandrou, C., Henström, M., Nilsson, E., Okely, A. D., El Masri, S. W., & Löf, M. (2020). International study of movement behaviors in the early years (Sunrise): Results from sunrise sweden's pilot and covid-19 study. *International Journal of Environmental Research and Public Health*, 17(22), 8491.
- Office for National Statistics. (2011). *Population denominators by broad ethnic group and for White British, local authorities in England and Wales: 2011 to 2019*. Available at : <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/adhocs/008781populationdenominatorsbybroadethnigroupan>

- [dforwhitebritishlocalauthoritiesinenglandandwales2011to2017](#). (Accessed 24th August 2021).
- Office for National Statistics. (2021). *Health state life expectancies by national deprivation deciles, England: 2017 to 2019*. Office for National Statistics. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/bulletins/healthstatelifeexpectanciesbyindexofmultipledeprivationimd/2017to2019>. (Accessed 1st November 2021).
- O'hara, N. M., Baranowski, T., Wilson, B. S., Parcel, G. S., & Simons-Morton, B. G. (1989). Validity of the observation of children's physical activity. *Research Quarterly for Exercise and Sport*, 60(1), 42-47.
- Ohayon, M., Wickwire, E. M., Hirshkowitz, M., Albert, S. M., Avidan, A., Daly, F. J., ... & Vitiello, M. V. (2017). National Sleep Foundation's sleep quality recommendations: first report. *Sleep health*, 3(1), 6-19.
- Ojiambo, R., Cuthill, R., Budd, H., Konstabel, K., Casajus, J. A., González-Agüero, A., ... & Pitsiladis, Y. P. (2011). Impact of methodological decisions on accelerometer outcome variables in young children. *International Journal of Obesity*, 35(1), S98-S103.
- Okely, A. D., Ghersi, D., Hesketh, K. D., Santos, R., Loughran, S. P., Cliff, D. P., ... & Tremblay, M. S. (2017). A collaborative approach to adopting/adapting guidelines-The Australian 24-Hour Movement Guidelines for the early years (Birth to 5 years): an integration of physical activity, sedentary behavior, and sleep. *BMC Public Health*, 17(5), 869.
- Okely, A. D., Kariippanon, K. E., Guan, H., Taylor, E. K., Suesse, T., Cross, P. L., ... & Draper, C. E. (2021a). Global effect of COVID-19 pandemic on physical activity, sedentary

- behaviour and sleep among 3-to 5-year-old children: a longitudinal study of 14 countries. *BMC Public Health*, 21(1), 940.
- Okely, A. D., Reilly, J. J., Tremblay, M. S., Kariippanon, K. E., Draper, C. E., El Hamdouchi, A., ... & Widyastari, D. A. (2021b). Cross-sectional examination of 24-hour movement behaviours among 3-and 4-year-old children in urban and rural settings in low-income, middle-income and high-income countries: the SUNRISE study protocol. *BMJ Open*, 11(10), e049267.
- Okely, A. D., Tremblay, M. S., Reilly, J. J., Draper, C. E., & Bull, F. (2018). Physical activity, sedentary behaviour, and sleep: movement behaviours in early life. *The Lancet Child & Adolescent Health*, 2(4), 233-235.
- Oliver, M., Schofield, G. M., Kolt, G. S., & Schluter, P. J. (2007). Pedometer accuracy in physical activity assessment of preschool children. *Journal of Science and Medicine in Sport*, 10(5), 303-310.
- Oliver, M., Schofield, G. M., & Kolt, G. S. (2007). Physical activity in preschoolers. *Sports Medicine*, 37(12), 1045-1070.
- Olshansky, E., Lakes, K. D., Vaughan, J., Gravem, D., Rich, J. K., David, M., ... & Cooper, D. (2012). Enhancing the construct and content validity of rating scales for clinical research: Using qualitative methods to develop a rating scale to assess parental perceptions of their role in promoting infant exercise. *The International Journal of Educational and Psychological Assessment*, 10(1), 36-50.
- Ong, A. A., & Gillespie, M. B. (2016). Overview of smartphone applications for sleep analysis. *World Journal of Otorhinolaryngology-head and Neck Surgery*, 2(1), 45-49.

- Oortwijn, A., Plasqui, G., Reilly, J. J., & Okely, A. D. (2009). Feasibility of an activity protocol for young children in a whole room indirect calorimeter: a proof-of-concept study. *Journal of Physical Activity and Health*, 6 633-637.
- Ordway, M. R., Sadler, L. S., Jeon, S., O'Connell, M., Banasiak, N., Fenick, A. M., ... & Redeker, N. S. (2020). Sleep health in young children living with socioeconomic adversity. *Research in Nursing & Health*, 43(4), 329-340.
- Ortega, F. B., Cadenas-Sánchez, C., Sánchez-Delgado, G., Mora-González, J., Martínez-Téllez, B., Artero, E. G., ... & Ruiz, J. R. (2015). Systematic review and proposal of a field-based physical fitness-test battery in preschool children: the PREFIT battery. *Sports Medicine*, 45(4), 533-555.
- Owen, A., & Bould, K. (2021). Reduced physical activity and increased sedentary behaviour: the damage on young people during the COVID-19 pandemic. *British Journal of Child Health*, 2(2), 64-68.
- Owens, J. A., Spirito, A., & McGuinn, M. (2000). The Children's Sleep Habits Questionnaire (CSHQ): psychometric properties of a survey instrument for school-aged children. *Sleep*, 23(8), 1043-1052.
- Paavonen, E. J., Fjällberg, M., Steenari, M. R., & Aronen, E. T. (2002). Actigraph placement and sleep estimation in children. *Sleep*, 25(2), 235-237.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88, 105906.
- Pagels, P., Boldemann, C., & Raustorp, A. (2011). Comparison of pedometer and accelerometer measures of physical activity during preschool time on 3-to 5-year-old children. *Acta Paediatrica*, 100(1), 116-120.

