

Durham E-Theses

*Effects of Time Pressure on Translation: A
Psycho-physiological, Behavioural-cognitive and
Performance-based Study*

YU WENG

How to cite:

WENG, YU (2021) Effects of Time Pressure on Translation: A Psycho-physiological, Behavioural-cognitive and Performance-based Study. 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/14359/> 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.



**Effects of Time Pressure on Translation:
A Psycho-physiological, Behavioural-cognitive and
Performance-based Study**

Yu Weng

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

School of Modern Languages and Cultures
University of Durham

September 2021

Declaration

The candidate confirms that the work submitted is her own and that it has not been submitted, in whole or in part, in any previous application for a degree. Appropriate credit has been given where reference has been made to the work of others. Part of the discussions in Chapter 4, Chapter 5 and Chapter 7 are developed from the author's preliminary work of this project which have been published in the form of refereed papers:

- Weng, Y., & B. Zheng. (2020). A multi-methodological approach to studying time-pressure in written translation: Manipulation and measurement. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 19, 218–236.
- Weng, Y., B. Zheng & Y. Dong. (in press). Time pressure in translation: Psychological and physiological measures. *Target*.

Copyright permissions have been granted for all the figures reproduced from others' work in this thesis.

The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.

Yu WENG

(Signature)



Table of Contents

DECLARATION	I
TABLE OF CONTENTS	II
ABSTRACT	VI
ACKNOWLEDGEMENTS	VII
LIST OF TABLES	VIII
LIST OF FIGURES	X
LIST OF ABBREVIATIONS	XII
CHAPTER 1. INTRODUCTION	1
1.1 RESEARCH BACKGROUND.....	2
1.2 METHODOLOGY	3
1.3 AIMS, RESEARCH QUESTIONS AND HYPOTHESES	3
1.4 STRUCTURE OF THE THESIS	7
CHAPTER 2. LITERATURE REVIEW	9
2.1 TIME PRESSURE IN TRANSLATION RESEARCH.....	10
2.1.1 <i>Time pressure and the translation product</i>	10
2.1.2 <i>Time pressure and the translation process</i>	14
2.1.2.1 Pausing behaviours	15
2.1.2.2 Translation speed.....	17
2.1.2.3 Time distribution	18
2.1.2.4 Attention distribution	19
2.1.2.5 Strategy and learning style.....	20
2.2 STRESS AND EMOTION IN TRANSLATION AND INTERPRETING RESEARCH.....	21
2.2.1 <i>Stress</i>	22
2.2.2 <i>Emotion</i>	26
2.3. SUMMARY	28
CHAPTER 3. THEORETICAL FRAMEWORK	30
3.1 TIME PRESSURE, STRESS AND HUMAN PERFORMANCE	31
3.1.1 <i>Time pressure in cognitive activities</i>	31
3.1.1.1 Time pressure, time pressure generation and the cognitive relevance	31
3.1.1.2 Mechanisms of time pressure functioning on cognitive processes	33
3.1.2 <i>Stress, coping and performance</i>	38
3.1.2.1 The nature of stress	39
3.1.2.2 The transactional theory of stress and coping	40
3.1.2.3 The relationship between stress and performance	43
3.2 COGNITIVE LOAD THEORY (CLT).....	46
3.2.1 <i>The theoretical basis of CLT</i>	47
3.2.2 <i>The construct of cognitive load in CLT</i>	48

3.2.3	<i>Three types of cognitive load in CLT</i>	49
3.2.4	<i>Expertise reversal effect</i>	50
3.2.5	<i>Cognitive or mental? Load or effort?</i>	52
3.3	AUTOMATICITY AND COGNITIVE RHYTHM	54
3.3.1	<i>Criteria of automaticity</i>	55
3.3.2	<i>Automaticity and the role of attention</i>	57
3.3.3	<i>Cognitive rhythm</i>	59
3.4	SUMMARY	61
CHAPTER 4. RESEARCH METHODOLOGY		64
4.1	MANIPULATION OF TIME PRESSURE	65
4.1.1	<i>Categorisation of time pressure inducement strategies</i>	66
4.1.2	<i>Objective strategies</i>	67
4.1.3	<i>Subjective strategies</i>	69
4.2	MEASURING TIME PRESSURE, EFFORT/COGNITIVE LOAD AND TRANSLATION BEHAVIOURS	70
4.2.1	<i>Subjective methods</i>	71
4.2.1.1	Measurement of time pressure using subjective methods	71
4.2.1.2	Measurement of effort using subjective methods	73
4.2.2	<i>Physiological methods</i>	75
4.2.2.1	Physiological measures used in this study	77
4.2.3	<i>Behavioural methods</i>	80
4.2.3.1	Keystroke logging metrics	81
4.2.3.2	Eye tracking metrics	86
4.2.3.3	Other behavioural metrics	89
4.2.4	<i>Analytical methods</i>	90
4.2.4.1	Application of CLT	91
4.3	TRANSLATION QUALITY ASSESSMENT	93
4.3.1	<i>Multidimensional Quality Metrics (MQM)</i>	93
4.3.2	<i>Functional-componential evaluation (FCE)</i>	96
4.3.3	<i>Inter-rater and internal reliability of the assessment</i>	98
4.4	SUMMARY	98
CHAPTER 5. EXPERIMENTAL SETUP AND DATA PROCESSING		101
5.1	PARTICIPANTS	102
5.2	SOURCE TEXTS	104
5.3	THE PRE-TEST AND TIME CONSTRAINTS	105
5.4	PROCEDURES AND APPARATUS	106
5.5	DATA ACQUISITION AND QUALITY ASSESSMENT	109
5.5.1	<i>Physiological data</i>	110
5.5.2	<i>Eye tracking data</i>	111
5.6	STATISTICAL ANALYSIS METHODS AND TOOLS	113
5.7	SUMMARY	114
CHAPTER 6. RESULTS		116
6.1	THE PSYCHO-PHYSIOLOGICAL EFFECTS	117
6.1.1	<i>The psychological effects</i>	118
6.1.2	<i>The physiological effects</i>	120
6.1.2.1	Comparison between baseline session(s) and task sessions	120
6.1.2.2	Comparison of the task sessions	123

6.1.3 Correlation tests	127
6.1.4 Summary of RQ1	128
6.2 THE BEHAVIOURAL-COGNITIVE EFFECTS	129
6.2.1 Automaticity	130
6.2.1.1 Processing speed	131
6.2.1.2 Parallel processing degree	133
6.2.1.3 Average TU effort	134
6.2.1.4 Attention allocation	136
6.2.1.5 Summary of RQ2	139
6.2.2 Cognitive rhythm	140
6.2.2.1 Pauses	141
6.2.2.2 Segments	146
6.2.2.3 Translation phases	147
6.2.2.4 Types of revision behaviour	150
6.2.2.5 Summary of RQ3	152
6.3 THE PERFORMANCE EFFECTS	153
6.3.1 Product quantity	153
6.3.2 Product quality	155
6.3.3 Summary of RQ4	156
6.4 THE ROLE OF COGNITIVE LOAD	157
6.4.1 Intrinsic and extraneous load	157
6.4.2 Overall cognitive load	158
6.4.3 Germane load	161
6.4.4 Behavioural metrics, cognitive load, and translation quality	165
6.4.6 Summary of RQ5	166
CHAPTER 7. DISCUSSION	167
7.1 PSYCHO-PHYSIOLOGICAL CHANGES ASSOCIATED WITH TIME PRESSURE	168
7.2 TRANSLATION-RELATED BEHAVIOURAL EFFECTS	171
7.2.1 Automaticity at the micro- and macro-level	171
7.2.2 Patterns of macro- and micro-strategy behaviours	176
7.2.3 Time distribution and revision type	179
7.3 THE STRESS EFFECT ON COPING EFFORT	182
7.4 COGNITIVE ACTIVATION EFFECT AND GERMANE LOAD	184
7.5 EXPERTISE REVERSAL EFFECT	186
7.6 TIME PRESSURE AND TRANSLATION PERFORMANCE	188
CHAPTER 8. CONCLUSION	192
8.1 SUMMARY OF THE MAJOR FINDINGS	194
8.2 IMPLICATIONS FOR TRANSLATION PRACTICE AND PEDAGOGY	198
8.3 STRENGTHS AND LIMITATIONS	200
8.4 AVENUES FOR FUTURE RESEARCH	201
BIBLIOGRAPHY	203
APPENDICES	224
APPENDIX 1: SOURCE TEXTS	224
APPENDIX 2: GLOSSARY LIST	226
APPENDIX 3: TRANSLATION GUIDELINES/BRIEF	227

APPENDIX 4: INSTRUCTIONS BEFORE TASK	228
APPENDIX 5: CONSENT FORM	229
APPENDIX 6: BACKGROUND INFORMATION QUESTIONNAIRE	230
APPENDIX 7: TASK ROTATION TABLE	231
APPENDIX 8: TASK INTERFACE	232
APPENDIX 9: EYE TRACKING DATA SCREENING TABLE	233
APPENDIX 10: GSR AND HRV DATA COMPLETENESS IN BASELINE 1 AND BASELINE 2	235
APPENDIX 11: ADAPTED NASA TLX QUESTIONNAIRE	237
APPENDIX 12: BIODATA OF PARTICIPANTS	238

Abstract

This empirical study provides a systematic investigation into how time pressure influences the translation process and product from psycho-physiological, behavioural-cognitive and performance-based perspectives. Taking a multi-methodological approach, this investigation is based on the analysis of data obtained from a total of 65 participants, who are grouped as novice and experienced translators, in an experiment containing three translation tasks with different time constraints. The project examines, firstly, the psychological and physiological consequences of time pressure indexed by self-reported Stress, Temporal Demand and State Anxiety, and several biomarkers recorded during and after the tasks. Reliable psycho-physiological indicators of time pressure in translation activities are identified by these analyses.

Secondly, combining the keystroke logging and eye tracking metrics, it explores how time pressure affects translators' observable behaviours from the aspects of automaticity and cognitive rhythm. Time pressure is found to have affected automaticity mainly at the micro-level while translators' experience makes a difference in affecting automaticity at the macro-level. Furthermore, the two groups of translators exhibit different processing patterns in terms of cognitive rhythm, especially in the moderately-time-pressured condition and the free-of-time-pressure condition. Thirdly, it looks into the direct effect of time pressure on translation quality, which generally features a negative relation despite slight traces of Yerkes-Dodson Law which are shown in the case of novice translators.

Lastly, based on Cognitive Load Theory, possible mechanisms that may underpin the identified relation between time pressure and translation quality are scrutinised, concerning the interplay of cognitive load (especially germane load), time pressure and translation quality. The result shows that time pressure could modulate the effect of germane load on novice translators' performance, with an essentially undermined effectiveness of germane load captured in the free-of-time-pressure condition. Such an effect is not visible in the case of experienced translators. These findings could provide pragmatic implications for translation practice and pedagogy.

Keywords: time pressure, psycho-physiological responses, automaticity, cognitive rhythm, cognitive load, translation quality

Acknowledgements

I would like to express sincere gratitude to my supervisors, Dr Bingham Zheng and Dr Sergey Tyulenev, for their intelligence, encouragement and invaluable guidance throughout my doctoral program. Many thanks to China Scholarship Council for funding my study and thanks to the faculty of the School of Modern Languages and Cultures at Durham University for providing all the support and help especially during the pandemic.

I would like to thank Prof. Yanping Dong, Prof. Arnt Lykke Jakobsen, Prof. Adolfo M. García, Prof. Ricardo Muñoz Martín, and Prof. Yanjing Wu for their insightful comments on my research. Thanks to all the reviewers in the three annual reviews who gave me a lot of constructive suggestions. I am also grateful to Prof. Michael Carl and Dr Moritz Schaeffer and all the staff and colleagues who participated the 2018 boot-camp held at University of Macau, where I learned a lot including R programming and eye tracking analysis skills. I also want to thank all the participants who took part in the experiment for their time and contribution.

Thank you to all my amazing friends and colleagues, Miss Ye Xia, Miss Yixiao Cui, Miss Bella Kuang, Mr Faustino Dardi, Miss Mingqing Xie and Miss Dariga Baktygereyeva, who have shared this experience along the way. I am deeply grateful for the immeasurable love and encouragement of my family and my loved ones who are always there standing behind me. Their unwavering support is much appreciated.

List of Tables

Table 4-1. Strategies of time pressure inducement	67
Table 4-2. NASA TLX item definitions and questions	74
Table 4-3. Summary of the physiological measures and corresponding indices	80
Table 4-4. Summary of the behavioural metrics.....	90
Table 4-5. MQM dimensions and error types applied in the present study	95
Table 4-6. Descriptions of components and categories in FCE provided by Colina (2009)	97
Table 4-7. Inter-rater and internal reliability of the translation quality assessment	98
Table 5-1. Statistics on the subjective evaluation of text difficulty level	105
Table 5-2. Datasets for physiological data analyses	111
Table 6-1. Type III ANOVA table of the LMER models of <i>State Anxiety</i> , <i>Stress</i> and <i>Temporal Demand</i>	118
Table 6-2. Type III ANOVA table of the LMER models of the physiological measures and indices	124
Table 6-3. Repeated measures correlation coefficients among the psycho-physiological measures that display a positive relationship with time stringency	128
Table 6-4. Summary of the metrics related to automaticity in translation.....	131
Table 6-5. Type III ANOVA table of the LMER models of <i>Keystroke production speed</i> , <i>Average IwTU</i> and <i>Average IbTU</i>	131
Table 6-6. Type III ANOVA table of the LMER models of <i>ST fixation duration % in TU</i> and <i>ST-TT fixation ratio in TU</i>	133
Table 6-7. Type III ANOVA table of the LMER models of <i>TU length (keystrokes)</i> , <i>TU time</i> , and <i>TU fixation</i>	135
Table 6-8. Type III ANOVA table of the LMER models of <i>Fixation duration % on ST/TT in drafting phase</i> with AOI and Condition as fixed effects	137
Table 6-9. Type III ANOVA table of the LMER model of <i>Fixation duration % on ST/TT in drafting phase</i> with AOI and Group as fixed effects.....	138
Table 6-10. Summary of the results for H4 and H5.....	140
Table 6-11. Summary of the metrics related to cognitive rhythm in translation	141
Table 6-12. Type III ANOVA table of the LMER models of SP, MP and LP count and length.....	142
Table 6-13. Type III ANOVA table of the LMER models of MP and LP count in the	

drafting and end revision phase	144
Table 6-14. Type III ANOVA table of LMER models of <i>Segment length (time)</i> and <i>Segment length (TU count)</i>	146
Table 6-15. Type III ANOVA table of the LMER models of Orientation, Drafting and End revision duration.....	148
Table 6-16. Type III ANOVA table of the LMER models of Orientation, Drafting and End revision percentage.....	149
Table 6-17. Type III ANOVA table of the LMER models of <i>Online deletions</i> and <i>End revision deletions</i>	151
Table 6-18. Type III ANOVA table of the LMER models of <i>TT word count</i> and <i>TU count</i>	154
Table 6-19. Type III ANOVA table of the LMER model of <i>Quality</i>	155
Table 6-20. Type III ANOVA table of the LMER model of <i>Mental Demand</i>	158
Table 6-21. Type III ANOVA table of the LMER models of <i>Effort</i> , <i>PWR</i> and <i>MFD</i>	159
Table 6-22. Repeated measures correlation coefficients among the effort indicators	160
Table 6-23. Repeated measures correlation coefficients between $Z_{\text{intrinsic+germane load}}$ and <i>Quality</i> with $Z_{\text{intrinsic+germane load}}$ being calculated by different combinations of overall cognitive load and extraneous load indicators	161
Table 6-24. Type III ANOVA table of the LMER models of <i>IGL_PWRTD</i> , <i>IGL_GSRTD</i> and <i>IGL_RMSSDTD</i>	162
Table 6-25. Type III ANOVA table of the LMER models of <i>Quality</i> with fixed effects of Condition, Group and $Z_{\text{Intrinsic+Germane}}$ represented by <i>IGL_PWRTD</i> , <i>IGL_GSRTD</i> and <i>IGL_RMSSDTD</i>	164
Table 6-26. Repeated measures correlation coefficients between the behavioural metrics and the overall cognitive load measures, $Z_{\text{intrinsic+germane load}}$ indicators and <i>Quality</i>	166

List of Figures

Figure 1-1. Scheme of the variables and research questions of the study	7
Figure 3-1. The control model of Variable State Activation Theory proposed by Maule and Hockey (1993), reproduced with permissions from the authors; copyright 1993 Springer.....	36
Figure 3-2. A flowchart of the transactional theory of stress and coping based on Lazarus and Folkman (1984) produced by Turner-Cobb and Hawken (2019). Permissions have been obtained to reproduce; copyright 2019 Cambridge University Press.....	42
Figure 3-3. The inverted U-shaped arousal-performance relation.....	44
Figure 3-4. Dimensions of the construct of cognitive load in CLT proposed by Paas and Van Merriënboer (1994), reproduced with permissions from the authors; copyright 1994 Springer.....	49
Figure 3-5. A model of attention and automaticity in reading proposed by Samuels (2013), reproduced with permissions from the author; copyright 2013 International Reading Association.	58
Figure 4-1. Rating scale of NASA TLX	75
Figure 4-2. Typing Chinese characters using the Windows 10 pinyin input method..	83
Figure 4-3. Thresholds of Short, Mid and Long Pauses	85
Figure 4-4. AOIs of <i>ST</i> , <i>TT</i> and <i>TM</i> drawn in the present study.....	87
Figure 4-5. Procedures of applying CLT as an analytical method	92
Figure 4-6. Dimensions and error types in MQM Core.....	94
Figure 4-7. Categorisation of data type in the present study	100
Figure 5-1. Scores of the objective measures of text difficulty	104
Figure 5-2. Experiment procedures.....	107
Figure 6-1. Effects of Condition and Group on <i>State Anxiety</i> , <i>Stress</i> and <i>Temporal Demand</i>	119
Figure 6-2. Box plots of the physiological measures in baseline and task sessions ..	121
Figure 6-3. Effects of Condition and Group on the physiological measures and indices	125
Figure 6-4. Effects of Condition and Group on <i>Keystroke production speed</i> , <i>Average IwTU</i> and <i>Average IbTU</i>	132
Figure 6-5. Effects of Condition and Group on <i>ST fixation duration % in TU</i> and <i>ST-</i>	

<i>TT fixation ratio in TU</i>	133
Figure 6-6. Effects of Condition and Group on <i>TU length (keystrokes)</i> , <i>TU time</i> , and <i>TU fixation</i>	135
Figure 6-7. Effects of Condition and AOI on <i>Fixation duration % on ST/TT in the drafting phase</i>	137
Figure 6-8. Effects of Group and AOI on <i>Fixation duration % on ST/TT in the drafting phase</i>	138
Figure 6-9. Effects of Condition and Group on <i>Fixation count on TM</i>	139
Figure 6-10. Effects of Condition and Group on SP, MP and LP count and length..	142
Figure 6-11. Effects of Condition and Group on MP and LP count in the drafting and end revision phase.....	144
Figure 6-12. Effects of Condition and Group on <i>Segment length (time)</i> and <i>Segment length (TU count)</i>	146
Figure 6-13. Effects of Condition and Group on Orientation, Drafting and End revision duration	148
Figure 6-14. Effects of Condition and Group on Orientation, Drafting and End revision percentage	149
Figure 6-15. Effects of Condition and Group on <i>Online deletions</i> and <i>End revision deletions</i>	151
Figure 6-16. Effects of Condition and Group on <i>TT word count</i> and <i>TU count</i>	154
Figure 6-17. Effects of Condition and Group on <i>Quality</i>	156
Figure 6-18. Effects of Condition and Group on <i>Effort</i> , <i>PWR</i> and <i>MFD</i>	159
Figure 6-19. Effects of Condition and Group on <i>IGL_PWRTD</i> , <i>IGL_GSRTD</i> and <i>IGL_RMSSDTD</i>	163
Figure 6-20. Effects of Condition, Group and <i>IGL_PWRTD</i> on <i>Quality</i>	164
Figure 7-1. Illustration of Yerkes-Dodson Law (left) in relation to the (flip-over presentation of) the stress-performance relation identified in the present study (right)	189

List of Abbreviations

μS	microsiemens
AIIC	International Association of Conference Interpreters
ANOVA	Analysis of Variance
ANS	Autonomic Nervous System
AOI	Area of Interest
BP	Blood Pressure
bpm	Beats per minute
CLT	Cognitive Load Theory
DBP	Diastolic Blood Pressure
DenDF	Denominator degrees of freedom
FCE	Functional-Componential Evaluation
GSF	Gaze Sample to Fixation percentage
GSP	Gaze Sample Percentage
GSR	Galvanic Skin Response
GTS	Gaze Time on Screen
H	Hypothesis
HR	Heart Rate
HRV	Heart Rate Variability
Hz	Hertz
IbTU	Interval between Typing Units
IGL	Sum of Intrinsic and Germane Load
IwTU	Interval within Typing Units
LMER	Linear Mixed-Effects Regression
LP	Long Pause
MAP	Mean Arterial Pressure
Mean Sq.	Mean square
MFD	Mean Fixation Duration
mmHg	Millimetre of mercury
MP	Mid Pause
MQM	Multidimensional Quality Metrics
NASA TLX	NASA Task Load Index
NumDF	Numerator degrees of freedom
PNN50	Percentage of successive normal-to-normal heartbeat intervals that differ by more than 50 milliseconds

PWR	Pause-Word-Ratio
PSNS	Parasympathetic Nervous System
RMSSD	Root Mean Square of Successive Differences between the normal-to-normal heartbeat intervals
RQ	Research Question
SBP	Systolic Blood Pressure
SCR	Skin Conductance Response
SD	Standard Deviation
SDNN	Standard Deviation of the normal-to-normal heartbeat intervals
SE	Standard Error
SG	Segment
Sig.	(Statistical) Significance
SNS	Sympathetic Nervous System
SP	Short Pause
ST	Source Text
STAI	State-Trait Anxiety Inventory
Sum Sq.	Sum of squares
TD	Temporal Demand
TM	Timing area (AOI)
TT	Target Text
TU	Typing Unit
VSAT	Variable State Activation Theory

Chapter 1. Introduction

1.1 Research background

In recent decades research interests in Translation Studies have substantially shifted from purely analysing textual elements to a more process-oriented approach. Especially, increasing importance has been attached to the cognitive aspects of human translators – the “black box” in their mind. The impetus of exploration of cognitive processes inevitably gave rise to the study of translators. As Bell (1991, p. 43) suggests, “description of the process and/or the translator [...] form the twin issues which translation theory must address: how the process takes place and what knowledge and skills the translator must possess in order to carry it out”. Additionally, Chesterman (2009) offers an outlook for researching the cultural, cognitive and sociological aspects of translators with mental processes, decision-making, emotions and attitudes included in the cognitive dimension. To describe the complex processes underpinning translators’ cognition in translation activities and to provide empirical evidence for theory formulation, observable data needs to be gathered. Nevertheless, while translation process research and “translator studies” have attracted much attention, specific and comprehensive investigations into translators’ affective interface with their work, especially the role time pressure plays, are still scarce in recent years.