- Paruthi, S., Brooks, L. J., D'Ambrosio, C., Hall, W. A., Kotagal, S., Lloyd, R. M., ... & Wise, M. S. (2016). Consensus statement of the American Academy of Sleep Medicine on the recommended amount of sleep for healthy children: methodology and discussion. *Journal of Clinical Sleep Medicine*, 12(11), 1549-1561.
- Pate, R. R., Almeida, M. J., McIver, K. L., Pfeiffer, K. A., & Dowda, M. (2006). Validation and calibration of an accelerometer in preschool children. *Obesity*, 14(11), 2000-2006.
- Pate, R. R., Mitchell, J. A., Byun, W., & Dowda, M. (2011). Sedentary behaviour in youth. *British Journal of Sports Medicine*, 45(11), 906-913.
- Pate, R. R., O'Neill, J. R., & Mitchell, J. (2010). Measurement of physical activity in preschool children. *Medicine and Science in Sports and Exercise*, 42(3), 508-512.
- Patel, P., Kim, J. Y., & Brooks, L. J. (2017). Accuracy of a smartphone application in estimating sleep in children. *Sleep and Breathing*, 21(2), 505-511.
- Pearce, A., Dundas, R., Whitehead, M., & Taylor-Robinson, D. (2019). Pathways to inequalities in child health. *Archives of Disease in Childhood*, 104(10), 998-1003.
- Pedišić, Ž., & Bauman, A. (2015). Accelerometer-based measures in physical activity surveillance: current practices and issues. *British Journal of Sports Medicine*, 49(4), 219-223.
- Pedišić, Ž., Dumuid, D., & S Olds, T. (2017). Integrating sleep, sedentary behaviour, and physical activity research in the emerging field of time-use epidemiology: definitions, concepts, statistical methods, theoretical framework, and future directions. *Kinesiology*, 49(2.), 252-269.
- Penpraze, V., Reilly, J. J., MacLean, C. M., Montgomery, C., Kelly, L. A., Paton, J. Y., ... & Grant, S. (2006). Monitoring of physical activity in young children: how much is enough?. *Pediatric Exercise Science*, 18(4), 483-491.

- Perpétuo, C., Fernandes, M., & Veríssimo, M. (2020). Comparison between actigraphy records and parental reports of child's sleep. *Frontiers in Pediatrics*, 606.
- Petrillo, J., Cano, S. J., McLeod, L. D., & Coon, C. D. (2015). Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value in Health*, 18(1), 25-34.
- Pfeiffer, K. A., Mciver, K. L., Dowda, M., Almeida, M. J. C. A., & Pate, R. R. (2006). Validation and calibration of the Actical accelerometer in preschool children. *Medicine and science in sports and exercise*, 38(1), 152-157.
- Pickett, K., Taylor-Robinson D., et al. (2021). The Child of the North: Building a fairer future after COVID-19, the Northern Health Science Alliance and N8 Research Partnership.
- Platt, L., & Warwick, R. (2020). Are some ethnic groups more vulnerable to COVID-19 than others. *Institute for Fiscal Studies*, 1(05), 2020.
- Poitras, V. J., Gray, C. E., Janssen, X., Aubert, S., Carson, V., Faulkner, G., ... & Tremblay, M. S. (2017). Systematic review of the relationships between sedentary behaviour and health indicators in the early years (0–4 years). *BMC Public Health*, 17(5), 868.
- Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Gorber, S. C., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 5(1), 56.
- Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1147-1157.

- Public Health England. (2018). *Inequalities in Health*. Available at: <https://www.gov.uk/government/publications/health-profile-for-england-2018/chapter-5-inequalities-in-health#inequalities-in-child-health>. (Accessed: 20 July 2019).
- Puhl, J., Greaves, K., Hoyt, M., & Baranowski, T. (1990). Children's Activity Rating Scale (CARS): description and calibration. *Research Quarterly for Exercise and Sport*, 61(1), 26-36.
- Puyau, M. R., Adolph, A. L., Vohra, F. A., & Butte, N. F. (2002). Validation and calibration of physical activity monitors in children. *Obesity Research*, 10(3), 150-157.
- Puyau, M. R., Adolph, A. L., Liu, Y., Wilson, T. A., Zakeri, I. F., & Butte, N. F. (2016). Energy cost of activities in preschool-aged children. *Journal of Physical Activity and Health*, 13(s1), S11-S16.
- Quante, M., Kaplan, E. R., Rueschman, M., Cailler, M., Buxton, O. M., & Redline, S. (2015). Practical considerations in using accelerometers to assess physical activity, sedentary behavior, and sleep. *Sleep Health*, 1(4), 275-284.
- Rao, D. P., Orpana, H., & Krewski, D. (2016). Physical activity and non-movement behaviours: their independent and combined associations with metabolic syndrome. *International Journal of Behavioral Nutrition and Physical Activity*, 13(1), 26.
- Reilly, J. J., Coyle, J., Kelly, L., Burke, G., Grant, S., & Paton, J. Y. (2003). An objective method for measurement of sedentary behavior in 3-to 4-year olds. *Obesity Research*, 11(10), 1155-1158.
- Reilly, J. J., Hughes, A., Janssen, X., Martin, A., Hesketh, K., Livingstone, S., ... & Okely, A. (2018). Expert Working Group Working Paper, Under 5s. *UK physical activity*