Time pressure in the professional work environment has become a predominant social-situational challenge for many occupations including translation. Particularly, given the pragmatic nature of translation work, translators almost always work with solutions which promise a maximum of effect with a minimum of effort (Levý, 1967, p. 1179). As a result, the ability of working to strict deadlines has become one of the critical elements of a professional translator’s competence. Taking a broader view, effective emotion regulation and stress management can be seen as a crucial aspect of the psycho-physiological components in the translation competence model developed by PACTE (2003). Thus, inquiring into the effects of time pressure, as a salient source of stress, on translation and translators across multiple dimensions is essential for improving translators’ performance in stress-provoking settings and providing pedagogical implications for translation training instructors.

This study is an empirical exploration of how time pressure affects the translation process, the translation product and the translators from psycho-physiological, behavioural-cognitive and performance-based perspectives. The investigation bases its analyses on a variety of quantitative data ranging from psycho-physiological measures

to behavioural metrics and translation quality assessments collected from two groups of participants, i.e., novice and experienced translators. Incorporating theoretical constructs from cognitive psychology and learning sciences such as models of stress and Cognitive Load Theory, this study sets out to provide new insights into the impact of time pressure on translation and thus shed light on translation practice, training and research.

1.2 Methodology

The investigation was carried out adopting a within-subject design, and all the data were collected from three translation tasks with comparable English source texts (ST) rendered into Chinese target texts (TT) under different time constraints (i.e., *Short*, *Standard* and *Free* condition). Apart from the objective strategy of imposing deadlines on the tasks, subjective strategies were used as well in terms of time pressure manipulation and elicitation. The investigation followed a multi-methodological approach complying with the principle of triangulation, i.e., applying several instruments of data gathering and analysis to strengthen the reliability and internal validity of an investigation (Alves, 2003; Creswell, 2009). Specifically, data collection combined eye tracking, keystroke logging, subjective scales (i.e., NASA Task Load Index and State-Trait Anxiety Inventory) and the measurement of several physiological markers (i.e., heart rate, blood pressure, galvanic skin response, heart rate variability, pupil size and skin temperature). The strengths of different types of methods converged to form a more robust methodological approach to the investigation.

The statistical analysis of data derived from these sources was mainly based on the inferential linear mixed-effects regression modelling, a rigorous statistical technique that accounts for both fixed effects and random effects (e.g., the effects relating just to the participants sampled in the study), rendering the findings more generalisable. Results of the analyses in this respect thus served to make the inquiry more comprehensive.

1.3 Aims, research questions and hypotheses

In light of the research background and the significance of investigating the effects of time pressure in translation from different dimensions, there are three specific aims that help to frame the present study:

Aim 1: to determine measures of time pressure/stress during translation with translators' psychological and physiological responses being tested as possible indicators, and to recognise possible differences between novice and experienced translators in these responses.

Aim 2: to examine the effects of time pressure on translators' behavioural patterns about automaticity and cognitive rhythm based on their eye-key activities and identify specific features underpinning the experienced translators' translation behaviours.

Aim 3: to explore the effects of time pressure on effort/cognitive load and reveal possible underlying mechanisms accounting for the relationship between time pressure and translation quality by looking into the functioning of cognitive load, especially germane load based on Cognitive Load Theory.

Five research questions (RQ) and fourteen corresponding hypotheses (H) are formulated to pursue these aims. Firstly, the psycho-physiological effects of time pressure/stress have rarely been touched upon in written translation research so far. Taking a stringent time constraint as a stressor, the present study explored this issue with several psychological measures and biomarkers. The first research question was thus drawn up as follows:

RQ1: What are the effects of time condition and experience on participants' psycho-physiological responses during translation?

H1: The psycho-physiological responses would show a significant effect of time condition and would vary in accordance with the stringency of time condition for both groups of participants.

H2: With the time condition getting more stringent, novice translators would have stronger psycho-physiological responses (with higher increments in the tested measures) than experienced translators.

H3: Those measures that corroborate H1 would be correlated with each other.

Secondly, translation-specific behavioural patterns were investigated from the perspectives of automaticity and cognitive rhythm with a range of elaborate behavioural metrics incorporating eye tracking and keystroke logging activities. Two research questions were formulated on this facet of the investigation:

RQ2: What are the effects of time condition and experience on participants' level of automaticity during translation as indicated by the behavioural metrics?

H4. A more stringent time condition gives rise to more automated processing, evidenced by faster processing speed, deeper parallel processing, less effort in each processing unit, as well as an increased proportion of attention on TT production and a decreased proportion of attention on ST comprehension.

H5. Experienced translators have a higher automaticity level than novice translators, evidenced by faster processing speed, deeper parallel processing, less effort in each processing unit, as well as a higher proportion of attention on TT production than on ST comprehension.

RQ3: What are the effects of time condition and experience on participants' cognitive rhythm during translation as indicated by the behavioural metrics?

H6. A more stringent time condition would shorten and reduce the number of all types of pauses.

H7. Time condition would affect the segment length differently for novice and experienced translators.

H8. When the time condition becomes more stringent, the orientation and end revision phases are more likely to be trimmed than the drafting phase.

H9. Time condition would affect the pattern of preference for online/end revision differently for novice and experienced translators.

Thirdly, translation quality was examined as a function of time pressure and translators' experience levels. Inspecting the relationship between time pressure and

translation product quality might be of pragmatic significance for translation practice and training. The research question was drawn up as follows:

RQ4: What is the relationship between time condition and translation quality for the two groups of participants?

H10. The stringency of time condition would negatively affect translation quality for both groups.

H11. Experienced translators can achieve better translation quality than novice translators regardless of the time conditions.

Lastly, based on Cognitive Load Theory, explorations were carried out to unveil the underpinning mechanisms of the relationship between time pressure and translation quality. The effect of time pressure on different types of cognitive load (e.g., overall cognitive load and germane load) was inspected, and more importantly, the role time pressure played in modulating the effect of cognitive load, especially germane load on translation quality, was looked into with the following research question:

RQ5: Within the framework of Cognitive Load Theory, how does time condition affect participants' overall cognitive load, germane load and the way in which germane load influences translation performance?

H12. The stringency of time condition would proportionately exacerbate the overall cognitive load.

H13. The effect of time condition on germane load would be similar to its effect on translation quality for the two groups of participants.

H14. The stringency of time condition could modulate the way germane load influences translation quality.

Figure 1-1¹ illustrates the scheme mapping the research questions to the variables that are investigated.

¹ Note that the variable “cognitive load” here is an omnibus construct containing some separate concepts such as “overall cognitive load” and “germane load” within the framework of Cognitive Load Theory (see more details in section 3.2).

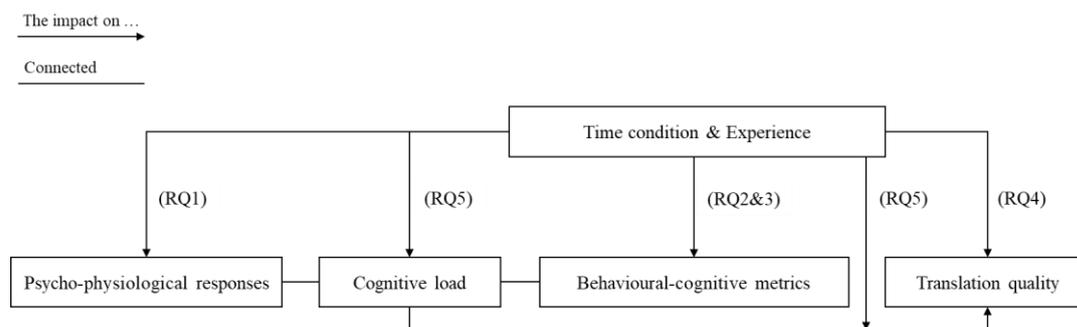


Figure 1-1. Scheme of the variables and research questions of the study

1.4 Structure of the thesis

The remainder of this thesis is structured as follows:

Chapter 2 provides a thorough review of previous literature on time pressure, stress and emotion in written translation and interpreting research. Specifically, the effects of time pressure on the translation product and process are examined in section 2.1; since stress and emotion are also highly relevant to the current investigation, an expansion of literature focusing on these topics is discussed in section 2.2.

Chapter 3 presents the theoretical framework of the present investigation, including detailed discussions of relevant theoretical concepts and models relating to time pressure, stress and their effects on human performance (section 3.1), the constructs in Cognitive Load Theory (section 3.2), and theories related to automaticity and cognitive rhythm (section 3.3), which, altogether, serve as the theoretical foundation of this study.

Chapter 4 describes the overall methodological approach adopted by the investigation, including the strategies for time pressure manipulation (section 4.1), the subjective, physiological, behavioural and analytical methods for measuring time pressure and effort/cognitive load (section 4.2), as well as the translation quality assessment instruments (section 4.3).

Chapter 5 gives details about the design, procedures and set-ups of the experiment (section 5.1 – 5.4), as well as the steps of data processing (section 5.5 – 5.6) required for the analyses.

Chapter 6 reports all the results of the analyses in response to the five research questions and the corresponding hypotheses: the psycho-physiological effects are reported in section 6.1, which corresponds to RQ1; the results about behavioural

patterns from the perspectives of automaticity and cognitive rhythm are reported in section 6.2 in response to RQ2 and RQ3; the relationship between time pressure and translation performance is examined in section 6.3 which answers RQ4; the results for RQ5, which is about the role of cognitive load plays under the influence of time pressure, is reported in section 6.4.

Chapter 7 discusses the results obtained in the analyses, connecting them with the empirical findings from previous studies and giving comprehensive and in-depth interpretations based on the theoretical constructs and models. The psychophysiological (section 7.1), behavioural-cognitive (section 7.2 – 7.5) and performance (section 7.6) impacts of time pressure and professional experience are inspected in detail with the research aims revisited.

Chapter 8 concludes this thesis by providing a summary of all the findings generated by the investigation (section 8.1), giving implications for translation practice and pedagogy (section 8.2), discussing the strengths and limitations of the study (section 8.3) and offering directions for future research (section 8.4).

Chapter 2. Literature Review

Since time pressure has been a remarkable social-situational factor in the workplace of translation practice, the influence of time pressure on translation has been noticed and researched from different dimensions. One of the most frequently visited topics is the relationship between time pressure and the final translation product quality, i.e., whether time pressure undermines or promotes translation performance. Substantial attention has also been attached to the effects of time pressure on the translation process, such as how time pressure affects translators' choice of translation strategies and how time pressure alters translators' allocation of attentional resources. Section 2.1 reviews relevant research from the perspective of time pressure's impact on the translation product and process, respectively. In addition, investigations on stress and emotion in translation and interpreting research are highly relevant to time pressure and could shed some light on the current research. Section 2.2 thus extends the examination to these areas. Section 2.3 gives a summary of the findings and methodologies of previous studies.

2.1 Time pressure in translation research

This section mainly focuses on previous literature in translation process research that investigated time pressure as a major variable in their experiments. Most of these studies employed different time constraints to elicit different levels of time pressure. The effects of time pressure on the translation product are examined in section 2.1.1, and the effects on the translation process are reviewed in section 2.1.2.

2.1.1 Time pressure and the translation product

The effects of stress or time pressure can be destructive or constructive to performance (see the theoretical discussions in section 3.1.2). However, empirical evidence from previous studies mainly embodied the adverse effects of time pressure on translation product quality. In one of Hansen's (Hansen & Hönig, 2000; Hansen, 2002a, 2002b, 2005, 2006a, 2006b) series of experiments, participants were asked to translate the exact text twice, one with a time constraint (the "short time translation" session) followed by the other without a time constraint (the "long time translation" session). The results showed that the quality of the translation product was evidently improved for most participants in the "long time translation" session when they were working without a time constraint. This result was not surprising, especially considering the fact

that the text was translated twice, and it was the second session that did not have a time constraint. However, one of the research objectives was to pinpoint students' first reaction, i.e., the spontaneous rendition offered immediately after reading through the source text. Therefore, what should be noticed in Hansen's research is that in many cases participants still produced the same number of, or even more errors, especially in the revision phase of the second task session when they had unlimited time at their disposal (Hansen, 2006b, p. 77). This finding indicated that not all the participants who performed well under time pressure achieved the same or better performance when working without a time constraint; sometimes, the free-of-deadline condition made them overthink and eventually overwrite the original correct renditions with erroneous ones.

Hansen (2006b) suggested that participants could have relied on their automatic skills, but unfortunately, the intuitive renditions were rejected when they had excessive time. The pedagogical implication is that critical awareness of behaviours during the translation process is vital for students. They should realise that they have the (potential) ability to perform well spontaneously, or they could consciously train their automatic translation skills to avoid destructive behaviours and habits. As a result, this could essentially improve the efficiency and effectiveness of translation training. Identifying such phenomena by comparing behaviours in time-pressured and free-of-deadline conditions during the translation process can undoubtedly be meaningful for translation training, which was the primary purpose of Hansen's research. These findings constitute a competitive and empirically supported explanation of why the lack of a time constraint may affect translation quality negatively for some translators. Nevertheless, since the experiment was designed to be a repetitive task for the participants, only limited inferences can be gained regarding the relationship between time pressure and the quality of the translation product.

Probes into the effects of time constraints on translation quality had also been the focus of De Rooze's (2003, 2008) study. In one of his experiments, each participant translated two texts (around 120 words) with a 10-minute and 15-minute time constraint respectively. It was surprisingly found that 19% of the student participants produced a better translation in the shorter session than in the longer session, and over 25% maintained the same level of quality under the two conditions (De Rooze, 2003, p. 88). This meant that a more stringent time constraint undermined the translation product quality for over half (around 55%) of the participants; but for the rest of them, which

was also a relatively large proportion, a more relaxed deadline did not guarantee a better translation product. In the later experiment with less time (8 minutes and 12 minutes) and longer source texts (around 240 words), a significant decrease of the translation quality score was observed in the student group. Again, the majority (around 79%) of the student participants performed worse under more intense time pressure while the rest (around 21%) of them showed no difference or the opposite trend. Similarly, the translation quality of 60% of the professional translators decreased under the more time-pressured condition while there was no change or an increase in the quality score for the rest (40%) of them. But, the overall effect of time condition on the professional group was not significant. These results clearly indicated that time pressure could lead to different consequences for different people in translation activities; for some participants, time pressure impaired their performance but for some others, it had no effect or spurred them to perform better. To some extent, this also echoed the findings from Hansen's (2002b) study in that there existed cases that time pressure could facilitate translation quality for a portion of translators.

Further, the results from De Rooze's (2003) study showed that translation quality dropped by more than 15% when the participants' translation speed exceeded 200 words per ten minutes. Taking acceleration as a result of intense time pressure, this finding also shed some light on the relevance of translation quality to translation speed, i.e., the quantity of translated material produced within a certain unit of time. There might exist a critical point of stress or time pressure beyond which their performance will drastically deteriorate.

Ghobadi et al. (2017) especially focused on whether time pressure could influence translation quantity and quality. Two groups of student participants were tested in their study. The experimental group (n=23) was given 30 minutes for the translation task (English to Persian) and the control group (n=21) had no time limit for the task. Their products produced in the first 15 minutes of the translation process were analysed. A trade-off effect was observed in that participants who executed the task with a time limit produced more translated materials but of inferior quality compared with the participants who completed the task without a time limit. This result was in line with De Rooze's (2003) finding to the extent that there might be a negative relationship between translation quality and translation speed when it is higher than a certain level (e.g., 200 words per ten minutes in De Rooze's research). Ghobadi et al. (2017) explained that the translator's mental resources had been unequally distributed to

different aspects (i.e., quality vs. quantity) of the translation task, which was influenced by the time conditions. Ghobadi et al. (2017) also indicated that there was no evidence of anxiety experienced by the participants under time pressure which may push them to perform the task with a high speed and negatively affect the product quality. Nevertheless, the underlying mechanisms of such a relationship between translation quality and time pressure (or translation speed) captured by these studies are worth investigating with more in-depth analyses.

In the same vein, the effects of time constraints on translation product quality were also thoroughly examined in Kourouni's (2012) study in which three tasks with three timeframes were designed in the experiment, namely 15 minutes, 20 minutes and 30 minutes. It was found that although reduced time lowered the overall product quality, the difference failed to reach a level of statistical significance. By further checking the effects on specific aspects of the product, it was found that more time led to particularly higher scores on Formal Correctness and Terminology/Vocabulary. Thus, it can be concluded that the change in timeframe may not directly correlate with the change in the overall translation quality but can negatively affect certain aspects of the translation product that seem to be less time-resistant. Therefore, a specific focus on these aspects in translation training was suggested by Kourouni (2012).

In addition to the empirical-experimental approach, Jiménez-Crespo (2012) adopted a product-based corpus methodology to examine the effects of time pressure on the quality of online published translations. A parallel corpus was compiled in the 12 hours after the former US President Obama's inaugural speech, which included ten different translations into Spanish collected from online media outlets. The translations were inspected with regard to error counts, lexical variety and creativity. It showed that although a clear association between errors and lexical variety could be captured, the effects of time pressure might vary among translations when all the factors under study were concerned. This meant that while those translations with the highest number of errors were also of lowest lexical variety and creativity, more translations showed a great range of variation, or inconsistency, in these aspects. For example, some translations were of high creativity but low lexical variety, and some were of high error counts but also high lexical variety. It was indicated that different translators coped with errors and creativity in different ways and when they were working under time pressure, they might not be able to attend to every facet of the product, but were likely to concentrate only on certain aspects (e.g., accuracy or creativity) depending on the

purpose and the context of the task.

Overall, evidence from previous studies mainly reflects the destructive effects of time pressure on translation quality, although not all these effects are significant. A few studies also show traces of the constructive effect of time pressure on translation quality, but only for a minor portion of the sampling group. Automatic skills and the unequal distribution of mental resources to different aspects of the translation task (e.g., translation quality vs. quantity) in different time conditions could possibly account for such phenomena. Meanwhile, a systematic investigation of the complex working mechanism which underlies the relationship between time pressure and translation quality for translators with different experience levels is still lacking.

2.1.2 Time pressure and the translation process

Probing into the effects of time pressure on the translation process is mainly associated with studying the translators' behaviours from which the cognitive processes can be inferred. In several pedagogy-oriented studies (e.g., Hansen & Höning, 2000; Hansen, 2002b, 2006b; Kourouni, 2012) conducted mainly with student translators, the effects of time pressure on the translation process have been researched with the purpose of facilitating and improving the effectiveness of translation training. For example, with the motivation of discovering the best methods that could be applied to translation classes to deal with the stagnation in the translation performance of the students, an effort was made in Hansen's (2006b) research to catch the knowledge from the process data which cannot be derived from the product alone. Her experiments were thus designed not to represent the "normal" translation process at the workplace, but to capture as much as possible the students' cognitive processes of translation reflecting their "competences, actual behaviour and habits" (Hansen, 2006b, p. 78).

Some research methods, such as think aloud protocols, retrospection, keystroke logging and eye tracking are frequently used in such inquiries. Different dimensions of the translation process have been researched with different methods. For example, topics such as patterns of segmentation and the frequency of problem-solving activities can be researched by applying the method of keystroke logging, while the distribution of visual and attentional resources can be studied by using the eye tracking technique. A systematic review of the findings in previous studies regarding the effects of time pressure on pausing behaviours, translation speed, time distribution to different

translation phases, (visual) attention distribution and the use of strategies during the translation process is presented in the following sections.

2.1.2.1 Pausing behaviours

Pausing behaviours during translation might signal different cognitive processes. For example, a distinction was made between the *orientation pauses*² and the *internal pauses*³ in Hansen and Hönig's (2000) study. The ratio of the total pause duration to the orientation pause duration in a translation task was calculated. Taking the *orientation pauses* as a sign of macro-strategy behaviours, it was found that thinking ahead (i.e., taking a longer orientation pause and rendering the product more spontaneously afterwards) did not necessarily lead to fewer mistakes for students from the two (German and Danish) institutions involved in that study. In fact, in the condition without a time limit, students from the German institution relied much more on their spontaneous ideas and their automatic translation skills than students from the Danish institution. In addition, students from the German institution who performed well in the automated processes (under time pressure) also predominantly succeeded when they were working without time pressure. Thus, the hypothesis that *those who perform well under time pressure will perform well without time pressure* was likely to be corroborated only for translators who showed the ability to render with a high automaticity. This ability requires the use of macro-strategy more frequently in the process of translation. Thus, the acquisition and development of such an ability is critical for translation trainees.

In another experiment, Hansen (2002b) defined three types of pauses, namely orientation pauses (at least 6 seconds immediately before the first word of a sentence), control pauses (at least 6 seconds immediately after the last word or the full stop of a sentence) and monitoring pauses (a type of control pause during which the translator checks not only the precedent sentence but also any part of the target text that have already been produced). It was found that the position, number and length of the pauses were not correlated with the product quality regardless of time condition and translation

² Pauses of at least 12 seconds, which are inserted after entering a new passage of text or before formulating sentences or sections. These pauses are used to get an overview of the sentence/section to be translated.

³ Pauses of at least 12 seconds, which are inserted within the formulation of an already-started sentence. These pauses are supposed to occur when the translator has difficulty in completing the formulation of the sentence.

direction. In addition, translators' behaviour patterns indicated a relatively fixed work rhythm with regard to the amount of time distributed to different translation phases and pauses. Such behaviour patterns certainly reflected the automaticity level and the routine or habits of the translators. However, there was no correlation between such patterns and the quality of the products as well. These findings implied a multi-faceted, rather than a one-fold relation between translation quality and the behaviours during the translation process for different translators.

While Hansen (2002b) identified and applied different types of pauses as indicators of different cognitive processes during translation, Jensen's (1999, 2000; Jensen & Jakobsen, 2000) studies took pauses with a fixed threshold of four seconds (recorded by Translog) as signals of problem-solving activities. Jensen's (2000) experiment adopted three fixed deadlines (15, 20 and 30 minutes) for translating three 120-word English texts into Danish by two non-translators, two students and two professional translators. It was found that when time decreased, problem-solving activities reduced significantly, but only during the revision phase, which meant that the revision phase was more flexible and therefore more likely to be compressed when time was constrained. In other words, "raw TT production has its own rhythm and speed, which corresponds to the time factor but is fairly independent of it" (Jensen & Jakobsen, 2000, p. 112).

The impact of time pressure on revisional behaviours was also examined in Alves and Liparini Campos's (2009) research, which focused on pause patterns and the types of resources accessed by professional translators during translation. The results showed that time pressure reduced the number of *revision pauses* in both drafting and end revision phases and indicated that time pressure affected the revision processes mostly, which was independent of the use of a translation memory system. Despite the effect of time pressure, De Rooze (2003) found that spelling mistakes tended to appear immediately after a pause. This can be seen as a possible indication that reallocation of cognitive resources was not automatic during the process and preoccupation with the latest problem took some time to fade away (Muñoz Martin, 2009, p. 31). Also, those spelling mistakes require additional revision behaviours, and time pressure could possibly influence the number of such revision behaviours as shown in other studies (e.g., Alves & Liparini Campos, 2009); such an influence may lead to a degraded translation product as a consequence.