- guidelines: Draft review and recommendations for the Under 5s*. Available at: [under-5s-technical-report.pdf \(bristol.ac.uk\)](#). (Accessed: 23rd April 2019)
- Reilly, J. J., Hughes, A. R., Janssen, X., Hesketh, K. R., Livingstone, S., Hill, C., ... & Martin, A. (2020). GRADE-ADOLOPMENT process to develop 24-hour movement behavior recommendations and physical activity guidelines for the under 5s in the United Kingdom, 2019. *Journal of Physical Activity and Health*, 17(1), 101-108.
- Reilly, J. J., Janssen, X., Cliff, D. P., & Okely, A. D. (2015). Appropriateness of the definition of 'sedentary' in young children: Whole-room calorimetry study. *Journal of Science and Medicine in Sport*, 18(5), 565-568.
- Reilly, J. J., Kelly, L. A., Montgomery, C., Jackson, D. M., Slater, C., Grant, S., & Paton, J. Y. (2006). Validation of Actigraph accelerometer estimates of total energy expenditure in young children. *International Journal of Pediatric Obesity*, 1(3), 161-167.
- Rennie, K. L., & Wareham, N. J. (1998). The validation of physical activity instruments for measuring energy expenditure: problems and pitfalls. *Public Health Nutrition*, 1(4), 265-271.
- Riazi, N., Ramanathan, S., O'Neill, M., Tremblay, M. S., & Faulkner, G. (2017). Canadian 24-hour movement guidelines for the early years (0–4 years): exploring the perceptions of stakeholders and end users regarding their acceptability, barriers to uptake, and dissemination. *BMC Public Health*, 17(5), 841.
- Ricardo, L. I. C., Wendt, A., Galliano, L. M., de Andrade Muller, W., Niño Cruz, G. I., Wehrmeister, F., ... & Crochemore M. Silva, I. (2020). Number of days required to estimate physical activity constructs objectively measured in different age groups: findings from three Brazilian (Pelotas) population-based birth cohorts. *PloS one*, 15(1), e0216017.

- Robinson, J. C., Temple, M. L., Duck, A., & Klamm, M. (2019). Feasibility and effectiveness of two built environmental interventions on physical activity among 3–5-year-old preschoolers. *Journal for Specialists in Pediatric Nursing, 24*(3), e12262.
- Rolfe, D. E., Ramsden, V. R., Banner, D., & Graham, I. D. (2018). Using qualitative health research methods to improve patient and public involvement and engagement in research. *Research Involvement and Engagement, 4*(1), 49.
- Rollo, S., Antsygina, O., & Tremblay, M. S. (2020). The whole day matters: understanding 24-hour movement guideline adherence and relationships with health indicators across the lifespan. *Journal of Sport and Health Science, 9*(6), 493-510.
- Romanzini, C. L. P., Romanzini, M., Batista, M. B., Barbosa, C. C. L., Shigaki, G. B., Dunton, G., ... & Ronque, E. R. V. (2019). Methodology used in ecological momentary assessment studies about sedentary behavior in children, adolescents, and adults: systematic review using the checklist for reporting ecological momentary assessment studies. *Journal of Medical Internet Research, 21*(5), e11967.
- Rosenberger, M. E., Fulton, J. E., Buman, M. P., Troiano, R. P., Grandner, M. A., Buchner, D. M., & Haskell, W. L. (2019). The 24-hour activity cycle: a new paradigm for physical activity. *Medicine and Science in Sports and Exercise, 51*(3), 454.
- Roth, M., & Stamatakis, E. (2010). Linking young people's knowledge of public health guidelines to physical activity levels in England. *Pediatric Exercise Science, 22*(3), 467-476.
- Rothman, M. L., Beltran, P., Cappelleri, J. C., Lipscomb, J., Teschendorf, B., & Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. (2007). Patient-reported outcomes: conceptual issues. *Value in Health, 10*, S66-S75.