In summary, some effects of time pressure on different cognitive processes

indicated by certain types of pauses are identified. For instance, there might be a relationship between the use of macro-strategy behaviours under time pressure and the final product quality (Hansen & Hönig, 2000). Also, a more plausible effect of time pressure is its unbalanced consequences on different components of the translation process. For example, the revision behaviours appear to be less resistant to the influence of time pressure (e.g., Jensen & Jakobsen, 2000; Alves & Liparini Campos, 2009).

2.1.2.2 Translation speed

Some of the findings mentioned above related to pausing behaviours essentially converge to a point that time pressure plays a role in activating automaticity, a core component of the competence of a translator. Apart from pauses, this ability is also reflected in the processing speed of translation.

Jakobsen (2002, 2003) conducted a series of experiments mainly investigating the effects of think-aloud protocols on translation speed, revision and segmentation of semi-professionals and professionals with the methods of keystroke logging and think-aloud protocols. The results showed that think-aloud protocols indeed affected translation speed to a great extent for both groups by comparing data with and without thinking aloud in the experiments. Thus, it was suggested that when dealing with time-related variables, such as time pressure or translation speed, the use of think-aloud protocols as a method should be treated with special care because it is likely to affect the results. More importantly, Jakobsen's research also showed that professionals tended to draft faster than semi-professionals, but they also spent more time on revision (Jakobsen, 2002, pp. 191-192). This corroborates the proposition that professional or expert translators are equipped with such an ability, which resembles the interpreters, to render with a higher level of automaticity resulting in a faster speed in the drafting process.

De Rooze (2003) found that the difference in time constraints did not affect the number of pauses and the number of keystrokes per minute for the student participants, indicating that the participants did not accelerate their work rate under time pressure. De Rooze thus emphasised the importance of speed when translating under time pressure because it indicates a higher level of automaticity and the ability of faster accessing the problem-solving resources. The traces of the advantageous effect of time pressure on translation quality for a portion of translators confirmed the conducive

impact of translation speed. This meant that a greater number of operations were performed per unit of time without overloading the information processing capacity composed of the long-term and short-term memory (*ibid.*). If the processing system was overloaded, the translator would be in danger of a collapse of their processing system, which may be the reason for most of the errors made in the task completed under time pressure. Therefore, specific focus in translation training should be put on correlating translation speed with the quality of information processing.

2.1.2.3 Time distribution

The above findings provide evidence that restricting the time resource available to the translators can impose an intrinsic influence on the way translators assign the resources during the process. Modulating the translation speed can be seen as one aspect of adjusting the overall pace of work or translation rhythm, which is relevant to the resource allocation scheme. The translation process is also widely researched from the perspective of the stages from the beginning to the end. These stages typically include the orientation, drafting and revision phases (Jakobsen, 2002). Modulating the time distribution to different phases of the translation process is another aspect of adjusting the translation rhythm.

As mentioned above, Jakobsen (2002) found that professionals drafted faster and spent more time on the revision phase than semi-professionals. Hansen (2002b) found a trade-off effect that if a large amount of time was consumed in one phase, it would be compensated by less time spent on other phases when time pressure was absent. Hansen's study (2002b) also showed that each individual translator might have developed a relatively fixed pattern (or habit) of time consumption in different translation phases, but no evidence was found that it is correlated with translation quality.

The results from Kourouni's (2012) study suggested that when available time decreased, the student participants tended to focus more directly on planning and translation work and to drop any time-consuming preparatory moves (e.g., preparing their workbench in various ways). Also, the decision to read the source text in advance, based on the participants' self-reported data, was also affected by the stringency of the time constraint. However, this decision did not necessarily result in an impact on the product quality. It seems that those participants who did not read their source text

consciously before translating compensated for it by other means, perhaps by focusing on the source text more as they went along. They also compromised on revision in a similar way: there were almost 60% of the participants who revised their product in the most relaxed time condition (30 minutes); when the available time was halved, so were the number of the participants (approximately 30%) who had a revision phase. On the other hand, participants resisted reducing the translation time. This result was in line with Jensen's (2000) conclusions. The minimum percentage for the translation (drafting) phase slightly increased when the time available decreased, while there were cases where participants devoted almost all their time to the translation phase. Revision did not constitute a separate phase for many participants and Kourouni (2012) suggested that such a behavioural pattern implied a blurring of phases, which Pym (2009) referred to as "blending".

The above findings show that the end revision phase appears to be the most flexible phase that is likely to be "trimmed" (possibly passively) under intense time pressure. This could result from the fact that the revision phase is the final stage of the translation process without much prospective control. But whether the revision time is associated with product quality is of relevance when such a compression of the revision phase under time pressure is so common. Künzli (2007) investigated the relationship between the final product quality and the time spent on an independent revision task of a legal text by ten professional translators. It was found that quality takes time because the two participants who spent the longest time on the revision task produced the best translations. However, spending a lot of time on revision did not necessarily always amount to a high-quality product.

2.1.2.4 Attention distribution

Sharmin et al. (2008) carried out a study concentrating on time pressure and text complexity in relation to visual attention distribution. Using the eye tracking technology, 18 student participants were involved in tasks of translating three texts of different complexity under the 6-, 5- and 4-minute timeframes. They found that time pressure mainly affected the fixation duration on the source text area compared with the target text area. This indicated that it was easier for translators to adapt their reading-for-comprehension to different time constraints, whereas it was more difficult for them to adapt their reading-and-monitoring of the target text. This is the first study on time

pressure which adopted the eye tracking technology and focused on how visual attention was affected by time pressure.

Hvelplund (2011) explored how time constraints can influence the distribution of attentional resources with the method of eye tracking as well. Conforming with Sharmin et al. (2008), it showed that under time pressure, less cognitive resources were allocated to ST processing than under no time pressure. A probable explanation was that the ST meaning was identified more quickly as less time was available. TT reformulation was affected little by time pressure. This meant that when a translator was working under time pressure, the amount of attention allocated to the ST comprehension, rather than the TT reformulation, was more likely to be affected or compressed. Also, indicated by pupil size, cognitive load was higher during translation under time pressure than under no time pressure. This comes as no surprise, since less time was available under time pressure to carry out the same cognitive operations involved in ST comprehension and TT reformulation.

The findings in the above-mentioned studies mainly show that time pressure is more likely to cast an impact on the attentional resources spent on reading comprehension, rather than on the production process. Despite these few studies, the examination of effects of time pressure on the allocation of visual attention during translation is still scarce.

2.1.2.5 Strategy and learning style

Apart from the distribution of problem-solving activities, Jensen and Jakobsen (2000) also investigated the effects of time pressure on the use of translation strategies with the methods of think-aloud protocols and keystroke logging. No significant correlation was found between time constraint and the use of direct (borrowing and literal translation) and indirect (paraphrase and adaptation) strategies. Jensen (2000) applied Scardamalia and Bereiter's (1987) models of Knowledge Telling and Knowledge Transforming into translation and the result showed that expert translators were able to apply Knowledge Telling strategies in a more consistent way than the other two groups, and when problems occurred, non-professionals applied Knowledge Telling strategies to solve them, whereas professional translators applied Knowledge Transforming strategies in response to problems (Jensen, 2000, pp. 180-181). These findings provide further insights into the differences in cognitive processing styles between groups of

translators with different levels of experience under time pressure. However, as mentioned earlier, the potential slowing-down effect of the concurrent think-aloud protocols may influence the accuracy of the results.

Incorporating the product-based analysis, management theory and educational psychology, Kourouni (2012) investigated student translators' learning styles (i.e., patterns of thinking, perceiving, problem-solving and remembering when approaching a learning task) in her study of time pressure and translation. Kourouni (2012) defined four procedures, i.e., omission, literalism, specification and generalisation, which are a set of steps leading to a specific result in order to solve a particular translation problem. It was found that the visual/verbal dimension of the learning styles was correlated with two cases of the procedures and a consistent pattern was shown: the stronger the preference for visual learning, the more the participants stuck to the text (i.e., likely to choose literalism). Conversely, the stronger the preference for verbal learning, the more the generalisation was used.

In addition, the results in Kourouni's (2012) study suggested that with a relatively convenient deadline, the visual/verbal dimension of learning styles (about the input modality) was significantly correlated with translation quality, showing that the more extreme the preference for visual learning, the worse the overall quality. With a tighter deadline, the sequential/global dimension (about the understanding of the input) was dominantly correlated with the translation quality, showing that the more sequential the participants, the higher the overall quality. Such empirical findings have shed some didactic light on translator training from a more learning-centred perspective.

Overall, the effect of time pressure on translators' choice of (direct or indirect) translation strategy was not evident (Jensen & Jakobsen, 2000). However, translators' experience level may modulate their choice of Knowledge Telling and Knowledge Transforming approaches to translation under different time conditions (Jensen, 2000). In addition, there might be a correlation between some dimensions of the student translators' learning style and their translation quality (Kourouni, 2012).

2.2 Stress and emotion in translation and interpreting research

Despite the above-mentioned works that focused specifically on time pressure in written translation, extensive attempts that sought insights into how stress and emotion could function to influence both translation and interpreting performance have been

made in recent decades. The findings and methodologies in those studies are undoubtedly relevant to the current exploration of the impact of time pressure on translators and translation performance.

2.2.1 Stress

The effects of stress on the translation product and process have been explored mainly through studying the various effects of time pressure as presented in previous sections. However, experimental investigations of the psychological and physiological effects of stress have been scarce. To the best of the author's knowledge, only one recent study (Baghi & Khoshsaligheh, 2019) has investigated the stress level experienced by translators in written and sight translation. Baghi and Khoshsaligheh (2019) employed heart rate and blood pressure as physiological indicators of stress in their study and found that sight translation was a more stressful modality for translators as evidenced by a significant rise in both heart rate and blood pressure, compared with written translation. Elsewhere, instead of conducting an experiment-based study, Courtney and Phelan (2019) employed a quantitative questionnaire consisting of seven questions to measure the self-perceived occupational stress and job satisfaction levels of 474 translators in the UK. The most pronounced stressors revealed by the respondents included perceived unfair treatment from agencies, poor remuneration and uncertainty about the future. Despite these negative aspects about stress, the respondents generally exhibited high to extremely high job satisfaction.

While time pressure can be taken as a common source of stress for translators in their routine work, it is inherently embedded in interpreting which always has a highly restricted flexibility of time for rendering. In other words, the task nature inevitably makes the interpreters experience imminent and ongoing time pressure throughout the process, which makes interpreting a highly stressful activity. As a result, explorations of stress in interpreting activities are in abundance and span a wide range of topics with diverse methods involved.

In the field of interpreting research, substantial attention has been paid to the effects of anxiety or stress on interpreters' task performance as well. For instance, Roziner and Shlesinger (2010) revealed a relatively small impact of stress on either the quality of the interpretation or interpreters' health. However, stress usually points to considerable psychological effects, including an increase in feelings of isolation and alienation.

Chiang (2009, 2010) investigated the scope and severity of foreign language anxiety experienced by Taiwanese student interpreters during the learning process. The results showed that about one-third of student interpreters suffered from foreign language anxiety, though their foreign language anxiety was less prevalent and less severe than other Asian L2 learners' (Chiang, 2009). In addition, foreign language anxiety was found to be significantly and negatively associated with their learning outcome (Chiang, 2010). However, Jiménez Ivars and Pinazo Calatayud (2001) found that although low confidence in public speaking was significantly related to high scores in the anxiety level of interpreter trainees, the anxiety level was not negatively related to their interpreting performance. The authors suggested that the relatively high level of anxiety shown by the sample could be a sign of "positive anxiety" and coping strategies may be the explanation for this lack of influence. Similarly, Kao and Craigie (2013) suggested that coping strategies significantly accounted for interpreting stress. Using the strategies of problem-solving coping, seeking-social-support coping and reducing reliance on avoidance coping may lead to a reduction in interpreters' stress. This, in turn, may improve student interpreters' performance. Therefore, more attention should be paid to improving the students' awareness and ability for the proper utilisation of coping strategies in interpreting training.

Another focus of previous stress-related studies in the field of interpreting research has been the psycho-physiological responses of stress reflected by various biological indicators such as salivary cortisol, heart rate and galvanic skin response. For example, by measuring interpreters' cardiovascular activity before and after four interpreting shifts during a day's work, Klonowicz (1994) found pronounced elevations after the working shifts and that the mobilisation wave (i.e., the transition from the baseline state to an activated state) was manifest in all the three measures of the cardiovascular activity (systolic blood pressure, diastolic blood pressure and heart rate) while the normalisation wave (i.e., the transition from an activated state to the baseline state) occurred systematically for systolic blood pressure only. This suggested that, methodologically, these measures might be more robust in reflecting mobilisation in interpreting activities.

It was found in the study conducted by Moser-Mercer et al. (1998) that prolonged turns of simultaneous interpreting risked a decline in the quality of the output, which appeared to be due to a combination of the psychological (increase in stress emotions) and physiological (mental overload, as indicated by cortisol and immunoglobulin A

secretion levels) factors. Using the self-report and physiological measures, Moser-Mercer (2003, 2005) confirmed that remote interpreting was more tiring than live interpreting and interpreters were tired significantly more quickly as evidenced by a faster decline in the quality of performance over a 30-minute turn. Kurz (2002, 2003) explored how novices would differ from expert interpreters in stress levels during simultaneous interpreting. The results indicated that the expert interpreters' pulse rate values remained within the normal range and no marked signs of elevated stress were observed. However, the student interpreters' pulse rate was clearly higher than that of the expert interpreters, indicating that for novices even an ordinary classroom situation (working with a text that had been available in advance) involved measurably higher physiological stress. Thus, it was suggested that experienced conference interpreters can be expected to cope with the high demands of a technical conference and that training and experience helped them adopt the right strategies.

With the biomarkers of heart rate and blood pressure, Korpál (2016) investigated whether the speaker's rate of delivery would influence the level of stress experienced by interpreting trainees in a simultaneous interpreting task. The results showed that the participants' level of stress was significantly related to heart rate. Similarly, Rojo et al. (2021) found that student interpreters' grades were negatively correlated with their levels of State Anxiety, and they had a significant increase in heart rate during the interpreting task phase as compared with the baseline and recovery phases. It was suggested that heart rate was a sensitive marker to detect the students' stress response, but not to the level of influencing their performance. The authors also indicated that there is a need to triangulate different measures to obtain a complete picture of how stress and anxiety may impact interpreting performance.

In addition to the empirical experimental studies, several survey studies have been conducted aiming to investigate the level of occupational stress experienced by interpreters as well as to identify the most pronounced stressors for interpreters. Cooper et al. (1982) undertook a study which combined in-depth interviews and a large-scale survey to investigate the stress state of the conference interpreters in the International Association of Conference Interpreters (AIIC). Four general sources of stress (physical environmental factors, task-related factors, inter-personal factors and home/work interface) were identified and researched through the interview study. It also examined AIIC interpreters' present physical health, job satisfaction, type A/B personality, perceived stress on the job and mechanisms for coping with stress through the survey

study. The results suggested that the respondents experienced high level of work stress but with a relatively high level of job satisfaction; they were in good physical health in general, but at some risk of mental ill-health; overall, they tended to be slightly Type A oriented, but not at a particularly vulnerable coronary risk level; their job stressors can be grouped into environmental, task-related, organisation/management, delegate and personal categories.

Mackintosh (2003) reported a comprehensive study conducted by AIIC on the work stress experienced by interpreters, which involved various stressors and physiological responses using a variety of methods such as survey questionnaire, salivary cortisol, blood pressure and heart rate. The results from the survey research showed that 88% of the respondents expressed overall satisfaction in their work, and the main stressors identified were poor documentation, speed of delivery, difficult accents, undisciplined speakers and poor booth conditions. Also, 66% of the respondents self-reported high levels of work-related stress. The results from the experimental research revealed that both blood pressure values and salivary cortisol levels of interpreters were consistent with high-stress occupations and their stress levels were highest when they were on mike. The survey research and the experimental study reported in Cooper et al. (1982) and Mackintosh (2003) gave a comprehensive overview of interpreters' occupational stress and offered both practical and methodological implications about the issue of stress in interpreting activities.

Recently, Korpala (2021) extended the respondents of the survey research to both translators and interpreters, aiming to investigate the psychological stress among sworn translators and interpreters working in the Polish-English language pair in Poland. Several stressors related to certified translators' and interpreters' work (such as the speaker's delivery rate, translation and interpreting rates and non-cooperative translators and interpreters) and their medical symptoms (such as fatigue, back pain and eye irritation) as well as their coping strategies (mainly task-oriented coping) were identified. The translation and interpreting rates have been a significant stressor and a main reason for their dissatisfaction in their work.

Overall, it appears that much more attention has been devoted to stress in interpreting than in written translation, which may result from the general impression that interpreting is considered a more stressful and demanding activity than written translation. This situation, on the other hand, reflects that stress in written translation is still understudied, let alone using the well-developed psychometric and physiological

methods. The empirical interpreting studies could also shed some light on the methodologies in researching time pressure or stress in written translation as well as the underlying mechanisms of the interplay between stress and cognition in language processing.

2.2.2 Emotion

Time pressure can be a potential emotional trigger and the resulted affect states are likely to influence the processing of translation as well as the quality of the final product (Rojo, 2017). As shown in the literature reviewed earlier, both time pressure and stress can have conducive and obstructive effects on translation or interpreting performance. In this sense, it is also of relevance to follow the role different emotions or affective factors play in translation or interpreting activities, in order to gain a more comprehensive understanding of the underpinning mechanism of time pressure or stress in cognitive processes.

Different emotions may influence translation performance differently. For example, Rojo and Ramos (2016, 2018) explored how positive and negative emotions impact translation performance. The results suggested that positive emotions encouraged creativity in terms of cognitive breadth and flexibility while negative emotions fostered meaning accuracy in terms of item-specific processing. Meanwhile, it was found in Rojo and Ramos's (2014) study that translators translated faster when the stimuli were congruent with their ideological beliefs, which meant that emotional states could promote mood-congruent thoughts and actions.

Translators' or interpreters' emotional reaction to different stimuli is also different. Rojo et al. (2014) found that metaphorical and non-metaphorical translations of four emotions (i.e., happiness, sadness, rage and fear) were significantly different in their emotional impact on the recipients of the translation as revealed by the heart rate measure. Jankowiak and Korpál (2018) focused on bilingual emotional language processing and they suggested a decreased reactivity to emotional stimuli in L2 than in L1, and more pronounced reactivity to visual than auditory stimuli in L1. Korpál and Jasielska (2019) investigated whether simultaneous interpreter's empathy could be affected by the speaker's emotions. The results showed that interpreters were indeed affected by the speaker's emotions when compared to neutral speech. Thus, it can be seen that the ability of properly regulating one's emotions at work is an important

element for translation or interpreting performance.

As such, the role of emotional intelligence and emotional stability in translation and interpreting activities has been extensively researched from various perspectives. For example, the results derived from Pahlavani and Asroush's (2013) study showed that there was a positive and significant relationship between the learners' level of emotional quotient and their success in oral translation performance. It was found in Shangarffam and Abolsaba's (2009) research, however, that there was no significant relationship between the level of emotional intelligence and the quality of translation among student translators. In the same vein, Varzande and Jadidi (2015) reported, based on the data collected from 100 professional translators, that no significant relationship was found between their emotional intelligence and their translation quality, but their academic experience significantly affected their translation quality. The implication of the impact of professional experience on emotional intelligence has been manifested in Hubscher-Davidson (2016) as well, which found that literary and non-literary translators were significantly different in terms of emotion regulation as revealed by their trait Emotional Intelligence scores. Similarly, skilled interpreters were different from novice interpreters in their self-regulatory processes with respect to metacognition, emotion regulation, self-observation and self-judgement (Hild, 2014). Based on these results, emotional intelligence, including the ability of emotion regulation, might be associated with translators' professional experience, but is not necessarily linked to their translation performance.

The difference of personality and emotional traits of translators and interpreters and how translation or interpreting performance can be affected by such traits have also attracted attention in recent years. For example, Rosiers et al. (2011) found that interpreting students reported higher levels of self-confidence with regard to oral tasks and translation students reported distinctly more language anxiety. The results from the study conducted by Cifuentes-Férez and Fenollar-Cortés (2017) revealed significant relations between the facets of expressive suppression and negative expressivity with regard to student translators' performance, suggesting that those who tended to hide or inhibit emotional states and responses were more likely to perform better in translation tasks and those who tended to display negative emotions were more likely to perform worse in translation tasks. However, no statistically significant correlations were found between the individual difference variables (e.g., anxiety levels, integrative motivation, self-perceived communication competence and language skills) and their sight

translation performance. In addition, Cifuentes-Férez and Meseguer Cutillas (2018) reported that the more creative the translators, the more creative the translations they produce, and the higher the translators' self-esteem, the lower the scores for accuracy. Thus, the personality and emotional traits of translators and interpreters could possibly modulate their performance from various dimensions. Identifying and exploiting such traits could be beneficial in their work.

These studies have demonstrated some important aspects of emotion which are significant factors that can regulate translators' and interpreters' minds and eventually influence their translation or interpreting performance. Some aspects, such as emotion regulation and stress management, are highly relevant to the capability of managing time pressure properly in translation practice. Researching time pressure, in turn, can also provide empirical evidence and valuable accounts for the various emotional impacts on the translation processes and product.

2.3. Summary

This chapter provides a systematic review of previous literature on topics of time pressure, stress and emotion in both written translation and interpreting research. Previous findings predominantly demonstrated a negative effect of time pressure on the translation product, with a few cases showing the facilitative function of time pressure. Automatic skills and an unequal distribution of mental resources to different aspects of the translation task (e.g., translation quality vs. quantity) under different time conditions might provide some insights for such phenomena. Time pressure could also affect the translation process in terms of pausing behaviours which can signal different cognitive processes such as problem-solving activities; but such effects were usually limited to a certain extent, for example, within the revision phase (e.g., Jensen & Jakobsen, 2000). The importance of maintaining a fast translation speed especially under intense time pressure has been emphasised, since it indicates a higher level of automaticity and the ability of faster accessing the problem-solving resources (De Rooze, 2003).

It appeared that the influence of time pressure on the pattern of time distribution to different translation phases varied: the orientation and the revision phases were more likely to be trimmed under time pressure, and translator's experience level might make a difference as well (e.g., Kourouni, 2012; Künzli, 2007). In addition, time pressure could modulate translators' visual attention distribution, especially lessening the

amount of attention that is distributed to ST comprehension compared with that distributed to TT production (Sharmin et al., 2008; Hvelplund, 2011). The effect of time pressure on translators' choice of (direct or indirect) translation strategy is not evident (Jensen & Jakobsen, 2000). However, translators' experience level may modulate their choice of Knowledge Telling and Knowledge Transforming approaches to translation under different time conditions (Jensen, 2000); certain dimensions of student translators' learning style might be associated with their translation quality, which calls for a specific focus on such aspects in translation training (Kourouni, 2012).

Examinations of the psycho-physiological effects of time pressure or stress on written translation have been scarce. In contrast, such investigations have been relatively extensive in interpreting research with more diverse methods involved. Basically, by employing various psycho-physiological measures of stress, the effects of common stressors such as delivery rate, prolonged interpreting turns and whether it is remote or live interpreting, on interpreters and on interpreting performance have been inspected (e.g., Korpál, 2016; Moser-Mercer et al., 1998; Moser-Mercer, 2003, 2005). Furthermore, taking a broader view, time pressure and stress could be emotional triggers for translators. Different emotions (e.g., of different valency) generated during the translation process could lead to varied consequences in translation performance in terms of, for example, accuracy, speed, and creativity (e.g., Rojo & Ramos 2014, 2016); similarly, translators' emotional traits such as the ability of regulating emotions during work could be potential factors regulating their performance (e.g., Cifuentes-Férez & Fenollar-Cortés, 2017).

From the point of view of research methodology, in most existing studies investigating the effects of time pressure on written translation, time pressure was assumed to be induced by constraining the given timeframe before the time-pressure effects on task performance were concluded. Actually, apart from imposing deadlines as a basic trigger of time pressure, supplementary manipulation strategies for time pressure inducement and intensification are available (see section 4.1). In addition, before examining the time-pressure effects on task performance, applying valid measures to verify the successful inducement of time pressure is necessary. The psycho-physiological measures (e.g., blood pressure and heart rate) which are validated and widely applied in relevant interpreting research are appropriate candidates. These methodological concerns have been addressed in the present study (see details in Chapter 4) to further improve the internal validity and reliability of the findings.

Chapter 3. Theoretical Framework

Time pressure, from a psychological point of view, is inevitably associated with stress, a psychological construct that has been extensively researched. Preliminary theoretical models in stress studies merit attention before delving into the effects of time pressure on translation. In addition, the translation process as a complex and higher-order cognitive activity resembles a learning process and involves different subprocesses such as reading comprehension and text production. Theoretical frameworks accounting for such cognitive processes, developed in disciplines such as learning sciences and reading research, could also shed light on understanding the cognitive processes of translation and how time pressure could affect them. Combining such theoretical constructs, models and frameworks, this chapter describes an overall theoretical framework for the thesis: first of all, concepts of time pressure and stress as well as relevant models accounting for their relationship with human performance are discussed in section 3.1; section 3.2 provides a detailed introduction to Cognitive Load Theory which is developed in the context of learning sciences; criteria and theories about automaticity and the theoretical scopes of cognitive rhythm are described in section 3.3; finally, section 3.4 summarises the whole chapter.

3.1 Time pressure, stress and human performance

Understanding the interaction between cognition and time pressure/stress would be the basis for understanding the effects or consequences of time pressure on performance. This section illuminates the core of the concepts of time pressure and stress, involving the essential cognitive mechanisms and fundamental theoretical models underpinning their effects on human behaviours and performance.

3.1.1 Time pressure in cognitive activities

3.1.1.1 Time pressure, time pressure generation and the cognitive relevance

Intuitively, time pressure usually results from an externally imposed time constraint on a task. Such a time constraint, or deadline, can be seen as the most direct trigger of time pressure. As shown in the previous chapter, imposing a deadline has been the most common way to induce time pressure in empirical translation studies. Strictly speaking, time pressure is defined as “the difference between the amount of available time and the amount of time required to resolve a decision task” (Rastegary & Landy, 1993, p.

225). It is, nevertheless, necessary to distinguish between time constraint and time pressure. Time constraint only refers to the objective deadline imposed on a task. However, time pressure as a more complex construct, involves more components, including not only the external and obligatory deadline, but also the internal and subjective experience of time urgency in the course of executing a task. In fact, perception plays a central role in the notion of time pressure – time pressure arises when the available time is perceived to be insufficient (*ibid.*, p. 226).

Determining the components of time pressure underscores the importance of understanding how time perception functions to produce time pressure. A few theoretical models have been put forward which attempted to unveil the mechanism that underlies the generation of time pressure. The acceleration hypothesis (Miller, 1960), for example, takes the view that people decompose task problems into smaller, more manageable subtasks that are organised in time. Time pressure is induced in such cases when accelerated information processing occurs, which overloads one's capacity.

Moving forward, time perception theories, such as the contextual change model (Fraisse, 1963; Block & Reed, 1978), the segmentation model (Poynter, 1983, 1989; Poynter & Homma, 1983), and the storage-size model (Ornstein, 1969), basically take the number of intervening events during an interval as cues of time passage evaluation. This means that the more that the events occur within a given time period, the longer the period is judged to be (MacGregor, 1993, p. 76). In other words, an increase in the perception of time pressure can result from the increase in the number of cognitive events taking place within a given time period. MacGregor (1993) gives a concise summary of this account:

Time pressure, in this sense, is experienced as a disturbance in one's internal clock or ability to judge the passage of time. Urgency is exacerbated by mental activity; the more the processing of information is accelerated, the greater the sense of time pressure. This hypothesis predicts that the perception of time pressure is not linear with veridical time but increases exponentially as deadlines draw near. (p. 77)

On the other hand, the attentional model (Frankenhaeuser, 1959; Priestly, 1968) posits that subjective duration is related to the degree of attention that one is devoting to the passage of time itself. It means that the attempt to monitor time availability as demands of a secondary task would compete for cognitive resources with the primary task demands. Time pressure is the result of the cognitive overload that people

experience and is associated with having to keep track of time. Thus, a negative relationship between information-processing load required during a target interval and its subjective duration is predicted (Zakay, 1993, p. 63). This prediction, however, is essentially opposite to the prediction made by the above-mentioned time perception theories which posit a positive linear relationship between the number of cognitive events or contextual changes that occur during an interval and the estimated duration.

By differentiating the prospective and retrospective paradigms in subjective time perception, these models were unified by Zakay (1993). In the case of the prospective paradigm, subjects are aware of the existence of the time constraint and keep monitoring how much time has passed throughout the task performing process. In contrast, within the paradigm of retrospective time perception, subjects are not aware of the time constraint and thus devote all the attentional resources into the completion of the task itself during the task performing process. As a result, in the retrospective paradigm, they can only retrieve the temporal information and estimate the time after completion of the task. Therefore, as pointed out by Zakay (1993, p. 66), time pressure occurs at the prospective paradigm of time perception when subjects are aware of and occupied with the passage of time. Thus, studies with deadlines investigating time pressure effects naturally fall in the prospective paradigm of time perception and consequently the attentional model can be adopted: the attentional resources are allocated to task processing and monitoring the passing of time simultaneously. As the attentional resources are limited, the more the resources are devoted to monitoring the passing of time, the less attention is left for performing the task and vice versa.

These theoretical models reveal that the generation of time pressure during task execution is indispensably associated with the disturbed subjective time perception where cognitive processes take place and cognitive resources are consumed. It is thus not difficult to understand that cognitive perturbations occur with the presence of time pressure. In response to such consequences, different patterns of cognitive strategy have been identified, as discussed in the following section.

3.1.1.2 Mechanisms of time pressure functioning on cognitive processes

Maule and Hockey (1993) summarised the major patterns of cognitive strategy selection in decision making tasks that can be affected by time pressure. Specifically, approaches to studying strategy change at a macro-level, a micro-level and in a

combination of the macro- and micro-level are discussed separately. The macro-level strategy selection mainly concerns the change from using a compensatory rule to a non-compensatory rule which can be induced by the imposition of time constraints. Compensatory rules refer to strategies that involve a global rating of each alternative that allows a trade-off between attributes, i.e., a poor aspect on one attribute can be compensated for by a good aspect on another; non-compensatory rules, in contrast, do not tolerate trade-offs, with alternatives being rejected immediately as they fail to meet an acceptable level on any attribute (Maule & Hockey, 1993, p. 87). In general, compensatory rules are considered to be more complex and requiring greater mental effort but are likely to result in better or more optimal decisions (Einhorn & Hogarth, 1981, as cited in Maule & Svenson, 1993, p. 15). It follows that the change of strategy in different time conditions is based on a cost/benefit analysis at the macro-level. As suggested by Maule and Hockey (1993, p. 86), time constraints can alter the choice of strategy since “the preferred strategy for any particular choice/benefit combination may be unable to be implemented in the time allowed”.

The micro-level strategy concerns people’s peculiar ways of adjusting to information overload. The most relevant and widely acknowledged modes of adjustment in tackling time pressure are filtering, omission and acceleration, which were put forward by Miller (1960). Maule and Hockey (1993) summarised the definitions of these three strategies:

Filtering occurs when individuals neglect certain low-priority categories of information while continuing to process others. (p. 86)

Omission occurs when individuals completely ignore aspects of task information, for example, when they stop processing all task-related processing temporarily. (p. 86)

Acceleration involves a general speeding up of mental activity in order to keep up with increasing task demands. (p. 87)

In a combination of the macro-level and micro-level approaches and the principle of the cost/benefit analysis, Maule and Hockey (1993, p. 88) suggested that a hierarchy of ways of adapting to deadlines, varying as a function of the deadline stringency, can be identified in various experimental findings: individuals would firstly try to accelerate processing, while maintaining the same macro-level strategy; if acceleration is not sufficient to meet these demands, filtering is used to reduce the overall information-

processing load; if both acceleration and filtering are not effective, individuals would change the strategy in the direction of an increased use of non-compensatory rules. It is thus considered that such strategy changes are adaptive, and the adaptation follows the decision rules that can produce the outcome with the highest value or benefit in that situation.

The consideration of time constraints as a stressor in addition to as a mere task variable was integrated within a control model – Variable State Activation Theory (VSAT) by Maule and Hockey (1993), which takes into account both the critical role of effort in decision making under stress, and the significance of a cost/benefit analysis in determining the ways of adapting to time constraints. In this control model of stress regulation, the cognitive state (containing the available resources) produced by a stressor (such as a time constraint) may or may not be appropriate for current processing demands. A disruption of equilibrium could be brought about by the stressor. It is assumed that “the (frequent) discrepancies within the normal range are managed automatically through the use of low-cost, routine corrections” (*ibid.*, p. 92). However, if the discrepancy exceeds a certain level and cannot be resolved by routine regulation, a high-level controlled response is involved. To be specific, such a high-level regulatory activity is featured by increased effort and the deliberate planning of actions, or by the conscious selection of alternative goals. Four general options for such regulatory activities are identified (route 1, 2, 3 and 4 in Figure 3-1).

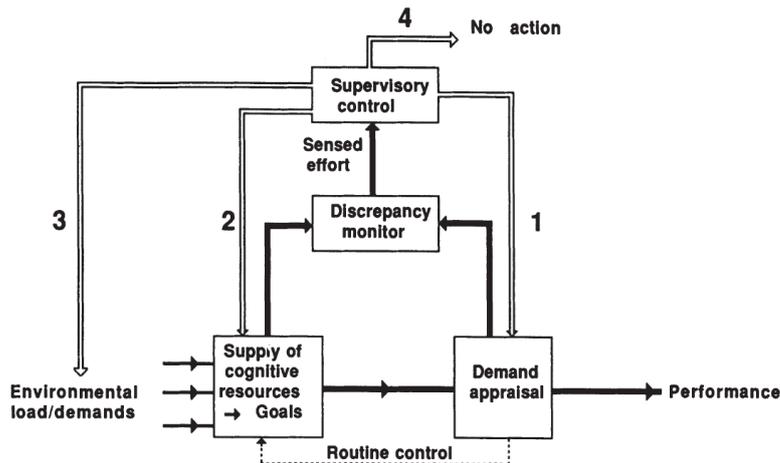


Figure 3-1. The control model of Variable State Activation Theory proposed by Maule and Hockey (1993), reproduced with permissions from the authors; copyright 1993 Springer.

Maule and Hockey (1993) described the first option as changing the cognitive state which is a means of direct control involving the maintenance of performance goals at the expense of increased effort (termed as a trying-harder reaction). In this case, the sensed discrepancy is managed through the recruitment of additional resources (i.e., effort), without changing the task goal. Such a strategy involves extended high-level executive (or supervisory) function (Hamilton et al., 1977; Hockey, 1986), and could be costly in terms of affective and physiological states, because maintaining effectiveness under continued stress or high workload is known to be associated with increased sympathetic activation or mental effort (e.g., Lundberg, 1982; Wilkinson, 1962; Frankenhaeuser, 1986). Therefore, this is the most effortful strategy among all the four options.

Maule and Hockey (1993) suggested that coping with a deadline (as a stressor) clearly demands an increase in the speed of processing, but acceleration is only feasible up to a point and there must be some upper limit on the speed of processing. This argument is empirically backed up by the finding of De Rooze (2003) in translation research (i.e., 200 words per ten minutes as the upper limit) as mentioned earlier in Chapter 2. Maule and Hockey (1993) explained that as the deadline becomes more stringent, the costs of further acceleration could outweigh the benefits gained, and thus allocating more central control resources in such a situation could be a less effective means of adapting. Therefore, two predictions were made: 1) in relatively important

situations, people should persevere longer with a control mode based solely on acceleration, because the costs of faster responding are more likely to be balanced by the benefits of using a complex and high-accuracy strategy; 2) acceleration induces an aversive state associated with high levels of effort that is difficult to sustain; this makes the direct control less likely to be successful when the decision process is extended over a prolonged period (*ibid.*, p. 94). This is why there must be some upper limit on the speed of processing.

The second option is to change the target state, which is a form of indirect control, involving “a downward adjustment of central motivational priorities, in order to meet available resources” (*ibid.*, p. 94). Adoption of such a low-effort regulation strategy may result from the consideration that the small marginal gains in decision accuracy from a trying-harder reaction may be outweighed by the high level of cognitive and emotional costs incurred. As such, a shift away from high-cost performance-oriented goals to the direction of low-cost goals (e.g., emphasizing personal comfort or emotional stability) may happen, which involves a reassessment of the costs and benefits of different goal-action plans (*ibid.*, p. 94). However, as indicated by Maule and Hockey (1993), changing the target state, or changing the goal for behaviours does not mean that the performance goals are abandoned altogether; rather, they are pursued with reduced priority (Schönplflug, 1983), or with modified specification requiring less use of high-cost processes such as working memory (Hamilton et al., 1977). In the case of dealing with strict deadlines as a stressor, increased use of lower-effort strategies such as filtering (at the micro-level) and a switch from compensatory to non-compensatory rules (at the macro-level) could amount to the adoption of this low-effort regulation strategy.

The third option, which represents the most direct means of control, is to change the environment. This strategy involves the modification or elimination of the stressor at source and thus is the most effective procedure for restoring the equilibrium. Obviously, this option is not always possible, especially in the case of externally imposed stressors such as working to externally imposed deadlines: it would involve a renegotiation of the deadline, or a reduction in workload, i.e., reducing the number or complexity of decisions to be made in the time given (Maule & Hockey, 1993, p. 95).

The last option is doing nothing in the face of the state of disequilibrium. This option reflects, as suggested by Maule and Hockey (1993), a lack of involvement in either the task or the environment which implies the individual’s inability to either

change the environment or implement a high- or low-effort operation (e.g., because of fatigue or unavailability of further cognitive resources). In a time-pressured condition, the lack of appropriate actions could be the result of, for example, a deadline being unreasonably strict for the work that must be done. In this situation, i.e., without access to normal coping resources, maladaptive states such as helplessness or panic (e.g., Seligman, 1975; Janis & Mann, 1977) may arise (Maule & Hockey, 1993, p. 96). Maule and Hockey (1993) summarised well the essence of this control model of stress regulation:

In conclusion, this approach assumes that individuals adapt to time constraints by comparing the demands of the situation to the current cognitive state and evaluating the costs and benefits of control activity to reduce any discrepancies identified by this comparison. Control activity involves expending effort, which is assumed to be limited in supply. If no appropriate control activity is available in a situation of some importance, then all the symptoms associated with panic and pressure would be expected. (p. 96)

In the context of an experimental setting in the present study where deadlines are externally imposed on the translation tasks, the third and the last options are generally not expected, because the deadlines are non-negotiable and doing nothing or giving up the task directly will result in a failure of the experiment. Thus, the focus of this study is put on the adoption of the first two options of stress regulation in terms of time pressure management, namely changing the cognitive state (a high-effort strategy) and changing the target state (a low-effort strategy). Specifically, by measuring the amount of effort devoted to the tasks with different time constraints, the effects of time pressure on the adoption of stress regulatory strategies can be illuminated.

Overall, the theoretical models and approaches about time perception and cognitive strategies associated with time pressure management elucidate the nature of how time pressure functions to affect complex cognitive activities. It is also evident, especially in the model of VSAT, that the construct of stress serves as a basis for understanding the effects induced by time pressure. The following section discusses some relevant and important aspects of stress and its relation to task performance.

3.1.2 Stress, coping and performance

3.1.2.1 The nature of stress

Although stress is a widely used term in everyday life, it is difficult to define. Cox (1978) delineated three categories of definitions of stress, namely, the response-based, stimulus-based and the interactional models. The response-based definitions generally take a physiological or biochemical approach, and stress in this approach is referred to as “a state manifested by a specific syndrome which consists of all the non-specifically induced changes within the biological system” (Selye, 1976, p. 64). The stimulus-based definitions (e.g., Broadbent, 1971; Hockey, 1979), as suggested by Sanders (1983), take stress as a convenient label and a collective noun indicating certain environmental and organismic conditions, such as noise, sleep loss and heat. Both the response-based and stimulus-based definitions lack the emotional component which is usually attached to the concept of stress in common sense understandings (*ibid.*). The interactional definitions, however, consider stress as an intervening variable referring to a state of divergence between perceived demands and capabilities to adapt (*ibid.*, p. 62). In this vein, Lazarus and Folkman (1984, p. 19) defined stress as “a particular relationship between the person and the environment that is appraised by the person as taxing or exceeding his or her resources and endangering his or her well-being”. The essence of the transactional definitions of stress is that a comparison between demands and resources leads to cognitive appraisals and subsequently to a stress response if the outcome of the comparison is unfavourable and difficult to correct (Sanders, 1983; Welford, 1973; Lazarus, 1966; Cox, 1978). Clearly, the control model of VSAT (Maule & Hockey, 1993) mentioned earlier on regulatory cognitive activities in coping with stress/time pressure follows such an interactional approach where the attempts to cope may come to terms with changing the perceived demands, or changing the required quality of adaptation, or changing both.

The nature of stress in the interactional model is the discrepancy that exists between the level of demand and the individual's ability or resources to cope. Cox (1985) epitomises the situation that is likely to be appraised as stressful taking a transactional approach:

A classic stressful situation is one in which the person's resources are not well matched to the level of demand and where there are constraints on coping and little social support. Stress, itself, is an individual psychological state. It is to do with the person's perception of the (work) environment and the (emotional) experience of it. (p. 1156)

As mentioned above, the emotional component of stress is emphasised in the transactional approach. Thus, an immediate response to a perceived stressful situation in the form of a negative emotional experience is common. The responses to stress can be categorised as psychological, behavioural or physiological, since the emotional experience of stress is often accompanied by changes in the person's perceptual and cognitive processes and in behavioural and physiological functions (*ibid.*, p. 1157). Chapter 4 discusses the details of these manifest responses as measures of stress applied in the current study from a methodological perspective.

3.1.2.2 The transactional theory of stress and coping

The most representative and influential theory taking the interactional approach to understanding and studying stress is Lazarus's transactional model of stress (Lazarus, 1966, 1976; Lazarus & Folkman, 1984). Stress is considered a product of the transactional system comprising two critical processes, i.e., cognitive appraisal and coping, that mediate the person-environment relationship (Figure 3-2). Cognitive appraisal involves a primary appraisal and a secondary appraisal. The primary appraisal refers to the process of evaluating and determining whether an event is stressful, i.e., to answer "Am I in trouble or being benefited, now or in the future, and in what way?" (Lazarus & Folkman 1984, p. 31). During primary appraisal, the event or situation can be categorised as irrelevant, benign-positive or stressful. An event is appraised as stressful when it is evaluated as a harm/loss, a threat or a challenge. A harm/loss refers to some damage that has already been sustained; a threat concerns harms or losses that have not yet taken place but are anticipated, which is associated with negative emotions such as fear and anxiety; similar to a threat, a challenge also calls for the mobilisation of coping efforts, but focuses more on the potential for gain or growth, which is associated with positive emotions such as eagerness and excitement. As emphasised by Lazarus and Folkman (1984, pp. 33-34), the relationship between threat and challenge appraisals can shift as an event unfolds: a situation that is appraised as more threatening than challenging can come to be appraised as more challenging than threatening because of the coping efforts that have been exerted, or through the changes in the environment that alter the troubled person-environment relationship for the better. This theory thus integrates both the motivational aspects of stress and the varying emotions that are associated with the experience of stress.

Secondary appraisal concerns the evaluation of what might and can be done in jeopardy, whether it be a threat or a challenge. A complex evaluative process takes place during the secondary appraisal, taking into account not only the available coping resources and options, but also the consequences of applying a particular coping option, and the potential internal and/or external demands that might be occurring simultaneously. Moreover, Lazarus and Folkman (1984, p. 35) suggested that secondary appraisals of coping options and primary appraisals of what is at stake interact with each other in shaping the degree of stress and the strength and quality (or content) of the emotional reaction. When there is a mismatch between the situational demands (e.g., risk, uncertainty and difficulty) and one's resources (e.g., physical resources, psychological resources and competencies) to cope with it, stress arises, and different kinds of emotional reaction could be elicited depending on the cognitive appraisal processes. This leads to an engagement in coping processes.

Cox (1985, 1987) suggests that coping denotes a particular type of stress response and can be conceived of as a problem-solving behaviour. Coping involves both cognitive and behavioural strategies, representing “either an adjustment to the situation, or an adjustment of the situation” (Cox, 1985, p. 1157). Lazarus and Folkman (1984, p. 141) defined coping as “constantly changing cognitive and behavioural efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person”. Coping functions in two ways: 1) by managing or altering the source of the problem which is referred to as problem-focused coping (Lazarus & Folkman, 1984) or direct action (Lazarus 1966, 1976); 2) by regulating the emotional response to the problem, which is referred to as emotion-focused coping (Lazarus & Folkman, 1984) or palliation (Lazarus, 1966, 1976). No matter whether the coping is successful or not, the events implied by coping will feed back to alter the individual's initial appraisal (i.e., making a reappraisal) of the environment.