- Rowe, D. A. (2011). Back to the future? Algorithms and equipment vs. simplicity and common sense in physical activity measurement. *International Journal of Human Movement Science*, 5(2), 25-45.
- Rowlands, A. V., Thomas, P. W., Eston, R. G., & Topping, R. (2004). Validation of the RT3 triaxial accelerometer for the assessment of physical activity. *Medicine and Science in Sports and Exercise*, 36(3), 518-524.
- Sadeh, A. (2011). The role and validity of actigraphy in sleep medicine: an update. *Sleep Medicine Reviews*, 15(4), 259-267.
- Sadeh, A. (2015). Iii. Sleep assessment methods. *Monographs of the Society for Research in Child Development*, 80(1), 33-48.
- Sadeh, A., & Acebo, C. (2002). The role of actigraphy in sleep medicine. *Sleep Medicine Reviews*, 6(2), 113-124.
- Sadeh, A., Sharkey, M., & Carskadon, M. A. (1994). Activity-based sleep-wake identification: an empirical test of methodological issues. *Sleep*, 17(3), 201-207.
- Sakzewski, L., Reedman, S., McLeod, K., Thorley, M., Burgess, A., Trost, S., ... & Boyd, R. N. (2021). Preschool HABIT-ILE: study protocol for a randomised controlled trial to determine efficacy of intensive rehabilitation compared with usual care to improve motor skills of children, aged 2–5 years, with bilateral cerebral palsy. *BMJ Open*, 11(3), e041542.
- Sallis, J. F. (1991). Self-report measures of children's physical activity. *Journal of School Health*, 61(5), 215-220.
- Santos, R., Mota, J., Okely, A. D., Pratt, M., Moreira, C., Coelho-e-Silva, M. J., ... & Sardinha, L. B. (2014). The independent associations of sedentary behaviour and physical

- activity on cardiorespiratory fitness. *British Journal of Sports Medicine*, 48(20), 1508-1512.
- Saris, W. H. M., & Binkhorst, R. A. (1977). The use of pedometer and actometer in studying daily physical activity in man. Part II: validity of pedometer and actometer measuring the daily physical activity. *European Journal of Applied Physiology and Occupational Physiology*, 37(3), 229-235.
- Sattler, M. C., Ainsworth, B. E., Andersen, L. B., Foster, C., Hagströmer, M., Jaunig, J., ... & Van Poppel, M. N. (2021). Physical activity self-reports: past or future?. *British Journal of Sports Medicine*, 55(16), 889-890.
- Schaefer, C. A., Nace, H., & Browning, R. (2014). Establishing wrist-based cutpoints for the actical accelerometer in elementary school-aged children. *Journal of Physical Activity and Health*, 11(3), 604-613.
- Schmidt, M. E., & Steindorf, K. (2006). Statistical methods for the validation of questionnaires. *Methods of Information in Medicine*, 45(04), 409-413.
- Schneider, P. L., Crouter, S. E., & Bassett, D. R. (2004). Pedometer measures of free-living physical activity: comparison of 13 models. *Medicine and Science in Sports and Exercise*, 36(2), 331-335.
- Schoch, S. F., Kurth, S., & Werner, H. (2021). Actigraphy in sleep research with infants and young children: Current practices and future benefits of standardized reporting. *Journal of Sleep Research*, 30, e13134.
- Scottish Government. (2019). *Scottish Health Survey 2019 - volume 1: main report*. Available at: [Chapter 7: Physical Activity - Scottish Health Survey 2019 - volume 1: main report - gov.scot \(www.gov.scot\)](https://www.gov.scot/publications/scottish-health-survey-2019-volume-1-main-report/pages/12.aspx). (Accessed: 10 February 2022).

- Sekine, M., Chen, X., Hamanishi, S., Wang, H., Yamagami, T., & Kagamimori, S. (2002). The validity of sleeping hours of healthy young children as reported by their parents. *Journal of Epidemiology*, 12(3), 237-242.
- Sen, T., & Spruyt, K. (2020). Pediatric sleep tools: an updated literature review. *Frontiers in Psychiatry*, 317.
- Sharma, S. V., Chuang, R. J., Skala, K., & Atteberry, H. (2011). Measuring physical activity in preschoolers: reliability and validity of the System for Observing Fitness Instruction Time for Preschoolers (SOFIT-P). *Measurement in Physical Education and Exercise Science*, 15(4), 257-273.
- Sharp, C. A., Mackintosh, K. A., Erjavec, M., Pascoe, D. M., & Horne, P. J. (2017). Validity and reliability of the Fitbit Zip as a measure of preschool children's step count. *BMJ Open Sport & Exercise Medicine*, 3(1), e000272.
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., ... & Bouter, L. M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, 7(1), 10.
- Shephard, R. J. (2003). Limits to the measurement of habitual physical activity by questionnaires. *British Journal of Sports Medicine*, 37(3), 197-206.
- Shin, J.C. (2015). Calibration of an accelerometer to measure physical activity in preschool children: A feasibility study. Masters Dissertation, University of Illinois at Urbana-Champaign.
- Sijtsma, A., Schierbeek, H., Goris, A. H., Joosten, K. F., van Kessel, I., Corpeleijn, E., & Sauer, P. J. (2013). Validation of the TracmorD triaxial accelerometer to assess physical activity in preschool children. *Obesity*, 21(9), 1877-1883.