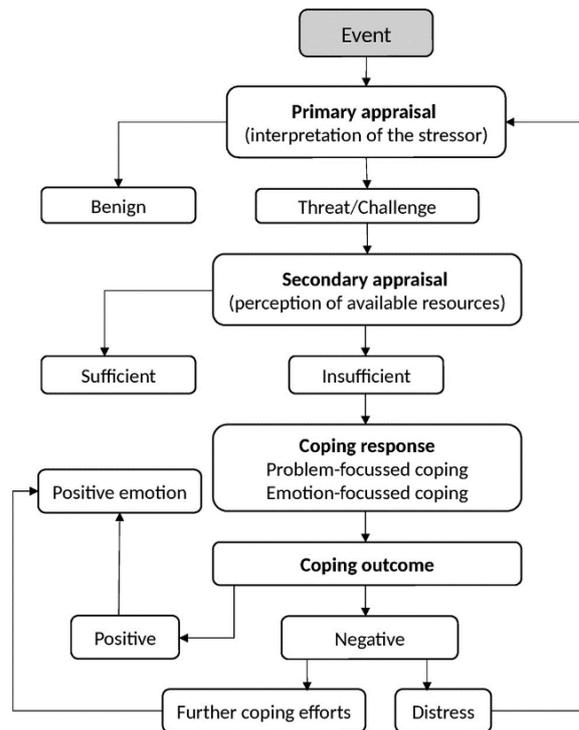


Figure 3-2. A flowchart of the transactional theory of stress and coping based on Lazarus and Folkman (1984) produced by Turner-Cobb and Hawken (2019). Permissions have been obtained to reproduce; copyright 2019 Cambridge University Press.

As the problem-focused coping in this framework of transactional theory of stress essentially refers to the cognitive and behavioural efforts that are devoted to managing the (task) demands, it converges closely with the construct of *effort* described in the VSAT model (Maule & Hockey, 1993), i.e., changing the cognitive state by recruiting additional resources, which serves as an option of direct control in regulating stress/time pressure. Thus, problem-focused coping is operationalised as *effort* in the present study, which can be measured by different behavioural and psycho-physiological metrics (see details in Chapter 4).

Overall, the transactional model of stress views the individual and the environment in a dynamic, mutually reciprocal, bidirectional relationship in which what is a consequence at the antecedent time point can become a causal variable at the subsequent time point; this can be caused by the ongoing coping operations associated with either changing the environment or the individual (Lazarus & Folkman 1984, p. 293). In this way – our appraisals drive our response, our response changes the situation

or ourselves, and this change itself affects our appraisals – the system is transactional.

3.1.2.3 The relationship between stress and performance

Taking the view of the transactional theory, stress occurs when there is a difficulty in adapting to the environmental demands. This may naturally imply a negative effect of stress on performance. When concentrating solely on the relationship between stress and performance despite the complex contextual factors, the negative linear theory gained more support with regard to empirical results (e.g., Allen et al., 1982; Friend, 1982; Greer & Castro, 1986; Harris & Berger, 1983; Jamal, 1984; Lagace, 1988; Westman & Eden, 1991, 1996). A basic premise for this theory is that stress at any level consumes an individual's time, energy and attention, taking away resources from the task; consequently, stress is conceived as adverse to an individual and dysfunctional for an organisation by its very nature (Jamal, 1985; Muse et al., 2003). Vroom (1964) also laid the theoretical groundwork for the negative linear theory with two potential explanations: 1) stress can narrow down an individual's perception, resulting in important information and cues being ignored; 2) stress can elicit involuntary physiological responses that interfere with performance. Thus, an impairment of performance under stress is predicted.

A positive linear relationship between stress and performance is postulated based on the belief that stress (or anxiety) is generally equated with challenge, and challenges are often viewed as occasions for constructive activity and improved performance (Meglino, 1977, as cited in Jamal, 1985, p. 410). In the transactional model of stress, challenge, as one type of cognitive appraisal in a stressful situation, is associated with an investment of effort striving for potential gains or growth accompanied by positive emotions such as eagerness and excitement. Thus, in the positive linear theory, stress functions as a motivational drive to promote performance. However, as mentioned, challenge only represents one facet of the construct of stress in the transactional theory – other facets of stress such as harms/losses and threats should be taken into consideration equally in regard to comprehensively assessing the relationship between stress and performance.

The famous Yerkes-Dodson Law (Yerkes & Dodson, 1908) identifies a curvilinear relationship between physiological arousal and performance (which is also referred to as the inverted U theory, see Figure 3-3). Essentially, the law suggests that an increase

in arousal up to a certain level can help to boost performance, while both under- and over-arousal impair task performance. Arousal refers to the psychological experience of energy, mobilisation, activity, tension, alertness or quietness (Russell & Barrett, 1999). The dimension of arousal ranges from deactivation to activation. The moderate level of arousal with regard to performance in the inverted U-shaped curve is usually considered as a “good stress” or “eustress” (Selye, 1976). The rationale for this relationship is that in a moderately aroused state, increased attentiveness activates and mobilises all cognitive resources which optimally facilitates performance (Westman & Eden, 1996). In contrast, when the physiological arousal is at a negligible level, the cognitive resources would be inactive, i.e., in a state associated with boredom and a lack of motivation, which leads to a deteriorated performance. On the other hand, in an excessively aroused state, more energy or cognitive resources are needed to manage the internal strain provoked by the external stressor, or “distress”, which is related to negative emotions such as anxiety and depression. This can also lead to an impaired performance because of little investment of effort in the task itself. Therefore, a moderate level of arousal can cause the individual to be activated and to expend maximal energy and cognitive resources on the task (*ibid.*, p. 166).

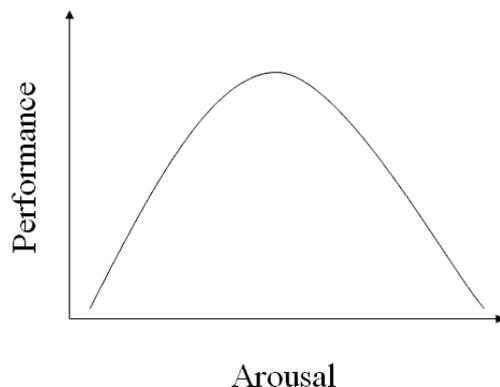


Figure 3-3. The inverted U-shaped arousal-performance relation

Other systematic attempts to uncover the underlying mechanism of the inverted-U theory include Easterbrook (1959) who suggested that the biphasic form of the arousal-performance function was mediated by an attentional mechanism, with increases in arousal reducing the range of cues used. As mentioned earlier, Vroom (1964) considered the narrowing of attention induced by increased stress as a detrimental force undermining performance, which serves as one of the accounts for the negative

relationship of stress and performance. However, Easterbrook's idea is that initially, the narrowing of attention elicits a reduction in the processing of task-irrelevant cues, resulting in an improvement in performance; when the arousal level goes higher, further perceptual narrowing leads to an increasing neglect of task-relevant information and a consequent impairment in effectiveness (Easterbrook, 1959, as cited in Maule & Hocky, 1993, p. 90). In this sense, such a mechanism of the narrowing of attention naturally comes with a filtering function to engender a higher priority of the processing of central or task-relevant information.

Eysenck (1982) indicated that changes in behaviour consistent with the narrowing of attention hypothesis occur only on some but not all occasions, so that the narrowing of attention is an active response under voluntary control instead of a passive and automatic process: it represents one among a number of ways in which individuals may adapt to stressors and therefore need not always occur following an increase in arousal. Frankenhaeuser (1986) thus distinguished two modes of active and passive coping as states of eustress and distress, respectively. Eustress is a positive responsive state involving effort and challenge, activation of the sympathetic nervous system and high levels of effectiveness, but without negative affective responses; distress, by contrast, is associated with anxiety and helplessness, as well as an increased activity in biochemical indices of stress such as cortisol (Frankenhaeuser, 1986, as cited in Maule & Hocky, 1993, p. 91). Integrating these insights, the inverted-U curve can be interpreted as suggesting that before arousal increases to the optimal level, the individual exerts active coping, focusing mostly on the task-relevant information to deal with the environmental demands, leading to an improved performance; but when arousal exceeds the optimal level, the coping style switches to the passive mode with more negative emotions generated, leading to impaired performance.

Among the numerous factors that may affect the perspective of the relation between stress and performance, the dimension of how stress is accounted for should be taken into consideration. For example, taking the response-based definitions (i.e., stress is considered as a response of our body), stress and arousal can hardly be distinguished. Consequently, the Yerkes-Dodson Law (an arousal/performance function) can directly denote one possible relation of stress and performance. Within the transactional paradigm of stress, however, stress as a response to the imbalanced person-environment relationship must reflect a reaction to the disturbed homeostatic state (i.e., arousal); hence, this reaction cannot fully coincide with that state (Sanders,

1983, p. 64). Theoretically, a positive relationship between stress and arousal should be illuminated first before mapping the concept of stress to arousal in the Yerkes-Dodson Law. However, stress and arousal are independent constructs which do not always overlap in theory. The intricate mechanisms linking the energetic mobilisation (arousal) and stress response have been researched extensively (e.g., Aston-Jones & Cohen, 2005; Cohen et al., 2004; Hockey, 1993, 1997, 2011, 2013; Kahneman, 1973; Kane & McVay, 2012; Kanfer & Ackerman, 1989; McVay & Kane, 2010; Sanders, 1983; Shallice et al., 2008; Stuss et al., 1995; Stuss et al., 2005; Unsworth & Robison, 2020), but a closer scrutiny is beyond the scope of this discussion. Basically, the ubiquitous discrepancy between stress and arousal may account for the most important intrinsic problem in the inverted-U theory between stress and performance and for the relatively scarce empirical support for this theory.

Overall, the above-mentioned theoretical hypotheses provide a primary framework for empirical investigations of the effect of stress on translation task performance and offer some valuable grounds that can facilitate our understanding for the observed phenomena. However, the divergencies which exist among the apparently different theoretical hypotheses (i.e., the negative linear, positive linear and the inverted-U theory) imply that the relationship between these two variables could be extremely complex and may vary across different situations and contexts.

3.2 Cognitive Load Theory (CLT)

The discussions above on both time pressure and stress reveal that cognition is innate in the process of time pressure/stress functioning on human behaviour and task performance. Effort, which is related to concepts of workload or cognitive load is a major variable that can be affected by time pressure/stress during the execution of cognitively demanding tasks. Cognitive Load Theory (CLT; Ayres & Paas, 2012; Paas, Tuovinen, et al., 2003; Paas et al., 2004; Paas & Sweller, 2012; Paas & Van Merriënboer, 1994; Sweller et al., 1998; Van Merriënboer & Sweller, 2005, 2010), which was developed in the context of learning sciences, provides a basic framework for investigations into the complex cognitive processes involved in learning activities. This framework aims at improving the effectiveness in instructional design for learning. Translation is a case of learning in complex cognitive domains (Paas & Van Merriënboer, 1994), which involves intricate problem-solving and decision-making

activities during the process and has the goal of achieving a better performance. When we perceive each translation task as a learning process, all the elements in a translation task setting, such as the difficulty of the translation materials, time constraints, access to external documentation resources can be taken as parts of the instructional design for the learning task. This section introduces CLT as a theoretical foundation for the present study in understanding the interplay between time pressure and (different) cognitive loads in translation tasks. The theoretical basics of CLT are presented in 3.2.1, followed by an illustration of the construct of cognitive load (3.2.2) and the three types of cognitive load delineated in the theory (3.2.3); the expertise reversal effect based on CLT is introduced in 3.2.4 and a delimitation of relevant theoretical concepts associated with cognitive load is given in 3.2.5.

3.2.1 The theoretical basis of CLT

The theoretical framework of CLT is built upon our current knowledge about human cognitive architecture which consists of a long-term memory and a working memory (Paas, Renkl, et al., 2003). All conscious cognitive processing occurs in the working memory, and it is severely limited in its capacity of information processing (Cowan, 2001, 2014; Miller, 1956). The limits of working memory, however, can be largely eliminated when it processes familiar information organised in cognitive schemas in the long-term memory. Schema is a pattern of thought or behaviour that organises categories of information and the relationships among them (DiMaggio, 1997). Cognitive schemas can store and organise knowledge by incorporating or chunking multiple elements of information into a single element with a specific function (Paas, Renkl, et al. 2003, p. 2). Schemas can develop by incorporating a huge amount of information, evolve from lower level to higher level and form further complex schemas. This is the process of schema construction. More skilled performance can be achieved when a schema is brought from long-term memory to working memory to govern an activity. Unconscious processing of those schemas, which is called schema automation, can further free up the capacity of working memory for more knowledgeable learners when they process previously learnt information stored in long-term memory (Choi et al., 2014, p. 227). It is by this process that human cognitive architecture handles complex material that appears to exceed the capacity of working memory (Paas, Renkl, et al. 2003, p. 2).

3.2.2 The construct of cognitive load in CLT

In the original model developed by Paas and Van Merriënboer (1994), cognitive load is a multidimensional construct involving a causal dimension and an assessment dimension (see Figure 3-4). The causal dimension reflects the interaction between the task characteristics and the learner characteristics. Specifically, task characteristics include factors such as task complexity and time constraints; learner-specific characteristics comprise factors such as expertise level (prior knowledge) and age. Later, Choi et al. (2014) revised the model by adding a third causal factor, i.e., the physical learning environment, which can interact with learner characteristics, task characteristics or a combination of both. The assessment dimension reflects the measurable concepts of mental load, mental effort and performance. The concept of mental load in this framework refers to the load which originates from the interaction of the task and learner characteristics (as well as the physical learning environment characteristics in the revised model); mental load can be determined by our current knowledge about these characteristics. Thus, mental load can be considered as an a priori estimate of cognitive load indicating the expected cognitive capacity demands (Paas, Tuovinen, et al., 2003, p. 64). Mental effort refers to the amount of cognitive capacity or resources that are actually allocated to the task demands. As such, mental effort can be measured while participants are working on the task to reflect the actual (instantaneous) cognitive load. Paas, Tuovinen, et al. (2003, pp. 64-65) pointed out that the intensity of effort being expended by learners can be considered the essence to get a reliable estimate of cognitive load. Performance is a third assessment aspect of cognitive load that can be referred to as the learner's achievements, such as number of errors or output quality.

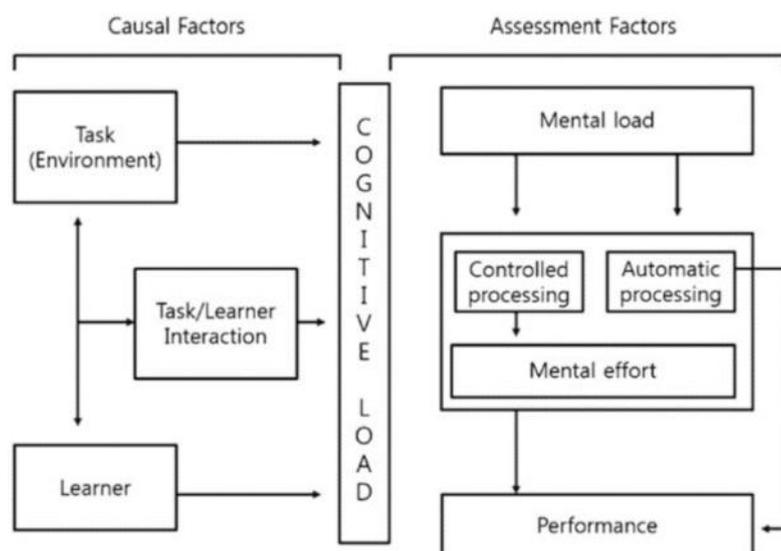


Figure 3-4. Dimensions of the construct of cognitive load in CLT proposed by Paas and Van Merriënboer (1994), reproduced with permissions from the authors; copyright 1994 Springer.

3.2.3 Three types of cognitive load in CLT

For the purpose of differentiating the invested cognitive resources during learning, CLT distinguishes three types of cognitive load, i.e., intrinsic, extraneous and germane load. In the context of optimising the learning process by instructional design, intrinsic cognitive load is the load that cannot be directly influenced by the way the task is presented whereas both extraneous and germane cognitive load can be manipulated by instructional design (Paas, Tuovinen, et al., 2003).

Intrinsic load is determined by the interaction between the nature of the material being processed in the cognitive task and the expertise of the learners. It is thus considered that both task complexity and learners' prior knowledge are important factors for the assessment of intrinsic cognitive load. Since experts are able to treat associated elements of information as one based on their existing schemas, they are able to effectively augment their working memory capacity; novices, by contrast, are not able to do so due to a lack of schema construction and automation (Artino, 2008; Klepsch & Seufert, 2020). Therefore, intrinsic cognitive load can only be influenced by manipulating the complexity of the task material, or by activating the learners' prior knowledge and existing schemata.

Extraneous load, which is also called ineffective load, refers to the extra load beyond the intrinsic cognitive load (i.e., task irrelevant elements) resulting from the instructional design of the cognitive task (i.e., the way the task is presented). Task irrelevant elements which may generate unproductive load include activities such as “searching for relevant and masking/ignoring the irrelevant information, or the need to process an unnecessarily high number of elements simultaneously in working memory because of instructional design factors” (Klepsch & Seufert, 2020, p. 47). As such, intensive time pressure imposed by task design can be a source of extraneous cognitive load during task execution because it could cause an unnecessarily high number of elements being processed simultaneously in working memory. High extraneous cognitive load therefore negatively affects the acquisition of knowledge or schema construction and should be reduced.

Germane load, which is also called effective load, refers to the cognitive resources dedicated to schema acquisition, construction and automation (Paas, Tuovinen, et al., 2003; Moreno & Park, 2010; Chen et al., 2016). Thus, germane load is a productive load that is directly devoted to processing task relevant elements. Germane load was also found to be positively related with motivation; higher intrinsic motivation resulted in a higher reported ability to devote cognitive resources to learning (Klepsch & Seufert, 2020, p. 49). Therefore, germane load is associated with the effective activation of the learner’s cognitive resources to help with schema construction and automation, which fosters learning.

A basic principle in CLT is that these three types of cognitive load are additive and the *overall cognitive load* (i.e., the sum of the three types of cognitive load) should stay within the working memory limits (Paas, Tuovinen, et al., 2003). Therefore, an increase of each individual type of cognitive load can contribute to the increase of the overall cognitive load. It is assumed that the level of extraneous load has a negative relation with learning outcome while the level of germane load exerted during the process has a positive relation with learning outcome. Since extraneous load is reducible and germane load is increasable by instructional design, extraneous load should be minimised to free up working memory capacity while germane load should be maximised to boost task performance.

3.2.4 Expertise reversal effect

The central goal of instructional design for complex cognitive tasks is to facilitate schema construction and automation. However, such a goal can be severely hindered by the limitations of working memory. Since novel information must be processed by working memory when it is presented to an individual, proper allocation of available cognitive resources to effectively utilise the limited capacity of working memory is essential. This means that if the limited resources are spent on activities that are not directly related to schema construction and automation (e.g., with high extraneous load), learning could be inhibited (Kalyuga et al., 2003, p. 24). By contrast, information can be processed more efficiently by enabling the use of schemas stored in long-term memory through which the limitations of working memory can be considerably reduced. Correspondingly, it is assumed in CLT that the limited capacity of working memory becomes effectively unlimited when dealing with familiar material previously stored in an immense long-term memory holding many schemas that vary in their degree of automation (Paas et al., 2004, p. 2). This underlines the difference in information processing manners between individuals who are equipped with the domain-specific schemas (experts) and those who are not (novices). The ability of automatic processing of schemas mastered by experts which requires minimal working memory resources allows problem-solving activities to proceed with minimal effort (Kalyuga et al., 2003, p. 25). Thus, it is clear that increasing one's level of expertise in a domain is a major means of reducing working memory load.

While experts can effectively reduce the burden on the limited capacity of working memory by enabling the acquired schemas held in long-term memory, novices, or inexperienced individuals, may experience high working memory load due to the lack of such sophisticated schemas in their long-term memory. However, the instructional procedures of a task can act as a substitute for such missing schemas and, if effective, act as a means of constructing schemas by reducing the extraneous load and increasing the germane load (*ibid.*, p. 25). In spite of the constructive effects of such kinds of instructional procedures designed for assisting schema construction for novices, they may not be needed by experts whose schemas have already provided full guidance for task processing. If experts are unable to avoid attending to the information provided by the additional instructional procedures, there will be an overlap between the schema-based and the redundant instruction-based components of guidance; processing such redundant components will require additional working memory resources and might cause a cognitive overload (*ibid.*, p. 25). As a result, instructional procedures, which

may be essential for novices to reduce extraneous load and increase germane load, may have negative consequences for experts leading to higher extraneous load. This is referred to as the *expertise reversal effect*.

3.2.5 Cognitive or mental? Load or effort?

Since concepts such as cognitive load, mental load, cognitive effort and mental effort have appeared extensively and variedly in several overlapping research domains, it is necessary to distinguish them and clarify their usage in the current study. It should be noted that overall, they are psychological constructs that help make sense of human performance including learning and other cognitive processes. Broadly speaking, cognitive load (in the psychology of learning) refers to the load experienced by working memory when humans engage in a variety of cognitively intensive tasks (Chen et al. 2016, p. 34). As shown in the framework of CLT, the construct of cognitive load represents the load imposed on the learner's cognitive system when performing a particular task. This term in the CLT framework is an overarching and more abstract construct that can be assessed from different and more concrete aspects, i.e., mental load, mental effort and performance.

Mental load, or mental workload, originally referred to the difference between task demands and the person's ability to master these demands (Moray, 1979). In this sense, the core of mental load is the perceived difficulty level of a task. Later, the construct of mental load was expanded to the extent that some psychological factors, such as demand expectations, the actual effort expended during performance and the perceived adequacy of performance, are taken into consideration; in other words, an individual's motivation, willingness or perceived relevance of the task to the individual's personal goals have become important aspects in estimating mental load (Moreno & Park, 2010, p. 10). Thus, mental load can be seen as a psychological experience resulting from the interaction between subjective individual characteristics and objective task and environment characteristics (Campbell, 1988; Kantowitz, 1987; Wood, 1986). The concept of mental load in CLT, however, focuses more on the objective characteristics of the task, environment and learners (e.g., age and prior knowledge). The subjective characteristics of learners, such as motivation, are reflected more in the concept of mental effort in CLT. As emphasised by Paas, Tuovinen, et al. (2003, p. 65), instructional manipulations to change the mental load will only be effective if people

are motivated and actually invest mental effort in them. Thus, mental effort, which refers to the amounts of resources being allocated to the task, is considered the essence to achieve a reliable estimate of cognitive load (*ibid.*, pp. 64-65).

The use of the term *cognitive effort* in translation process research is gaining increasing popularity especially in recent years. This use may be traced back to Giles' (1995) influential effort models for interpreting which was later applied into written translation research (e.g., Dragsted, 2010). Also, Krings (2001, p. 179) adopted the idea of (different types of) effort required in post-editing process and defines cognitive effort as a concept that “involves the type and extent of [the] cognitive processes that must be activated in order to remedy a given deficiency in a machine translation”. Studies related to post-editing thus adopted this concept as well (e.g., Vieira, 2015). Westbrook and Braver (2015, p. 397) indicated that the psychological construct of cognitive effort bears the motivational quality and volitional nature, concentrating on “decisions about whether to engage, and also about the intensity of engagement”. This nature is similar to the construct of *mental effort* described in CLT. In the context of translation process research, cognitive effort generally refers to “the mental effort involved in reading the texts, thinking about how to translate and how to correct mistranslations, selecting the desired product, and reflecting on the chosen solutions” (Lacruz, 2017, p. 386). This definition comes with more contextual features associated with the activity of translation.

As mentioned in section 3.1.2.2, the term *effort* is used as the operationalised construct of problem-based *coping* in the transactional theory of stress (Lazarus & Folkman, 1984). Meanwhile, *effort* also corresponds to the active regulatory activity which involves changing the cognitive state by recruiting additional resources in the control model of stress regulation (Maule & Hockey, 1993): *effort* refers to such resources that are recruited to respond to the state of disequilibrium in a stressful condition.

Within the framework of CLT, the construct of (*overall*) *cognitive load* refers to the sum of intrinsic, extraneous and germane load. Since one of the measurable concepts of (*overall*) *cognitive load* is mental effort, which is the amount of cognitive capacity or resources that are actually allocated to the task demands, (*overall*) *cognitive load* thus appears to converge closely to *effort* (referring to problem-based coping and active regulatory activity) as well. However, (*overall*) *cognitive load* (when measured by mental effort) is distinguished from *effort* in the sense that (*overall*) *cognitive load*

encompasses certain correlational relationships with its components (e.g., it is the sum of the three types of cognitive loads; extraneous load being inversely related with performance; germane load being positively related with performance) in CLT, whilst *effort* does not necessarily have such attributes. This distinction is made for the purpose of filtering the measures of (*overall*) *cognitive load* from a larger pool of *effort* measures by utilizing the internal relationships of different cognitive loads with performance. A detailed illustration of this methodological application of CLT is presented in Chapter 4. In summary, *effort* is used in discussions detached from the CLT framework, whereas (*overall*) *cognitive load* (rather than mental effort, to avoid confusion) is used in discussions within the CLT framework in the present study.

3.3 Automaticity and cognitive rhythm

Automaticity is the ability to perform a task with little or no central cognitive control (Anderson, 2009, p. 86). Skill acquisition or development of expertise is usually closely related with one's degree of automaticity: complex skills gradually evolve towards the direction of becoming more automated and requiring fewer processing resources. Central to automaticity is repetitive practice, and the general effect of practice is to reduce the central cognitive component of information processing (*ibid.*, p. 85). When the central cognitive component of a task has been practised so much that the task requires little or no conscious thought, doing the task becomes automatic. Automaticity has gained attention in translation research in recent years because of its relevance to the acquisition of translation competence or expertise (e.g., Jääskeläinen & Tirkkonen-Condit, 1991; Dragsted, 2004; Campbell & Wakim, 2007; Hvelplund, 2011, 2016; Deckert, 2017). Cognitive rhythm describes various temporal features associated with the distribution pattern of different types of events at both the macro-level and micro-level during the production process of translation. With a temporal nature, cognitive rhythm could be an important aspect that is easily influenced by time pressure in translation. This section examines the formal criteria of automaticity in section 3.3.1 and introduces the automaticity theory on reading proposed by Samuels (2006, 2013) in section 3.3.2, which could foster our understanding of automaticity in translation; the theoretical scopes of cognitive rhythm in translation research are presented in section 3.3.3.

3.3.1 Criteria of automaticity

Logan (1997) delineated five criteria (i.e., speed, effortlessness, autonomy, unconsciousness, co-occurrence of properties) of automaticity by thoroughly reviewing the relevant literature. These criteria can serve as the basis for distinguishing automatic processing from non-automatic (e.g., controlled, effortful or strategic) processing. Firstly, automatic processing is fast, and an increase in speed is a characteristic of the development of automaticity. Although it is difficult to defend an absolute criterion for speed in order to be considered automatic, it is clear that performance gets faster with practice; thus, the speed criterion is relative. An example given by Logan (1997) is that high-frequency words, which are more practised, are read more rapidly than low-frequency words (Seidenberg & McClelland, 1989). In translation research, De Rooze (2003) also emphasised high processing speed as an important characteristic of automaticity level especially under time pressure, and it deserves more attention in the training of translation students to improve their competence. Processing speed can be relatively conveniently measured in different cognitive activities using basic techniques (e.g., calculating the reading/writing/translating speed in relevant activities and tracking the reaction time in perception and response related activities).

Secondly, automatic processing is effortless. Logan (1997, p. 125) suggested that the effortless property of automatic processing reflects first as a sense of ease and second as the ability to do another task while performing an automatic one; thus, if two tasks are done simultaneously without interference, at least one of them is automatic. In a translation activity, this property of automaticity reflects, similarly, firstly a sense of ease in processing, and secondly the ability of parallel processing, which has been suggested as an explanation of why skilled translators can produce translation faster and better than novice translators (Jääskeläinen & Tirkkonen-Condit, 1991; Hvelplund, 2011). Parallel processing occurs in translation when some subprocesses (e.g., reading and typing) overlap, whereas sequential processing characterises the type of processing when those subprocesses take place in sequence and do not overlap (Balling et al., 2014). Thus, two aspects in translation activities can be examined in terms of the property of effortlessness in automaticity: the sense of ease in processing and the degree of parallel processing.

The third criterion for automatic processing is autonomy. This means that an automatic process begins and runs on to completion without intention, and once the

automatic process runs, to stop it would require additional inhibition efforts. This property of automaticity can be seen in the classic Stroop effect: the automatic process of reading takes precedence over colour naming, which is more cognitively demanding. Hvelplund (2011, 2016) indicates that at least two subprocesses in a translation activity can be automated, namely reading and typing, and while reading and typing rely on intentional initiation, the continuation of these activities can occur automatically. This may signal the autonomy property of automaticity in the translation process, which could, to some extent, be reflected in the translator's translation processing speed as well.

The fourth criterion mentioned by Logan (1997) is that automatic processing is not available to consciousness. Automatic processes such as semantic priming are unconscious in that they can occur without the subject being aware of the stimulus that produced them. But this property remains controversial both theoretically and empirically (see Logan, 1997 for more details). Lastly, automatic processes should share all the properties associated with automaticity (i.e., they should be fast, effortless, autonomous and unconscious). However, there are cases of violations and exceptions of this property (e.g., Regan, 1981). Logan (1997, p.128) suggested that automaticity can be viewed as a continuum rather than a dichotomy; in this case, one can expect the co-occurrence of properties at the beginning and the end of the continuum but not in the middle because different properties may change or evolve at different rates. That is, if autonomy develops before effortlessness, then it may be easy to find cases of effortful autonomous processes. Thus, the emphasis should be shifted from defining automaticity in terms of the co-occurrence of properties to investigating the practice effects and the acquisition of automaticity.

In the area of language processing, Segalowitz and Hulstijn (2005) suggested that automaticity is characterised as rapidity, effortlessness, unconsciousness and of a ballistic (unstoppable) nature, but they do not necessarily have to bundle together. Apparently, most of these properties are overlapped with those proposed by Logan (1997); however, the divergence appears at the last property, i.e., whether the properties of automaticity should be treated as a unitary construct. Segalowitz and Hulstijn (2005) argued that:

It would be an error, therefore, to assume without first doing the requisite empirical research that extensive practice leading to expertise will unfailingly result in

performance that has all the characteristics typically associated with automaticity. (p. 372)

Therefore, the property of co-occurrence of all the properties associated with automaticity remains controversial and is lacking a base in reality. Based on these theoretical accounts for the properties of automaticity, the current study will focus mainly on the measurable aspects or properties of automaticity in the translation process, including processing speed and effortlessness (from the perspectives of the ease of processing and parallel processing degree). The operationalisation of these aspects is introduced in Chapter 4 in detail.

3.3.2 Automaticity and the role of attention

Automaticity is a ubiquitous construct in reading research, which has been associated closely with how attention is allocated to complete a reading task. Basically, three components are considered essential in the reading process according to Samuels (2002, 2006, 2013):

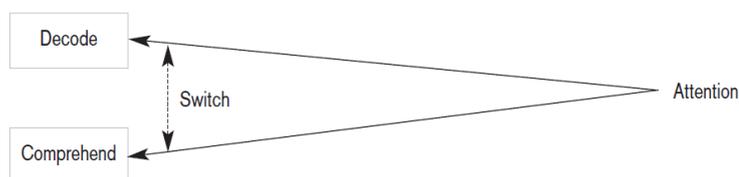
- 1) Decoding, which is the ability to pronounce the printed words, or to translate the printed words into spoken words, though the “spoken words” are not necessarily uttered aloud.
- 2) Comprehension, which is a process in which meaning is constructed using information from the printed words and the knowledge stored in the head of the reader as the building materials, much as individual bricks are combined in the construction of a house.
- 3) Attention, which is the cognitive energy or capacity required for mental processing tasks such as decoding and comprehending, but is limited in quantity.

It is assumed in the theory that getting meaning from printed words involves a two-step process: first, the printed words must be decoded; second, the decoded words must be comprehended. Whereas one may go from print to the meanings of individual words automatically, the acts of integrating, relating and combining these meanings in the unique ways demanded by sentences are required for comprehension (Samuels, 2013, p. 702).

As beginners or disfluent readers are not automatic in decoding (word recognition), they have to devote significant portions of their finite cognitive resources to that task with less left for comprehension; thus, the dual tasks of decoding and comprehension may exceed their attention capacity. This could lead to the situation that they overload or overextend their cognitive capacity to attend to the meaning of what they are reading so that the comprehension task is not getting done. Since the end product of reading is comprehension, the beginning reader is faced with a formidable problem (*ibid.*, p. 702). As shown in Figure 3-5, they put their attention on the decoding task first and then switch attention to comprehension to understand what they have decoded. Thus, although the beginning reader is able to comprehend by switching attention back and forth in this way, the process is slow, laborious and frustrating.

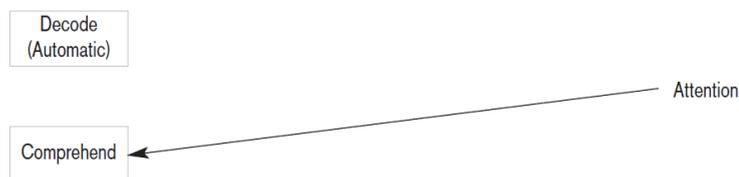
Fluent or more experienced readers, on the other hand, are able to automatically decode the words which frees up their cognitive resources and enables them to attend to and comprehend what they are reading (Figure 3-5). As mentioned in the previous section, automaticity has an effortless property: if two tasks can be performed simultaneously, at least one of them is being done automatically. The reading process of fluent readers is a clear manifestation of this property.

A. Beginning Reading



In beginning reading, attention is switched alternately from decoding to comprehension. Only one task can be done at a time.

B. Fluent Reading



In fluent reading, decoding is done automatically and attention remains on comprehension. Both tasks get done at the same time.

Figure 3-5. A model of attention and automaticity in reading proposed by Samuels (2013), reproduced with permissions from the author; copyright 2013 International Reading Association.

In translation activities, effort or attentional resource is generally considered to be distributed to reading, writing or coordination (cf. effort model in Gile, 1995). Although coordination efforts are required for task switching and language switching (Whyatt et al., 2016), such efforts or attentional resources are distributed, essentially, to either the ST or the TT when considering the virtual location of these resources. In other words, it could be difficult to tell whether a certain amount of attention allocated to the ST is for comprehension or for coordination; the same case also applies to the TT. Thus, from the perspective of attention allocation, the current discussion takes the subprocesses of reading comprehension and text production in translation as dual tasks that resemble the two-step process of decoding and comprehension in a reading activity. Meanwhile, the resources devoted to reading comprehension and text production in translation also incorporate the coordination effort.

Whereas the end product of reading is comprehension, the end product of translation is the typed-out target text. For a reading activity, the first step of decoding serves as the basis for comprehension; similarly, for a translation activity, reading comprehension (in the two-step process of translation) serves as the basis for text production. Thus, in theory, if the reading comprehension step is automated sufficiently (i.e., requiring considerably less attention resources) during the translation process, then more attention resources could be devoted to the final step of the activity, i.e., text production. This constitutes one of the hypotheses about automaticity regarding the attention allocation pattern in the present study, i.e., a more automated translation process would feature a lower proportion of attention on ST comprehension and correspondingly, a higher proportion of attention on TT production. Thus, together with the other three properties of automaticity mentioned earlier, automaticity in translation is examined in the current study from four perspectives: speed, effortlessness (the ease of processing), parallel processing degree and attention allocation to reading comprehension/text production.

3.3.3 Cognitive rhythm

Rhythm, which can be considered as an organising principle that structures events, is omnipresent in human activities, especially in language and speech (Magne et al., 2005). Broadly speaking, the way translators allocate their attention and coordinate various tasks manifested in their processing patterns can be delineated as cognitive rhythm

(Whyatt et al., 2016, p. 184). At the macro-level, cognitive rhythm in translation includes the amount of time allocated to various phases (e.g., orientation, drafting and revision phases defined in Jakobsen, 2002), or various subprocesses such as reading, writing and searching for information; at the micro-level, it can be linked to text typing and pausing behaviours (*ibid.*, p. 184). Notions such as segmentation (Jakobsen, 2003), production units (Carl & Dragsted, 2012), or translation units (Buchweitz & Alves, 2006) are often closely related to cognitive rhythm at the micro-level. Such notions normally refer to flows of successive text typing of varying size which are flanked by predefined pauses. To this end, cognitive rhythm in translation concerns the aspect of the “bursts of creativity in between pauses” (Saldanha & O’Brien, 2014, p. 112), or the alternating phases of pausing/production (Schilperoord, 1996, p. 5). One of the underpinnings behind such an idea is that translation is envisioned as a process where problem-solving (e.g., long pauses) and the uneventful or unmarked processing (typing flow) are clearly distinguished (Muñoz Martín & Martín de Leon, 2018).

Both the idea of dividing the translation process into different clear-cut stages (i.e., an aspect of cognitive rhythm at the macro-level) and the idea of taking the translation process as an orderly picture comprising of sharply different default procedures and problem-solving activities (i.e., an aspect of cognitive rhythm at the micro-level) have received criticism in several ways (e.g., Carl et al., 2011; Muñoz Martín, 2014; Muñoz Martín & Martín de Leon, 2018). Nevertheless, such features of text production, either at the macro- or micro-level, may yield recursive patterns of translation that are worth exploring. Drawing upon arguments that cognitive activities change with the unfolding interaction of the translator with the text and the environment, Muñoz Martín and Martín de Leon (2018, p. 45) adopted a novel approach for determining the pause thresholds individually and found “an uneven continuum of intermittent, alternating subtasks, rather than a binary automaton switching between two processing modes”. To this end, it would be of interest to gauge to what extent cognitive rhythm or the tempo translators structure and organise different subprocesses during translation can be regulated by constraints of time and their experience level. The present study adopts a similar methodology based on Muñoz Martín and Martín de Leon’s (2018) work in regard to pausing and segmentation in text production and the details about the methods are presented in section 4.2.3.1.

In writing research, Bereiter and Scardamalia (1987) proposed two models of composing, namely the Knowledge Transforming and Knowledge Telling models. The

Knowledge Transforming strategy features a more problematic or effortful process of writing with which the writer needs to constantly reprocess and transform existing knowledge in order to solve novel problems (*ibid.*, p. 11). In contrast, the Knowledge Telling strategy makes writing a natural and unproblematic task with which the writer makes “maximum use of existing cognitive structures and minimize the extent of novel problems that must be solved” (*ibid.*, p. 5). These two models thus exhibit two different processing modes in terms of cognitive rhythm. Schilperoord (1996) verified several properties that could characterise the Knowledge Telling model in writing process, such as the lack of a correlation between start-up times and text length/production time, rare online revisions, linear production and stable pause/production ratio. Basically, processing with the Knowledge Telling strategy is likely to proceed with a more steady and fixed rhythm, producing the text linearly and with rare online revisions, while processing with the Knowledge Transforming strategy could feature the opposite of such properties. As mentioned in section 2.1.2.5, Jensen (2000) applied these two models to examine the processing styles of translation displayed by different groups of translators. Such models could be useful in explaining the observed phenomena regarding how time pressure affects cognitive rhythm in translation process.

3.4 Summary

This chapter establishes a theoretical framework for the present study by examining and illustrating relevant theoretical models and concepts, firstly, on time pressure, stress, coping and performance, secondly, on the framework of Cognitive Load Theory developed in learning sciences, and thirdly, on the constructs of automaticity and cognitive rhythm. The perceptual nature of time pressure implies its cognitive relevance for its origin: it is usually the perception of shortage of time that makes the difference in behaviours and performance. Typically, time pressure affects the patterns of cognitive strategy selection in decision making tasks, and approaches to studying such strategy change have been investigated at a macro-level, a micro-level and in a combination of the macro- and micro- levels. A control model of stress regulation proposed by Maule and Hockey (1993) is an approach integrating the macro- and micro-level cognitive strategy selection patterns. The use of a direct control by changing the cognitive state which involves increased effort in this model is of particular interest when looking into the time pressure effects in translation activities.

Time pressure is a form of stress, and the transactional theory of stress (Lazarus, 1966, 1976; Lazarus & Folkman, 1984) provides an interactional approach to understanding and studying stress. Stress is considered as a product of the transactional system comprising two critical processes, i.e., cognitive appraisal and coping, that mediate the person-environment relationship. Cognitive appraisals determine whether a situation is stressful or not and (emotion-focused and problem-focused) coping is concerned with the way in which an individual responds to it. The problem-focused coping in the transactional model, like the direct control of stress by investing more cognitive resources in the control model of stress regulation, represents a more active means of dealing with the adverse stressful environment. The construct of effort is thus used to describe such a way of responding to the demands in a time-pressured or stressed environment. Hypotheses about the effect of stress on task performance have differed. The positive, negative and the inverted-U shape theories are proposed and are backed up by different theoretical accounts. This implies that the relationship between stress and performance could be extremely complex and may vary across different situations and contexts.

Cognitive Load Theory is built upon the human cognitive architecture which consists of long-term memory and working memory. Working memory limitations profoundly influence the character of human information processing and, to a considerable extent, shape human cognitive architecture (Sweller, 2003). Three types of cognitive load are distinguished in CLT, namely intrinsic, extraneous and germane load. Given the limitation of processing capacity, central to instructional design for effective learning is to reduce the (ineffective and task-irrelevant) extraneous load and increase the (effective and task-relevant) germane load in a task. While certain instructional procedures can effectively lower the extraneous load and facilitate schema construction for novices, they can burden the processing capacity of experts, since the schemas have already been stored in the experts' long-term memory and there will be an overlap between the schema-based and the redundant instruction-based components of guidance. Processing such redundant components will require additional working memory resources and might cause a cognitive overload. Thus, the same instructional procedure may have different effects on novices and experts.

Complex skills gradually evolve towards the direction of becoming more automated and requiring fewer processing resources. Typical properties or criteria of automaticity in cognitive activities include high speed, effortlessness, autonomy and

unconsciousness. However, these properties may not always be bundled together in reality. Automaticity can inevitably modulate attention, which has been researched extensively in reading activities. Generally, beginners have to switch attention between decoding and comprehension in sequence, so that considerable resources are spent on the decoding process leading to insufficient resources left for comprehension; by contrast, the decoding process is automated for fluent readers and thus the majority of their attention focuses on comprehension which is the final product of reading. This pattern of attention allocation regulated by automaticity may be applied to translation, a similar two-step process which involves reading comprehension and text production. It is hypothesised that generally, higher automaticity level in translation could lead to less attention devoted to comprehension leaving more attention to text production which is the final product of a translation activity.

Core to the notion of cognitive rhythm in translation, either at the macro- or micro-level, are the temporal features that are related to the production process. The distribution of cognitive activities during translation could be the result of the interplay between the translator, text and the environment, and may be more varied and complex than a binary model of alternating phases of pausing/production. Thus, the effects of time pressure and experience on cognitive rhythm at different levels and from various aspects are worth investigating. In addition, models such as Knowledge Telling and Knowledge Transforming in writing process could provide more theoretical insights into the observed behaviours in translation.

Chapter 4. Research Methodology

This chapter presents three aspects of the research methodology, namely, the manipulation of time pressure, the measurement of time pressure, effort/cognitive load and translation behaviours, and the assessment of translation quality. Since time pressure bears a perceptual nature, the manipulation of time pressure in the experiment must concern both objective time constraints and a subjective appraisal of time shortage. These strategies are presented in section 4.1. A range of subjective, physiological, behavioural and analytical methods are employed for measuring time pressure and effort/cognitive load as well as translation behaviours, which are illustrated in section 4.2. These measures include two self-report instruments, several physiological indices, a variety of eye tracking and keystroke logging metrics and the application of CLT as an analytical tool for estimating cognitive load. Section 4.1 and part of section 4.2 are based on a preliminary work (Weng & Zheng, 2020) from the present study. In addition, two translation quality assessment systems, the Multidimensional Quality Metrics (MQM, Lommel et al., 2013) and the Functional-Componential Evaluation (FCE, Colina, 2008, 2009), used in the present study are demonstrated in section 4.3. Finally, section 4.4 summarises the whole chapter.

4.1 Manipulation of time pressure

To investigate the effect of time pressure on human behaviour or task performance, researchers first need to ensure the existence of time pressure. This makes the experimental manipulation of time pressure inducement necessary. Only by properly manipulating time pressure can researchers be reasonably assured that the observed changes in task performance are indeed related to the variation of time pressure. The notion of time pressure involves not only the objective time constraints, but also “the feeling of time pressure, where individuals perceive the time available to them to be insufficient” (Rattat et al., 2018, p. 81). So, it is necessary to differentiate between the actual (objective) and the perceived (subjective) time availability because they all affect the degree of time pressure.

This distinction between the two origins of time pressure underlies the availability of manipulating both the objective time constraint and the subjective time perception to induce or intensify time pressure in an experiment. As mentioned earlier, Zakay (1993, p. 66) pointed out that time pressure occurs at the prospective paradigm of time perception when participants are “aware of and occupied with the passage of time”.

Therefore, in studies with deadlines investigating time pressure effects on task performance, the attentional resources are allocated to task processing and monitoring the passing of time simultaneously. If more resources are devoted to monitoring the passing of time, less attention is left for performing the task, since the attentional resources are limited. Time pressure can therefore be induced by constraining the timeframe, increasing the task demand or intensifying the participants' perception of time pressure by, for example, enhancing their awareness of the passing of time.

4.1.1 Categorisation of time pressure inducement strategies

Rendón-Vélez et al. (2016) provided a summary of strategies that researchers had used to induce time pressure in general laboratory settings. Their work shows that time pressure can be induced by controlling the amount of time available and the participants' motivation. The amount of time available can be manipulated objectively (by implementing a fixed or an individual/adaptive time constraint) and subjectively (by simply giving instructions such as "you must work quickly"). Motivation can be manipulated by designing tasks that engage the participants psychologically and stimulate their performance. These can be realised, for instance, by creating a gaming or competition situation, or implementing external cues such as a visual/audio feedback in the task.