- Silverman, D. (2017). *Doing qualitative research*. Sage.
- Sims, J., Milton, K., Foster, C., & Scarborough, P. (2021). A profile of children's physical activity data from the 2012 and 2015 Health Survey for England. *Research Square*.
- Sinha, I., Bennett, D., & Taylor-Robinson, D. C. (2020). Children are being sidelined by covid-19. *Bmj*, 369, m2061.
- Sirard, J. R., & Pate, R. R. (2001). Physical activity assessment in children and adolescents. *Sports Medicine*, 31(6), 439-454.
- Sirard, J. R., Trost, S. G., Pfeiffer, K. A., Dowda, M., & Pate, R. R. (2005). Calibration and evaluation of an objective measure of physical activity in preschool children. *Journal of Physical Activity and Health*, 2(3), 345-357.
- Sitnick, S. L., Goodlin-Jones, B. L., & Anders, T. F. (2008). The use of actigraphy to study sleep disorders in preschoolers: some concerns about detection of nighttime awakenings. *Sleep*, 31(3), 395-401.
- Skouteris, H., Dell'Aquila, D., Baur, L. A., Dwyer, G. M., McCabe, M. P., Ricciardelli, L. A., & Fuller-Tyszkiewicz, M. (2012). Physical activity guidelines for preschoolers: a call for research to inform public health policy. *Medical Journal of Australia*, 196(3), 174-176.
- Slater, J. A., Botsis, T., Walsh, J., King, S., Straker, L. M., & Eastwood, P. R. (2015). Assessing sleep using hip and wrist actigraphy. *Sleep and Biological Rhythms*, 13(2), 172-180.
- Smith, B. (2018). Generalizability in qualitative research: Misunderstandings, opportunities and recommendations for the sport and exercise sciences. *Qualitative Research in Sport, Exercise and Health*, 10(1), 137-149.
- Smithson, J. (2000). Using and analysing focus groups: limitations and possibilities. *International Journal of Social Research Methodology*, 3(2), 103-119.

- Sport England. (2020). *Active Lives, Children and young people survey*. Available at: <https://www.sportengland.org/know-your-audience/data/active-lives?section=methodology#childrenandyoungpeoplesurvey-11678>. (Accessed 8th October 2021).
- Spruyt, K., & Gozal, D. (2011). Pediatric sleep questionnaires as diagnostic or epidemiological tools: a review of currently available instruments. *Sleep Medicine Reviews*, 15(1), 19-32.
- Stanley, R., Jones, R., Swann, C., Christian, H., Sherring, J., Shilton, T., & Okely, A. (2020). Exploring stakeholders' perceptions of the acceptability, usability, and dissemination of the Australian 24-hour movement guidelines for the early years. *Journal of Physical Activity and Health*, 17(1), 120-125.
- Staples, A. D., Bates, J. E., Petersen, I. T., McQuillan, M. E., & Hoyniak, C. (2019). Measuring sleep in young children and their mothers: Identifying actigraphic sleep composites. *International Journal of Behavioral Development*, 43(3), 278-285.
- Steenbock, B., Wright, M. N., Wirsik, N., & Brandes, M. (2019). Accelerometry-based prediction of energy expenditure in preschoolers. *Journal for the Measurement of Physical Behaviour*, 2(2), 94-102.
- Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., ... & Higgins, J. P. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj*, 366.
- Strain, T., Milton, K., Dall, P., Standage, M., & Mutrie, N. (2020). How are we measuring physical activity and sedentary behaviour in the four home nations of the UK? A narrative review of current surveillance measures and future directions. *British Journal of Sports Medicine*, 54(21), 1269-1276.

- Strain, T., Wijndaele, K., Pearce, M., & Brage, S. (2022). Considerations for the use of consumer-grade wearables and smartphones in population surveillance of physical activity. *Journal for the Measurement of Physical Behaviour*, 5(1), 8-14.
- Summerbell, C. D., Moore, H. J., Vögele, C., Kreichauf, S., Wildgruber, A., Manios, Y., ... & ToyBox-study group. (2012). Evidence-based recommendations for the development of obesity prevention programs targeted at preschool children. *Obesity Reviews*, 13, 129-132.
- Sun, D. X., Schmidt, G., & Teo-Koh, S. M. (2008). Validation of the RT3 accelerometer for measuring physical activity of children in simulated free-living conditions. *Pediatric Exercise Science*, 20(2), 181-197.
- SUNRISE. (2020). *International Study of Movement Behaviours in the Early Years*. Available at: <https://sunrise-study.com/> (Accessed 29th March 2022).
- Sweetser, P., Johnson, D., Ozdowska, A., & Wyeth, P. (2012). Active versus passive screen time for young children. *Australasian Journal of Early Childhood*, 37(4), 94-98.
- Tan, T. X., Wang, Y., Cheah, C. S. L., & Wang, G. H. (2018). Reliability and construct validity of the Children's Sleep Habits Questionnaire in Chinese kindergartners. *Sleep Health*, 4(1), 104-109.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53.
- Taylor, R. W., Haszard, J. J., Healey, D., Meredith-Jones, K. A., Taylor, B. J., & Galland, B. C. (2021). Adherence to 24-h movement behavior guidelines and psychosocial functioning in young children: a longitudinal analysis. *International Journal of Behavioral Nutrition and Physical Activity*, 18(1), 110.