A slightly different categorisation of these strategies is proposed based on the rationales of time pressure inducement. Time pressure occurs when there is an imbalance between the task demand and the amount of time required. Thus, time pressure can be induced by increasing the actual/perceived task demand, by reducing the actual/perceived time available, or by improving the participants' awareness of time passing, which is also a way of increasing the perceived total task demand by its very nature. Table 4-1 below illustrates the relationship between these dimensions and lists the corresponding strategies of time pressure inducement.

Table 4-1. Strategies of time pressure inducement

	<i>Objective Strategies</i>	<i>Subjective Strategies</i>
Time availability ↓	<i>Deadlines:</i> shorten the time available	<i>Instruction:</i> reduce the perceived time availability
Task demand ↑	Increase the complexity of the actual task	<i>Motivation and Visualisation:</i> increase the perceived task demand

Firstly, time availability can be reduced to elicit time pressure when the task complexity is consistent; conversely, when time availability is consistent, task complexity can be increased to intensify time pressure. These two strategies belong to the objective manipulation of time pressure. Secondly, to reduce the perceived time availability, researchers can simply inform the participants in the pre-task *Instruction* that the time given for the task is shorter than is normally required. Thirdly, one way to increase the perceived task demand is to motivate the participants as much as possible (e.g., providing incentives or making it a competition); the other way is to insert visual/audio feedback during the task, which can effectively improve the participants' awareness of the passing of time by constantly reminding them how much time is left. The present study adopts the objective strategy of imposing deadlines and subjective strategies of giving deliberately designed *instructions*, scoring the translation products to motivate the participants (*motivation*) and inserting a visible stopwatch/timer as a visual feedback (*visualisation*) in the task. The following sections illustrate the details of how to operationalise each of these aspects.

4.1.2 Objective strategies

As shown in Table 4-1, in theory, the two objective manipulation strategies are to either shorten the timeframe or render the task more complex. However, increasing the complexity of a task seems to shift the focus of the study to another variable – task complexity – rather than remaining on time pressure. In practice, complicating a task is not as feasible as imposing a tight deadline on a given task for the study of time pressure effects. Therefore, this study adopts the objective strategy of imposing a deadline to induce time pressure. Although it seems self-evident that constraining the timeframe for a task straightforwardly triggers time pressure, there are difficulties in establishing how much time to constrain in order to observe the effects of time pressure on task performance. Bayer-Hohenwarter (2009) identified two types of deadlines that have

been applied in previous empirical translation research, i.e., the fixed and individual/adaptive deadlines. Most of the previous translation research on time pressure (e.g., Jensen, 2000; De Rooze, 2003; Kourouni, 2012) employed a *fixed deadline* to induce time pressure in the experiment. They would identify an appropriate time interval from a pre-test or a pilot study and then assume that a reasonable fraction of that time interval would naturally impose time pressure on the participants. Below is an example taken from Jensen (2000):

Two professional translators took part in the pilot studies and I subsequently assumed that non-professional translators, who have no training or routine in translation, would need more time than professionals. The texts averaged about 120 words in length, and 30 minutes was chosen as a time interval that would be sufficient time for both groups. Similarly, it was expected that all informants would feel pressed for time when it was reduced to 10 minutes. (p. 81)

Another method used in some studies is the *individual deadline*, which means that the deadlines applied to each participant are individualised in line with their own translating speed. An example from Hansen (2006b) illustrates this:

The students were asked to translate texts of different degrees of difficulty using Translog at home under normal working conditions. Using the log files of these translations, I could register the period of time they had used for translation of 10 lines when not using dictionaries or the Internet. Based on that information, I calculated their average translating speed. My pilot experiments had shown that they needed some time for orientation – which I gave them in relation to the number of lines they had to translate. (p. 73)

Translation scholars (e.g., Bayer-Hohenwarter, 2009; Hansen, 2006b) argue that individual deadline is preferred in empirical translation research because it takes individual difference, the individual's *translating speed*, into account. The time pressure effects would in principle be captured at a higher level of efficacy by this method. In practice, however, it is only possible to test a participant's *translating speed* for a specific text, and this speed is not necessarily representative for the *overall translating speed* of the same person at all times. In this respect, an individual's *translating speed* in a pre-test can only represent their speed for that specific task; determining the timeframe for the subsequent experimental task based on this pre-tested

translating speed may not fully guarantee the inducement of time pressure as well, because the text has to be changed in the experimental task. Consequently, so far as inducing time pressure is concerned, the individual deadline may not necessarily function better than a fixed deadline.

Considering these factors, the present study adopted a fixed deadline as an objective strategy together with three subjective strategies (detailed in the following section) to induce time pressure for the time constrained tasks. The deadlines (for the highly and moderately time-constrained tasks) were determined as quartile one and the average of a dataset obtained from a pre-test. More details of the deadlines and the pre-test are presented in section 5.3.

4.1.3 Subjective strategies

As shown in Table 4-1, three supplementary strategies – *instruction*, *motivation* and *visualisation* – that could help elicit time pressure by decreasing the perceived time availability or by increasing the perceived task demand are identified and applied in the present study.

Instruction refers to the deliberately designed pre-task guidance which can make the participants feel pressured by decreasing the perceived time availability. An extreme instance of such a manipulation is presented in De Dreu (2003), a study that focused on the effect of time pressure on negotiation performance, in which *instruction* turned out to be a successful manipulation:

Participants in both experiments were given the same amount of time for the negotiation, regardless of experimental conditions. However, in the high time pressure conditions, participants were led to believe that this amount of time was relatively tight, while those in the low time pressure conditions were led to believe that this same amount of time was more than enough to reach agreement. (p. 282)

As mentioned before, different deadlines were employed in the present study as an objective manipulation strategy to induce time pressure. In addition, an *instruction* to intensify time pressure was designed (see Appendix 4) to make the participants aware that for the Short session, the time available to complete the task is less than is normally required and for the Standard session, the time given is the average time interval that is

normally needed. This could possibly reduce the perceived time availability to the participants to complete the task especially in the Short task session.

Motivation refers to the strategy that can stimulate the participants to be more psychologically engaged in a task. This can consequently increase the perceived task demand within a certain timeframe. Examples in behavioural studies using this strategy for time pressure manipulation are Young et al. (2012) and Brown and Miller (2000), both of whom informed the participants that their rewards (money and a raffle prize) would depend on the quality of their task performance (i.e., decision-making). In the present study, the participants were informed that their translation product would be assessed and marked by an expert after the experiment, and they would receive their feedback later based on their performance. This would to some extent increase their level of engagement, as well as the perceived task demand.

Visualisation refers to external visual feedback, namely, the numeric displays of time or progress bars that visually represent the amount of time that has elapsed. The underlying mechanism, as mentioned earlier, is that attempting to monitor the passing of time consumes the limited attentional resources and increases the total task demand. This strategy is well illustrated in Maule et al. (2000) and Furlan et al. (2016) in which a countdown was displayed to remind the participants of the time left for the decision-making or problem-solving tasks. In the present study, the passing of time was visualised by a countdown timer with a progress bar presented at the bottom of the screen (see Appendix 8). In doing so, the participants were expected to pay more attention to monitoring the passing of time. As a result, time pressure would be intensified.

In summary, the means of time pressure inducement by experimental manipulation are expanded from the mere imposition of time constraints to strategies that can subjectively enhance an individual's perception of time pressure. By applying both the objective and subjective strategies, time pressure could be optimally induced in the experiment.

4.2 Measuring time pressure, effort/cognitive load and translation behaviours

The approaches to measuring time pressure are identified based mainly on the physiological and psychological responses to stress. The response-based definition of stress implies a biological approach to studying stress. The sensitivity of some

biomarkers to the activation of the sympathetic nervous system justifies their use as physiological indicators of stress or time pressure. In addition, self-report questionnaires and psychometric instruments constitute a straightforward approach to acquiring individuals' subjective feelings about time pressure.

Following the accounts about effort/cognitive load assessment put forward in Paas & Van Merriënboer (1994), Paas, Tuovinen, et al. (2003) and Seeber (2013), the present study employed subjective, physiological, behavioural and analytical methods to assess the effort/cognitive load in translation tasks. As mentioned in section 3.2.5, although the construct of *effort* in the present study (corresponding to problem-based coping and the direct control of stress by recruiting more resources, i.e., changing the cognitive state) converges closely to the construct of *overall cognitive load* in CLT, they are of an inclusion relation in the present study: not all the subjective, physiological and behavioural measures of *effort* can be equally used as measures of *overall cognitive load*, because of the inherent relationships between different types of cognitive load and performance assumed in CLT. This means that only those *effort* measures that meet the assumptions in CLT can be taken as valid *overall cognitive load* measures in the present study. The operationalisation of this distinction is discussed in detail in section 4.2.4.

The behavioural metrics presented in 4.2.3 are mainly for measuring certain aspects of translation behaviours especially with respect to automaticity and cognitive rhythm. Overall, in order to produce reliable results, a multi-method approach triangulating different types of data was adopted in the present study. This section presents the four types of methods, namely, subjective (4.2.1), physiological (4.2.2), behavioural (4.2.3) and analytical (4.2.4), that have been used for measuring time pressure, effort/cognitive load and translation behaviours.

4.2.1 Subjective methods

4.2.1.1 Measurement of time pressure using subjective methods

Subjective methods constitute an important type of measure for both time pressure and effort level. One of the most commonly used and effective approaches for assessing time pressure levels is to administer self-report psychometric instruments. As the emotional state of anxiety is “usually regarded as a product of stress” (Dobson, 1982,

p. 9), measuring how anxious an individual is in a certain situation can reflect their time pressure level at that moment. This legitimises the measurement of anxiety as a psychological indicator of time pressure. The current study employs one of the most widely used psychometric instruments of anxiety, i.e., the State–Trait Anxiety Inventory (STAI) developed by Spielberger (1972), as one of the subjective methods for time pressure measurement. This instrument comprises two components, namely, the State and the Trait Anxiety Scales. The State Anxiety Scale evaluates the current state of anxiety, particularly in the face of a threatening event, by asking how respondents feel “right now”. The items mainly measure subjective feelings of apprehension, tension, nervousness, worry and activation/arousal of the autonomic nervous system. The Trait Anxiety Scale evaluates relatively stable aspects of their anxiety proneness or baseline, including general states of calmness, confidence and security. By involving the baseline anxiety level in the analysis, this instrument can take individual differences into account. STAI is relatively brief to administer and does not require costly or time-consuming scoring or interpretation procedures. Therefore, this measure lends itself well to general use in various research fields and has been proved a useful and reliable psychometric tool (Julian, 2011). In the present study, the State Anxiety form was completed after each task session, and the Trait Anxiety form was completed after the whole experiment as recommended in the instrument’s instructions.

In addition, the method of a retrospective questionnaire or interview is an efficient alternative for acquiring the participants’ psychological experience of time pressure. As indicated by Minkel and Phillips (2015), measuring the subjective responses to probes into emotion is essential for validating that the chosen stimuli work as intended; this is because the core feature of an emotion is the participant’s subjective feeling associated with it and there is no substitute for asking participants to report their own subjective experiences. In the field of translation process research, such questionnaires have been applied in a few studies for the purpose of time pressure verification (e.g., Hansen, 2002a, 2002b; Kourouni, 2012). The present study examined participants’ subjective feelings on *Stress* and *Temporal Demand* respectively by two self-reported questions adapted from the NASA Task Load Index (NASA TLX) questionnaire. The question on *Temporal Demand* inquires how much time pressure the participants felt during the corresponding task and the question on *Stress* (adapted from the original question of NASA TLX on *Frustration*) inquires how stressed they were during the corresponding

task. More details about the application of NASA TLX in the present study are presented in the following section.

4.2.1.2 Measurement of effort using subjective methods

Subjective rating scales are widely used in assessing effort in cognitive tasks (Paas, Tuovinen, et al. 2003; Paas, 1992; Skulmowski & Rey, 2017). The use of subjective methods is based on the assumption that “people are able to introspect on their cognitive processes and to report the amount of mental effort expended” (Paas, Tuovinen, et al. 2003, p. 66). Both multidimensional and unidimensional scales have been used in a variety of studies (e.g., Paas & Van Merriënboer, 1994; Paas et al., 1994). One of the widely adopted measures in translation process research is NASA TLX which is a multidimensional scale assessing a group of associated variables including Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration levels. An example of its application is Sun and Shreve (2014) and Sun (2015) who employed and adapted this scale to obtain participants’ subjective workload data in order to assess translation difficulty and it was proved to be reliable.

As mentioned, the current study adopted NASA TLX as a subjective method to assess the effort that participants invested in each translation task. Specifically, four dimensions out of six in the original version of the questionnaire were included in the research to fit the purpose of this study. Table 4-2 below illustrates the comparison between the items in the original version of NASA TLX and the ones used in the current study.

Table 4-2. NASA TLX item definitions and questions

<i>NASA TLX item definitions</i>	<i>NASA TLX questions</i>	<i>Current study</i>
<p>Mental Demand (low/high)</p> <p>How much mental and perceptual activity was required (for example, thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, forgiving or exacting?</p>	<p>Mental Demand</p> <p>How much mental and perceptual activity did you spend for this task?</p>	✓
<p>Physical Demand (low/high)</p> <p>How much physical activity was required (for example, pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?</p>	<p>Physical Demand</p> <p>How much physical activity did you spend for this task?</p>	✗
<p>Temporal Demand (low/high)</p> <p>How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?</p>	<p>Temporal Demand</p> <p>How much time pressure did you feel in order to complete this task?</p>	✓
<p>Performance (good/poor)</p> <p>How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?</p>	<p>Performance</p> <p>How successful do you think you were in accomplishing the goals of the task?</p>	✗
<p>Effort (low/high)</p> <p>How hard did you have to work (mentally and physically) to accomplish your level of performance?</p>	<p>Effort</p> <p>How hard did you have to work to accomplish your level of performance?</p>	✓
<p>Frustration Level (low/high)</p> <p>How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?</p>	<p>Frustration Level</p> <p>How insecure, discouraged, irritated, stressed and annoyed were you during the task?</p>	<p>Stress</p> <p>How stressed were you during this task?</p>

The question on *Effort* served as a subjective measure of effort invested in the translation task in the present study. The question on *Temporal Demand* was used to assess participants' subjective feelings of time pressure. The question on *Frustration Level* was further adapted to focus on the aspect of *Stress* in order to examine the association between time pressure and stress. The question on *Mental Demand* was used to assess the intrinsic difficulty or the subjective perception of complexity of the

translation task/material. The two dimensions on *Physical Demand* and *Performance* were eliminated as they are not closely related to the purpose of this study.

Originally, administering the NASA TLX involves two steps. First, participants reflect on the task they are asked to perform and look at each paired combination of the six dimensions to decide which is more related to their personal definition of workload. This results in 15 paired comparisons. The second step involves participants rating each of the six dimensions on scales from Low to High. The raw score for each of the six items is multiplied by the weight from step one to generate the overall workload score per task. In practice, however, many researchers skip the weighting step (see Hart, 2006), which reduces the amount of time needed to administer the scales, and analyse the raw responses only. The current study adopted the raw score calculation method without the weighting step to avoid an excessively long experiment process.

The rating scale of NASA TLX is essentially a line with 21 marks (Figure 4-1). The method of calculating the score is: the number of lines a participant marked subtracting one, and multiplying by five. For example, in Figure 4-1, the score of the item would be: $(5-1) * 5=20$.



Figure 4-1. Rating scale of NASA TLX

4.2.2 Physiological methods

As a type of stress, time pressure could lead to a range of physiological responses. Bayer-Hohenwarter (2009) argued that measuring such physiological processes triggered by time pressure could yield more reliable results than other time pressure measurement approaches. Generally, the physiological responses to stress span the hypothalamic–pituitary–adrenal axis effects (including the adrenocorticotrophic hormone, cortisol, vasopressin and dehydroepiandrosterone), the immune system effects, the sympathetic-adrenal-medullary system effects, the cardiovascular effects and the brain–gut axis effects (Allen et al., 2014). Although assessing most of the stress-related biomarkers still remains relatively invasive, requiring, for example, taking blood samples from the individual, some of the other biomarkers have become easier to assess over time (Slavich et al., 2019). For instance, portable research-oriented

smartwatches can continuously and accurately monitor the biomarkers of heart rate, galvanic skin response and skin temperature, which are related to the activation of the sympathetic nervous system. Such technological innovations open up valuable channels for studies requiring high ecological validity in human experimentation.

The use of physiological methods for estimating effort is based on the assumption that changes in cognitive functioning are reflected by physiological variables, such as heart, brain and eye activities (Paas, Tuovinen, et al., 2003). These activities can reflect the arousal state of a person. Since most of these biomarkers can be continuously recorded, these methods allow for a moment-to-moment analysis of events. This is important because cognitive load is assumed to fluctuate locally in language processing, and the individual operations inherent in language processing are performed very rapidly (Mitchell, 2004, as cited in Seeber, 2013, p. 25). The continuously recorded biomarkers offer data that can be precisely analysed in fractions of seconds.

Despite the objective nature of physiological methods, their main drawback is “the difficulty of identifying and determining what is actually measured” (Seeber, 2013, p. 25). Although some brain activity techniques (e.g., fMRI) may constitute relatively direct physiological methods for cognitive load measurement, most other physiological measures only have an indirect causal link to the intensity or activation state of cognitive processing. For instance, “high-cognitive load may lead to high stress in an individual, which may lead to changes in heart rate – as may the individual’s emotional response to the learning materials” (Brünken et al., 2003, p. 56). In this regard, the underlying assumption made for the use of physiological methods in assessing both time pressure and intensity of cognitive processing is essentially on a similar ground – both originate from a human being’s physiological responses or the arousal state resulting from the activation of the sympathetic nervous system.

Consequently, the interplay between time pressure and the intensive cognitive processing expended during translation in the current study may lead to an intricate complex of responses reflected in the physiological arousal state. It can be expected that some of the biomarkers may be more sensitive to the level of time pressure experienced by the participants, while some others may be prone to reflect the intensity of cognitive processing. Thus, through triangulation of different types of methods (e.g., subjective and behavioural methods) and statistical analysis such as testing the correlations between different types of measures, the validity of the physiological methods for measuring either time pressure or effort can be ascertained. The following

subsection illustrates the details of all the physiological measures employed in this study as potential measures for time pressure and/or effort.

4.2.2.1 Physiological measures used in this study

Generally, four requirements should be considered when evaluating the physiological methods used in the experiment, namely, noise-resistance, non-invasiveness, temporal resolution and affordability (Seeber, 2013). With an attempt to create a natural scenario that translators may experience in their workplace, a smartwatch, i.e., an Empatica E4 Wristband, was used to continuously record several biomarkers (i.e., heart rate, heart rate variability, galvanic skin response and skin temperature) during the experiment which could minimize the interruption or influence from external devices; an Omron M7 Intelli IT automatic upper arm blood pressure monitor was used for blood pressure measurement immediately after each task session; and an eye tracker (either Tobii TX300 or Tobii Pro X3-120) was used for measuring pupil size as well as collecting other eye movement data during translation. Details about each physiological measure are demonstrated below.

Heart rate (HR) and blood pressure (BP)

Heart rate arises from the integration of an individual's intrinsic rate with the influences of Sympathetic Nervous System (SNS) and Parasympathetic Nervous System (PSNS), which are the two divisions of the Autonomic Nervous System (ANS). In a stressful situation, our body produces a surge of stress hormones (adrenaline and cortisol) into the blood to prepare the body for the "fight-or-flight" response, which boosts the body's alertness and activates the SNS, causing the blood vessels to narrow, engendering an increase in blood pressure and accelerating the heartbeat (see Taelman et al., 2009; Akselrod et al., 1981; Sherwood & Carels, 2007). The sensitivity of heart rate to the activation of the SNS provides the justification for its use as an indicator of psychological stress (Jennings, 2007). The HR and BP indices that were employed in the current study were the mean value of HR, Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP) and Mean Arterial Pressure (MAP).

Heart rate variability (HRV)

HRV refers to the variation in the time interval between consecutive heartbeats. The time between beats measured in milliseconds is called the R-R interval or inter-beat interval. When the artifacts are excluded, the intervals are referred to as normal-to-normal intervals. HRV is also regulated by the ANS and manifests the interactions between ANS and the cardiovascular system. Decreased HRV, indicating a disturbed ANS function (Horsten et al., 1999), has been associated with mental processing or stress in a number of empirical studies (e.g., Myrtek et al., 1996; Sloan et al., 1994). There are two categories of HRV indices, the time-domain and frequency-domain measures (see Shaffer & Ginsberg, 2017 for a comprehensive review). The indices that were included in the current study are RMSSD (root mean square of successive normal-to-normal interval differences), SDNN (standard deviation of the normal-to-normal intervals) and PNN50 (percentage of successive normal-to-normal intervals that differ by more than 50 milliseconds), all of which are the most frequently used time-domain measures.

Galvanic skin response (GSR)

Galvanic skin response, or Electrodermal Activity, is the measurement of the continuous dynamic variations of the electrical properties of the skin and it originates from the autonomic activation of sweat glands in the skin, which is a function of SNS (Norman et al., 2016). Pijeira-Díaz (2018) indicated that the electrodermal system is the only one in the entire body solely innervated by the SNS, which reflects a state of activation. As such, GSR is considered as a well-validated, widely accepted and readily accessible measure of arousal (Critchley, 2002; Neumann & Blanton, 1970). There are two components of GSR data, namely the slowly varying tonic skin conductance level and the rapidly changing phasic Skin Conductance Response (SCR). Analysing these two components reveals a considerable amount of information and different characteristics about the arousal level (see Braithwaite et al., 2013 for more details). The indices used in the current study included the mean value of GSR, the mean value of SCR Amplitude and SCR Frequency (i.e., SCR number per minute).

Pupil size and skin temperature

The human pupil is also known to reflect the ANS activities in that mydriasis has been found to be related with mental effort exertion and psychological stress in a number of

studies (e.g., Yang & Kim, 2018; Pedrotti et al., 2014; Partala & Surakka, 2003; Ren et al., 2014). Acute stress can also trigger sympathetically mediated peripheral vasoconstriction which causes an obvious drop in skin temperature; in the meantime, this influx of peripheral blood, along with stress-induced thermogenesis increases core temperature (Herborn et al., 2015). Generally speaking, human faces easily get hot while the fingertips get cold under acute stress (Kataoka et al., 1998). There are plenty of studies that found skin temperature decreased under stress (e.g., Cho et al., 2017; Rickard, 2004; Hui & Sherratt, 2018). The indices employed in the current study for these two measures were the mean value of pupil size and the mean value of skin temperature.

Overall, based on existing studies, the measures of HR, BP and skin temperature are more likely to reflect stress levels whereas the measures of GSR, HRV and pupil size are more possible to reflect effort/cognitive load. The present study examined how different time constraints could modulate these measures. Table 4-3 presents the indices of the physiological measures, whether or not they are continuously recorded during the translation tasks and the devices used to record them.