- Taylor, R. W., Haszard, J. J., Meredith-Jones, K. A., Galland, B. C., Heath, A. L. M., Lawrence, J., ... & Taylor, B. J. (2018). 24-h movement behaviors from infancy to preschool: cross-sectional and longitudinal relationships with body composition and bone health. *International Journal of Behavioral Nutrition and Physical Activity*, 15(1), 118.
- Taylor, R. W., Williams, S. M., Farmer, V. L., & Taylor, B. J. (2015). The stability of sleep patterns in children 3 to 7 years of age. *The Journal of Pediatrics*, 166(3), 697-702.
- Telama, R., Yang, X., Leskinen, E., Kankaanpää, A., Hirvensalo, M., Tammelin, T., & Raitakari, O. T. (2014). Tracking of physical activity from early childhood through youth into adulthood. *Medicine and Science in Sports and Exercise*, 46(5), 955-962.
- Telford, A., Salmon, J., Jolley, D., & Crawford, D. (2004). Reliability and validity of physical activity questionnaires for children: The Children's Leisure Activities Study Survey (CLASS). *Pediatric Exercise Science*, 16(1), 64-78.
- Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., ... & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34-42.
- Terwee, C. B., Mokkink, L. B., Hidding, L. M., Altenburg, T. M., van Poppel, M. N., & Chinapaw, M. J. M. (2016). Comment on "Should we reframe how we think about physical activity and sedentary behavior measurement? Validity and reliability reconsidered". *International Journal of Behavioral Nutrition and Physical Activity*, 13(1), 66.
- Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W., Bouter, L. M., & de Vet, H. C. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research*, 21(4), 651-657.

- Thompson, D., Peacock, O., Western, M., & Batterham, A. M. (2015). Multidimensional physical activity: an opportunity not a problem. *Exercise and Sport Sciences Reviews, 43*(2), 67.
- Tikotzky, L., & Sadeh, A. (2001). Sleep patterns and sleep disruptions in kindergarten children. *Journal of Clinical Child Psychology, 30*(4), 581-591.
- Timmons, B. W., LeBlanc, A. G., Carson, V., Connor Gorber, S., Dillman, C., Janssen, I., ... & Tremblay, M. S. (2012). Systematic review of physical activity and health in the early years (aged 0–4 years). *Applied Physiology, Nutrition, and Metabolism, 37*(4), 773-792.
- Tolonen, H., Dobson, A., & Kulathinal, S. (2005). Effect on trend estimates of the difference between survey respondents and non-respondents: results from 27 populations in the WHO MONICA Project. *European Journal of Epidemiology, 20*(11), 887-898.
- Tolonen, H., Helakorpi, S., Talala, K., Helasoja, V., Martelin, T., & Prättälä, R. (2006). 25-year trends and socio-demographic differences in response rates: Finnish adult health behaviour survey. *European Journal of Epidemiology, 21*(6), 409-415.
- Tomaz, S. A., Okely, A. D., Van Heerden, A., Vilakazi, K., Samuels, M. L., & Draper, C. E. (2020). The South African 24-hour movement guidelines for birth to 5 years: results from the stakeholder consultation. *Journal of Physical Activity and Health, 17*(1), 126-137.
- Townsend, M. S., Shilts, M., Ontai, L., Leavens, L., Davidson, C., & Sitnick, S. (2021). Obesity risk for young children: development and initial validation of an assessment tool for participants of federal nutrition programs. Available at: [Obesity risk for young children: Development and initial validation of an assessment tool for participants of federal nutrition programs \(escholarship.org\)](https://escholarship.org/uc/item/9kz7d) (Accessed: 28th March 2022).

- Tracy, D. J., Xu, Z., Choi, L., Acra, S., Chen, K. Y., & Buchowski, M. S. (2014). Separating bedtime rest from activity using waist or wrist-worn accelerometers in youth. *PLoS One*, 9(4), e92512.
- Tracy, J. D., Donnelly, T., Sommer, E. C., Heerman, W. J., Barkin, S. L., & Buchowski, M. S. (2021). Identifying bedrest using waist-worn triaxial accelerometers in preschool children. *PloS One*, 16(1), e0246055.
- Tremblay, M. S. (2020). Introducing 24-hour movement guidelines for the early years: a new paradigm gaining momentum. *Journal of Physical Activity and Health*, 17(1), 92-95.
- Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., ... & Chinapaw, M. J. (2017a). Sedentary behavior research network (SBRN)—terminology consensus project process and outcome. *International Jjournal of Behavioral Nutrition and Physical Activity*, 14(1), 75
- Tremblay, M. S., Barnes, J. D., González, S. A., Katzmarzyk, P. T., Onywera, V. O., Reilly, J. J., & Tomkinson, G. R. (2016). Global matrix 2.0: report card grades on the physical activity of children and youth comparing 38 countries. *Journal of Physical Activity and Health*, 13, S343-S366.
- Tremblay, M. S., Chaput, J. P., Adamo, K. B., Aubert, S., Barnes, J. D., Choquette, L., ... & Carson, V. (2017b). Canadian 24-hour movement guidelines for the early years (0–4 years): an integration of physical activity, sedentary behaviour, and sleep. *BMC Public Health*, 17(5), 874.
- Tremblay, M. S., Colley, R. C., Saunders, T. J., Healy, G. N., & Owen, N. (2010). Physiological and health implications of a sedentary lifestyle. *Applied Physiology, Nutrition, and Metabolism*, 35(6), 725-740.

- Tremblay, M. S., & Gorber, S. C. (2007). Canadian health measures survey. *Canadian Journal of Public Health, 98*(6), 453-456.
- Tremblay, M. S., LeBlanc, A. G., Carson, V., Choquette, L., Connor Gorber, S., Dillman, C., ... & Timmons, B. W. (2012). Canadian physical activity guidelines for the early years (aged 0–4 years). *Applied Physiology, Nutrition, and Metabolism, 37*(2), 345-356.
- Tremblay, M. S., & Ross, R. (2020). How should we move for health? The case for the 24-hour movement paradigm. *Canadian Medical Association Journal, 192*(49), E1728-E1729.
- Troiano, R. P. (2009). Can there be a single best measure of reported physical activity?. *The American Journal of Clinical Nutrition, 89*(3), 736-737.
- Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British Journal of Sports Medicine, 48*(13), 1019-1023
- Troiano, R. P., Stamatakis, E., & Bull, F. C. (2020). How can global physical activity surveillance adapt to evolving physical activity guidelines? Needs, challenges and future directions. *British Journal of Sports Medicine, 54*(24), 1468-1473..
- Trost, S. G. (2007). State of the art reviews: measurement of physical activity in children and adolescents. *American Journal of Lifestyle Medicine, 1*(4), 299-314.
- Trost, S. G., Mciver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and Science in Sports and Exercise, 37*(11), S531-S543.
- Truelove, S., Vanderloo, L. M., & Tucker, P. (2017). Defining and measuring active play among young children: a systematic review. *Journal of Physical Activity and Health, 14*(2), 155-166.

- Tso, W., Rao, N., Jiang, F., Li, A. M., Lee, S. L., Ho, F. K. W., ... & Ip, P. (2016). Sleep duration and school readiness of Chinese preschool children. *The Journal of Pediatrics*, 169, 266-271.
- Tudor-Locke, C., Barreira, T. V., Schuna Jr, J. M., Mire, E. F., & Katzmarzyk, P. T. (2014). Fully automated waist-worn accelerometer algorithm for detecting children's sleep-period time separate from 24-h physical activity or sedentary behaviors. *Applied Physiology, Nutrition, and Metabolism*, 39(1), 53-57.
- Tulchin-Francis, K., Stevens Jr, W., Gu, X., Zhang, T., Roberts, H., Keller, J., ... & VanPelt, J. (2021). The impact of the coronavirus disease 2019 pandemic on physical activity in US children. *Journal of Sport and Health Science*, 10(3), 323-332.
- Twomey, P. J., & Kroll, M. H. (2008). How to use linear regression and correlation in quantitative method comparison studies. *International Journal of Clinical Practice*, 62(4), 529-538.
- U.S. Department of Health and Human Services. (2018). *Physical Activity Guidelines for Americans, 2nd edition*. Available at: [Physical Activity Guidelines for Americans, 2nd edition \(health.gov\)](https://www.health.gov/physical-activity-guidelines-for-americans-2nd-edition). (Accessed: 30th March 2022).
- Vale, S., & Mota, J. (2020). Adherence to 24-hour movement guidelines among Portuguese preschool children: The prestyle study. *Journal of Sports Sciences*, 38(18), 2149-2154.
- Vale, S., Santos, R., Silva, P., Soares-Miranda, L., & Mota, J. (2009). Preschool children physical activity measurement: importance of epoch length choice. *Pediatric Exercise Science*, 21(4), 413-420.
- Vale, S., Silva, P., Santos, R., Soares-Miranda, L., & Mota, J. (2010). Compliance with physical activity guidelines in preschool children. *Journal of Sports Sciences*, 28(6), 603-608.

- Van Cauwenberghe, E., Gubbels, J., De Bourdeaudhuij, I., & Cardon, G. (2011a). Feasibility and validity of accelerometer measurements to assess physical activity in toddlers. *International Journal of Behavioral Nutrition and Physical Activity*, 8(1), 67.
- Van Cauwenberghe, E., Labarque, V., Trost, S. G., De Bourdeaudhuij, I., & Cardon, G. (2011b). Calibration and comparison of accelerometer cut points in preschool children. *International Journal of Pediatric Obesity*, 6, e582-e589.
- Van Cauwenberghe, E., Wooller, L., Mackay, L., Cardon, G., & Oliver, M. (2012). Comparison of Actical and activPAL measures of sedentary behaviour in preschool children. *Journal of Science and Medicine in Sport*, 15(6), 526-531.
- Van den Bulck, J. (2015). Sleep apps and the quantified self: blessing or curse?. *Journal of Sleep Research*, 24(2), 121-123.
- van der Ploeg, H. P., & Hillsdon, M. (2017). Is sedentary behaviour just physical inactivity by another name?. *International Journal of Behavioral Nutrition and Physical Activity*, 14(1), 142.
- Van Ekris, E., Wijndaele, K., Altenburg, T. M., Atkin, A. J., Twisk, J., Andersen, L. B., ... & Chinapaw, M. (2020). Tracking of total sedentary time and sedentary patterns in youth: a pooled analysis using the International Children's Accelerometry Database (ICAD). *International Journal of Behavioral Nutrition and Physical Activity*, 17(1), 65.
- Van Loon, A. J. M., Tijhuis, M., Picavet, H. S. J., Surtees, P. G., & Ormel, J. (2003). Survey non-response in the Netherlands: effects on prevalence estimates and associations. *Annals of Epidemiology*, 13(2), 105-110.
- van Poppel, M. N., Chinapaw, M. J., Mekkink, L. B., Van Mechelen, W., & Terwee, C. B. (2010). Physical activity questionnaires for adults. *Sports Medicine*, 40(7), 565-600.

- Vanderloo, L. M., Di Cristofaro, N. A., Proudfoot, N. A., Tucker, P., & Timmons, B. W. (2016). Comparing the Actical and ActiGraph approach to measuring young children's physical activity levels and sedentary time. *Pediatric Exercise Science, 28*(1), 133-142.
- Vanhelst, J., Béghin, L., Turck, D., & Gottrand, F. (2011). New validated thresholds for various intensities of physical activity in adolescents using the Actigraph accelerometer. *International Journal of Rehabilitation Research, 34*(2), 175-177.
- Vaughn, B. E., Elmore-Staton, L., Shin, N., & El-Sheikh, M. (2015). Sleep as a support for social competence, peer relations, and cognitive functioning in preschool children. *Behavioral Sleep Medicine, 13*(2), 92-106.
- Welk, G. J., Corbin, C. B., & Dale, D. (2000). Measurement issues in the assessment of physical activity in children. *Research Quarterly for Exercise and Sport, 71*, 59-73.
- Wells, G., & Shay, B. (2000). Data extraction for non-randomised systematic reviews. University of Ottawa, Ottawa.
- Welsh Government. (2018). *National Survey for Wales 2017-18 questionnaire*. Available at: <https://gov.wales/sites/default/files/statistics-and-research/2020-01/national-survey-wales-questionnaire-2017-18.pdf>. (Accessed: 10 February 2022).
- Wen, L. M., Van Der Ploeg, H. P., Kite, J., Cashmore, A., & Rissel, C. (2010). A validation study of assessing physical activity and sedentary behavior in children aged 3 to 5 years. *Pediatric Exercise Science, 22*(3), 408-420.
- Werner, H., Molinari, L., Guyer, C., & Jenni, O. G. (2008). Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns. *Archives of Pediatrics & Adolescent Medicine, 162*(4), 350-358.

- Westerterp, K. R. (2017). Doubly labelled water assessment of energy expenditure: principle, practice, and promise. *European Journal of Applied Physiology*, 117(7), 1277-1285.
- Whitehead, M., Taylor-Robinson, D., & Barr, B. (2021). Poverty, health, and covid-19. *Bmj*, 372.
- Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., ... & QUADAS-2 Group. (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529-536.
- Wickström, G., & Bendix, T. (2000). The "Hawthorne effect"—what did the original Hawthorne studies actually show?. *Scandinavian Journal of Work, Environment & Health*, 363-367.
- Williamson, A. A., Gould, R., Leichman, E. S., Walters, R. M., & Mindell, J. A. (2021). Socioeconomic disadvantage and sleep in early childhood: Real-world data from a mobile health application. *Sleep Health*, 7(2), 143-152.
- Willig, C. (2013). *Introducing qualitative research in psychology*. McGraw-hill education:UK.
- Willumsen, J., & Bull, F. (2020). Development of WHO guidelines on physical activity, sedentary behavior, and sleep for children less than 5 years of age. *Journal of Physical Activity and Health*, 17(1), 96-100.
- Wilson, K. E., Miller, A. L., Lumeng, J. C., & Chervin, R. D. (2014). Sleep environments and sleep durations in a sample of low-income preschool children. *Journal of Clinical Sleep Medicine*, 10(3), 299-305.
- World Health Organization. (2007). *A guide for population-based approaches to increasing levels of physical activity: implementation of the WHO global strategy on diet,*

physical activity and health. Available at: [9789241595179_eng.pdf \(who.int\)](#).

(Accessed: 31st December 2021).

World Health Organization. (2018). *Global Action Plan on Physical Activity 2018-2030*.

Available at: [9789241514187-eng.pdf \(who.int\)](#). (Accessed: 2nd December 2021).

World Health Organization. (2019). *Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age*. Available at: [Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age \(who.int\)](#).

(Accessed: 1st May 2019).

(Accessed: 1st May 2019).

World Health Organization. (2021). *Noncommunicable diseases*. Available at:

[Noncommunicable diseases \(who.int\)](#). (Accessed: 2nd December 2021).

World Health Organization. (2017). *Report of the Commission on Ending Childhood Obesity:*

implementation plan: executive summary. Available at: [WHO-NMH-PND-ECHO-17.1-eng.pdf](#). (Accessed: 14th January 2021).

World Health Organization (2020). *WHO guidelines on physical activity and sedentary*

behaviour. Available at: [WHO guidelines on physical activity and sedentary behaviour](#). (Accessed: 10th November 2021).

World Bank Data. (2020). *Primary School Starting age*. Available at:

<https://data.worldbank.org/indicator/se.prm.ages?view=map>. (Accessed: 22nd May 2020).

World Bank Data. (2021). *World Bank Country and Lending Groups*. Available at: [World Bank](#)

[Country and Lending Groups – World Bank Data Help Desk](#). (Accessed: 22nd May 2020).

Zinkhan, M., Berger, K., Hense, S., Nagel, M., Obst, A., Koch, B., ... & Stang, A. (2014).

Agreement of different methods for assessing sleep characteristics: a comparison of

two actigraphs, wrist and hip placement, and self-report with polysomnography. *Sleep Medicine*, 15(9), 1107-1114.