Table 4-3. Summary of the physiological measures and corresponding indices

<i>Measures</i>	<i>Indices</i>	<i>Continuous</i>	<i>Device</i>
Heart rate (HR)	The mean value of heart rate	✓	Empatica E4
Blood pressure (BP)	Systolic blood pressure (SBP)	✗	Omron M7 Intelli IT
	Diastolic blood pressure (DBP)		
	*Mean arterial pressure (MAP)		
Pupil size	The mean value of pupil size	✓	Tobii TX300/ Tobii Pro X3-120
Skin temperature	The mean value of skin temperature	✓	Empatica E4
Galvanic Skin Response (GSR)	The mean value of the GSR raw data (GSR Mean)	✓	Empatica E4
	The mean Amplitude of Skin Conductance Response (SCR Amplitude)		
	The Frequency of Skin Conductance Response per minute (SCR Frequency)		
Heart Rate Variability (HRV)	Root Mean Square of Successive Differences between the normal-to-normal heartbeat intervals (RMSSD)	✓	Empatica E4
	Standard Deviation of the normal-to-normal heartbeat intervals (SDNN)		
	Percentage of successive normal-to-normal heartbeat intervals that differ by more than 50 milliseconds (PNN50)		
<i>Note:</i> *Mean arterial pressure is calculated by the formula [(systolic blood pressure + 2×diastolic blood pressure) ÷ 3]; the mean value of other measures (e.g., heart rate, pupil dilation, skin temperature and GSR) refers to the average of the continuously recorded data during a translation task session.			

4.2.3 Behavioural methods

Research has shown that acute psychological stress impairs the higher-order cognitive functions such as working memory (e.g., Qin et al., 2009). Changes in such aspects of the cognitive process may severely influence behaviours, as reflected by, for example, the eye-movement patterns and fixation duration (e.g., Henckens et al., 2009; Wang et al., 2010), and keystroke and mouse dynamics (e.g., Lim et al., 2014). It appears that time pressure or stress could give rise to behavioural changes in eye movements and keystroke dynamics owing to its influence on an individual's cognitive functions. To this end, certain behavioural metrics could indicate the intensity of effort exertion. In the present study, behavioural methods mainly refer to the techniques of eye tracking and keystroke logging which have been increasingly applied in translation process

research in recent years. These methods were used to gauge the cognitive processes by various eye and key related metrics in the present study, which are presented in the following sections.

4.2.3.1 Keystroke logging metrics

Keystroke logging is the electronic recording of all operations (including keypresses, editing functions and cursor movements) made by the writer as he or she types. A keystroke logging software allows the storage and subsequent retrieval of a large number of features of the writing activity. All the keys pressed during a task, such as text production characters, punctuation symbols, navigational characters (e.g., up and down, right and left arrows), editing and revision indicators (e.g., cut and paste, backspace and deletion), are recorded with corresponding timestamps. Thus, “the output of the logfile is a highly detailed record of the temporal features of the writing activity, which provides a rich source of online data for analysis” (Miller & Sullivan, 2006, p. 5). The keystroke logging data in the present study was recorded by its embedded function of the eye tracking software Tobii Studio/Tobii Pro Lab in combination with Translog II (Carl, 2012), a widely used keystroke logging software in translation process research.

Pausing, as one of the most important temporal aspects of language production, has attracted most attention from keystroke logging studies. Since there is always a time interval (inactivity) between two successive keystrokes in typing, all such inter-key intervals are possible candidates for pauses. In order to identify what kind of inter-key intervals should be counted as pauses, Wengelin (2006) gave a working definition and an interpretation for it:

“[...] a *pause* is a transition time between two keystrokes, which is longer than what can be expected to be necessary for the time needed to merely find the next key. Thus, in order for a transition to be considered as a pause, a writer has to “interrupt” her typing considerably longer than that “normal” transition time between two keystrokes.” (p. 111)

It is clear that certain criteria or thresholds should be stipulated to operationalise a pause in a specific study. Although pauses in text production may have multiple causes, they are generally assumed to be “behavioural reflections of the cognitive processes

involved in changing attentional states” (Schilperoord, 1996, p. 9). It follows that the longer the pausing time, the more the cognitive operations are required by the output (Butterworth, 1980, p. 156). Based on such underpinning assumptions, features such as the frequency, duration and positioning of pauses have been used to signal problems or difficulty, as well as increased effort in translation.

Whilst it is relatively widely agreed that intensive cognitive processing occurs at long pauses during the writing or translation process, there is little consensus about the threshold of a pause. In most cases in translation research, pause thresholds are tailored, ranging from one to five seconds or longer to best suit the aims of a research study (e.g., Alves & Vale, 2009; Immonen, 2011). It is also considered that the distribution of cognitive activities is changing over time and between different people, because of the unfolding interaction of the translator with the text and the environment (Muñoz Martín & Martín de Leon, 2018, p. 31). Thus, it is of great interest to take into account the dynamic feature of pauses when the pause threshold is determined.

Based on Dragsted (2005) in translation and Rosenqvist (2015) in writing research, Muñoz Martín and Martín de Leon (2018) and Muñoz Martín and Cardona Guerra (2019) described an approach to determining pause thresholds that can better account for the individual differences in typing between different participants and different task conditions. This approach distinguishes three levels of thresholds: 1) a baseline threshold at 200 milliseconds which can capture the minimal cognitive, perceptual and action operations; 2) a lower threshold at $[2 * \text{the median pause within words of a single task file}]$; and 3) an upper threshold at $[3 * \text{the median pause between words}]$ (Muñoz Martín & Cardona Guerra, 2019). By establishing the three levels of thresholds, pauses are classified into *Short Pauses* (SPs, between 200 milliseconds and the lower threshold), *Mid Pauses* (MPs, between the lower and the upper threshold) and *Long Pauses* (LPs, above the upper threshold). The collection of all the activities between two LPs is defined as a *task segment*. Muñoz Martín and Cardona Guerra (2019) suggested that SPs are mostly physical, mechanical or strategic behaviours related to keyboarding; MPs mostly hint at monitoring activities; and LPs flag problem-solving strategies as well as some other cognitive processes. The present study basically adopted this approach for pause analysis. Changes were only made to the notions of “pause within words” and “pause between words” in the definitions of the thresholds to fit the characteristics of typing flow in Chinese.

Since typing in (simplified) Chinese is usually done with the pinyin system, i.e., the Romanisation transcriptions of Mandarin, the typing procedures are different from typing in European languages. Typing in Chinese requires the typist to know the pinyin letters of the characters he or she wants to type out and to be able to distinguish them from the other characters that pop up as recommendations. When typing the commonly used characters or phrases, the input system is usually intelligent enough to make them regularly the first one of several recommendations. The input systems allow the typist to use the spacebar to select the first recommendation which could be one character or a combination of several characters depending on the pinyin letters that have been typed. This means that a confirmation key (i.e., usually the SPACE bar to select the first recommendation or the number representing the correct recommendation) has to be pressed to type out the chosen character(s). In most cases, especially for skilled typists, typing out a phrase (containing two or more Chinese characters) at one time is more common than typing out only one character each time. As shown in Figure 4-2 below, one can type out only one Chinese character at one time (upper in Figure 4-2), or more characters (such as a phrase) at one time (lower in Figure 4-2).



Figure 4-2. Typing Chinese characters using the Windows 10 pinyin input method

The length of each unit typed out in Chinese is dependent mostly on one’s typing habit, e.g., some may tend to type out one phrase at a time whilst others might prefer to type out a slightly longer unit each time before the final confirmation key is hit. Thus, the notion of “pause within/between words” in Muñoz Martín and Cardona Guerra’s (2019) original approach can hardly be applicable to the typing flow of Chinese if we take a Chinese “character” as an English “word”. As such, the most natural way of chunking the typing flow of the Chinese text production process is to distinguish the inter-key intervals within and between such *typing units* rather than within and between Chinese characters.

A *Typing Unit* (TU) in the present study is defined as a string of keystrokes starting with a pinyin letter keystroke, ending with the (first⁴) confirmation keystroke. *Intervals between TUs* (*IbTU*) are thus the time intervals starting with the offset of a TU (i.e., the confirmation keystroke), ending with the onset of the following TU (i.e., a pinyin letter keystroke). *Intervals within TUs* (*IwTU*) are the time intervals between any two successive keystrokes within a TU.

It should be noted that keystrokes within a TU are not necessarily all pinyin letters but could include functional keys (e.g., Deletions) as well. It happens quite often when the participant makes a typo in typing down the pinyin letters and revises it immediately before the end of the production of the current TU. Intervals between TUs are not always empty – punctuations and functional keystrokes could occur in these intervals. Muñoz Martín and Martín de Leon (2018) suggested that translators may carry out certain monitoring activities before punctuating a translation unit or a segment. Similarly, punctuation might be followed by planning activities preparing for the following stage of production. Thus, the interval from the offset of the first TU (a confirmation keystroke) until the onset of the next TU (starting with a letter keystroke) could be filled with (different) cognitive processing activities, no matter whether the punctuation or functional keystrokes are hit or not during this interval.

IbTUs and IwTUs were used in determining the pause thresholds in the present study instead of the “pause between/within words” to fit for the features of typing flow in Chinese. IbTUs and IwTUs were calculated individually based on each task file completed by each participant. Following Muñoz Martín and Cardona Guerra’s (2019) method, the lower threshold is $[2 * \text{the median IwTU of a translation task}]$; the upper threshold is $[3 * \text{the median IbTU of a translation task}]$. Consequently, those pauses between two successive keystrokes that are between 200 milliseconds and the lower threshold are called *Short Pauses*; those pauses between two successive keystrokes that are between the lower and the upper threshold of one task are called *Mid Pauses*; and those pauses between two successive keystrokes that are longer than the upper threshold are called *Long Pauses* (Figure 4-3). Similar to Muñoz Martín and Martín de Leon (2018), LPs are used to define *Segments*, which refer to the collection of any activities (including keystrokes and SPs/MPs) between two LPs.

⁴ There are situations that multiple confirmation keys are pressed to complete the selection of correct characters/phrases from the recommendations. In such cases, the first confirmation key is deemed as the end of a typing unit, because it signals an end of the virtual typing flow of a unit.



Figure 4-3. Thresholds of Short, Mid and Long Pauses

Apart from different types of pauses that are usually used to flag different cognitive activities in the translation process, Vieira (2016) reported that the measure of *Pause-Word-Ratio* (PWR, proposed originally by Lacruz & Shreve, 2014) had been a reliable indicator of cognitive effort in post-editing. In the present study, PWR was obtained by dividing the total number of LPs by the number of TUs produced during the task. Together with the pauses (i.e., SPs, MPs and LPs), TUs and Segments as defined above, different metrics featuring the typing or text production dynamics during translation were examined. Explanations and the calculation formulas of these metrics are presented below:

- *Keystroke production speed*⁵: [total number of keystrokes ÷ drafting⁶ and end revision duration (minutes) of a task].
- *Average IbTU*: average duration of Intervals between TUs.
- *Average IwTU*: average duration of Intervals within TUs.
- *TU length (keystrokes)*: average number of keystrokes in a TU.
- *TU time*: average time duration spent on a TU.
- *SP/MP/LP count and length*: total number and average duration of SPs/MPs/LPs.
- *Segment length (time)*: average duration of a segment.
- *Segment length (TU number)*: average number of TUs in a segment.
- *PWR*: [total number of LPs ÷ total number of TUs in a task].
- *TT word count*: total number of Chinese characters in the final product.
- *TU count*: total number of TUs produced during the whole translation process.

⁵ Keystroke production speed is distinguished from the notion of overall translation speed in that it concerns the stages in which typing activities are stretched (i.e., either producing or modifying the target text); thus, the orientation or planning stage without any typing activities is discarded from the formula.

⁶ Determination of the three translation phases, i.e., orientation, drafting and end revision, is presented in section 4.2.4.3.

The metrics of *Keystroke production speed*, *Average IbTU* and *Average IwTU* were used to measure the speed property of automaticity. It is assumed that a higher *keystroke production speed* along with shorter *Average IbTU* and *Average IwTU* features a higher automaticity level. *TU length* was examined to see whether time conditions could influence translators' habitual typing behaviours. The metric of *TU time* (along with the metric of *TU fixation* which is presented in the following section) was used to measure the effortlessness property of automaticity when processing each TU. It is assumed that processing a TU is effortless when the translator consumes little time and devotes little attention to it.

The metrics of *SP/MP/LP count and length*, and *Segment length* measured by time and TU number, were used to portray different aspects (such as occurrences of different micro- and macro-strategy behaviours indicated by MPs and LPs) of cognitive rhythm. *PWR* was used as a behavioural metric indicative of the level of effort exertion in a translation task. *TT word count* and *TU count* were used to depict the aspects of product quantity reflected in the end product and generated during the course of the translation process respectively.

4.2.3.2 Eye tracking metrics

Eye tracking is the technique which can record the point of gaze of a person and the movement of the eyes from one point to another (Saldanha & O'Brien 2014, p. 136). Eye fixation behaviours have been extensively researched since the 1970s regarding their connection with cognitive processes. For example, Just and Carpenter (1976, p. 441) found in a number of cognitive tasks that "the eye fixation behaviour is linked to a processing model for the task by assuming that the eye fixates the referent of the symbol being operated on". Rayner (1977) found that during reading, eye movements are affected by cognitive processes occurring at the time of the fixations. Later, the widespread eye-mind hypothesis was formulated by Just and Carpenter (1980, p. 331) which posits that "there is no appreciable lag between what is being fixated and what is being processed". Most studies in translation research that attempt to uncover the complex mental processes during translation with the help of the eye tracking technique are based on such an assumption.

Overall, the recording of eye movements can reveal, instantaneously, where the attention of a person focuses on and for how long this focus lasts. In the present study,

the Areas of Interests (AOIs) were drawn on the source text (ST), the target text (TT) and the timing (TM, i.e., the stopwatch/timer) areas on the screen (see Figure 4-4). The most commonly used features of eye tracking include fixation duration, fixation count and Mean Fixation Duration (MFD). Fixation duration and fixation count on an AOI represent the relative engagement with the objects in the AOI, while MFD (i.e., total fixation duration divided by fixation count) which typically ranges from 150 to 300ms (Tullis & Albert, 2013, p. 173), can represent the level of overall engagement with an object. That is, the greater the MFD of an AOI, the higher the level of overall engagement with the object in that AOI (*ibid.*, p. 173). For these properties of the eye tracking or fixation metrics, they are also widely used in relevant translation research as indicators of mental effort. For example, Hvelplund (2017, 2019) investigated how mental processing varied in translation drafting, revision and resource consultation, and found that digital consultation was associated with heavier processing than the other two types of activities as indicated by a longer MFD.



Figure 4-4. AOIs of *ST*, *TT* and *TM* drawn in the present study

These fixation related metrics were also applied in the current study to inquire into relevant research questions relating to patterns of attention distribution and effort exertion. In addition, some metrics integrating both the typing dynamics and the gazing behaviours were designated for certain purposes as illustrated below:

- *ST fixation duration percentage in TU*: [(fixation duration on ST when producing a TU ÷ fixation duration on ST and TT when producing a TU) × 100%].
- *ST-TT fixation ratio in TU*: [fixation duration on ST when producing a TU ÷ fixation duration on TT when producing a TU].
- *TU fixation*: average fixation duration on ST and TT when producing a TU.
- *Fixation duration percentage on ST/TT in the drafting phase*: fixation duration on ST/TT as a percentage of the total fixation duration on ST and TT during the drafting phase.
- *Fixation count on TM*: number of fixations on the timing area of the screen.
- *Mean fixation duration (MFD)*: [total fixation duration on the whole screen during a task ÷ total fixation count on the whole screen during a task].

ST fixation duration percentage in TU and *ST-TT fixation ratio in TU* were inspected as metrics to indicate the degree of parallel processing in translation, which is one aspect of the effortlessness property of automaticity. It is assumed that, with a high parallel processing degree, the translator could read the ST and produce the TT in a more concurrent manner. Thus, compared with the sequential processing of ST comprehension and TT production, a higher degree of parallel processing should feature a higher portion of attention paid to ST reading during the course of typing down the TUs, or formulating the TT. *TU fixation* (in addition to *TU time* as mentioned earlier) was also used to measure the effortlessness property of automaticity but in terms of the ease of processing each TU.

Fixation duration percentage on ST/TT in the drafting phase was for assessing how translators allocated their attention to the subprocesses of reading comprehension and text production in translation. The attention allocation pattern is also associated with automaticity in that in the two-step cognitive process of translation, if reading comprehension is highly automated, more attention or resources will naturally be allocated to text production. Thus, the evaluation of how visual attention is distributed to ST/TT can reveal, to a certain extent, the overall automaticity level in translation. The metric of *Fixation count on TM* was used to inspect how time constraints affected the attention paid to the passing of time. Lastly, the widely applied metric of *Mean fixation duration (MFD)* was used to indicate the intensity of cognitive processing (i.e., as a measure of *effort*) in a task.

4.2.3.3 Other behavioural metrics

Jakobsen's (2002) approach of explicitly dividing the translation process into three production phases, namely orientation, drafting and revision, has been influential and has inspired a variety of in-depth investigations into the subprocesses of translation in recent decades. With the use of the keystroke data, identification or definition of the three phases becomes more feasible. According to Jakobsen (2002), these three production phases are delimited as follows:

- *Orientation phase*: from the appearance of the source text on the screen to the typing of the first text production key.
- *Drafting phase*: from the first text production keystroke until the first typing of the final punctuation mark (or equivalent keystroke).
- *End revision phase*: From the end of the drafting phase until submission.

The present study adopted this method of identifying different stages of the translation process by involving the metrics of the time duration of the three phases. Moreover, the revision behaviours during the entire translation process were examined in both the drafting phase (termed as *online revision*) and the end revision phase (termed as *end revision*), and the revision behaviours were indicated by the number of deletion keystrokes of Chinese characters⁷. These additional behavioural metrics about translation phases and types of revisional behaviours are demonstrated as follows:

- *Orientation/Drafting/End revision duration and percentage*: time duration of each phase and time duration of each phase as a percentage of the total task time.
- *Online deletions and end revision deletions*: number of deletion keystrokes in the drafting phase and end revision phase respectively.

Overall, Table 4-4 summarises all the behavioural metrics that were used in the present study to explore automaticity, cognitive rhythm and effort exertion under the influence of time pressure.

⁷ Deletion keystrokes that are inserted within TUs (i.e., deleting the pinyin letters rather than the typed-out Chinese characters) are eliminated since they are most likely the corrections of typos which may not signal a meaningful revision of the produced target text; only the deletions of typed-out Chinese characters are seen as meaningful revision behaviours in the present study.

Table 4-4. Summary of the behavioural metrics

<i>Metric name</i>	<i>Aspects</i>	<i>Dimension</i>	
Keystroke production speed	Processing speed	Automaticity	
Average IwTU			
Average IbTU			
ST fixation duration percentage in TU	Parallel processing degree (effortlessness)		
ST-TT fixation ratio in TU			
TU time	Average TU effort (effortlessness)		
TU fixation			
Fixation duration percentage on ST/TT in the drafting phase	Attention allocation		
SP count & length	Pauses	Cognitive rhythm	
MP count & length			
LP count & length			
Segment length (time)	Segments		
Segment length (TU number)			
Orientation duration & percentage	Translation phases		
Draft duration & percentage			
End revision duration & percentage			
Online deletions	Types of revision behaviour		
End revision deletions			
Mean fixation duration (MFD)	Behavioural metrics		Metrics for <i>effort</i>
PWR			
TU length (keystrokes)	Habitual typing behaviour		Other
TU count Fixation count on TM	Attention on the passing of time		
TT word count	Product quantity		
TU count			

4.2.4 Analytical methods

Analytical methods which are used to estimate cognitive load are often based on one or more well-developed and widely accepted theoretical framework in conceptualising the cognitive processes or components of the task being researched. Such a method usually combines the subjective data elicited through expert opinion and analytical data generated with mathematical models or task analysis. For example, Sweller (1988) used the technique of *production system* (see Langley & Neches, 1981 for details) as a computational model to measure cognitive load; Seeber (2013) provided two examples on applications of Gile's (1997) Effort Model and Seeber's (2011) Cognitive Load Model as analytical methods to assess cognitive load during simultaneous interpreting. The current study adopted CLT as an analytical method to estimate the extraneous, intrinsic and germane cognitive load, during translation under different time constraints. The following subsection demonstrates how CLT was applied.

4.2.4.1 Application of CLT

So far, all the subjective (the item of Effort in NASA TLX), physiological⁸ and behavioural (PWR and MFD) metrics proposed for estimating the intensity of cognitive processing of a task can be considered as measures of *effort* from different perspectives; measures for overall cognitive load can be filtered out of this pool with the assumptions of CLT being met. As mentioned earlier, CLT assumes that extraneous, intrinsic and germane load are additive, and the overall cognitive load is the sum of the three types of loads. While extraneous load is destructive to performance, germane load is facilitative to performance. Thus, a negative correlation between extraneous load and performance, and a positive correlation between germane load and performance are assumed.

In addition, time pressure, as the major manipulated variable in the present study, is postulated to be the major source of extraneous load. Since the difficulty levels of the source materials have been controlled to be comparable (see section 5.2), the intrinsic load of the three tasks is hypothesised to be consistent across different time conditions. With different subjective and physiological measures of time pressure and the subjective, physiological and behavioural measures of effort as introduced in previous sections, valid overall cognitive load indicators can be determined based on their internal relationships. The procedures (also in Figure 4-5) are:

- 1) determine the extraneous load measures: successful time pressure measures⁹ are the candidates of extraneous load measures; meanwhile, the extraneous load measures must also demonstrate a negative correlation with translation quality based on its nature defined in CLT.
- 2) determine the candidates of overall cognitive load measures: they are the psycho-physiological and behavioural measures of effort included in this study.
- 3) calculate the sum of intrinsic and germane load: this can be determined by subtracting the z score of an extraneous load measure from the z score of a candidate of an overall cognitive load measure (effort measures), i.e., $Z_{\text{intrinsic+germane}}$

⁸ Specific physiological measures of *effort* need to be verified in the analysis by distinguishing them from the measures of time pressure with statistical tests.

⁹ Successful time pressure measures are those measures that are proved to have represented the variation of time pressure in the analysis.

load = $Z_{\text{overall cognitive load}} - Z_{\text{extraneous load}}$. Since there could be multiple extraneous load measures and effort measures, the output of this step includes all the different combinations of relevant measures. One example could be: $Z_{\text{intrinsic+germane load}} = Z_{\text{PWR}} - Z_{\text{Temporal Demand}}$.

4) test the output of $Z_{\text{intrinsic+germane load}}$ to validate the overall cognitive load measures: the variation of $Z_{\text{intrinsic+germane load}}$ should be consistent with the variation of germane load, because the intrinsic load is considered to be comparable across tasks; thus, if the output of $Z_{\text{intrinsic+germane load}}$ is positively correlated with translation quality, then the corresponding effort measure used in the formula can be taken as a valid overall cognitive load measure; if the output $Z_{\text{intrinsic+germane load}}$ is not positively correlated with translation quality, the corresponding effort measure used in that formula cannot be considered as a valid overall cognitive load measure, because it violates the assumption in CLT that germane load should have a positive correlation with performance.

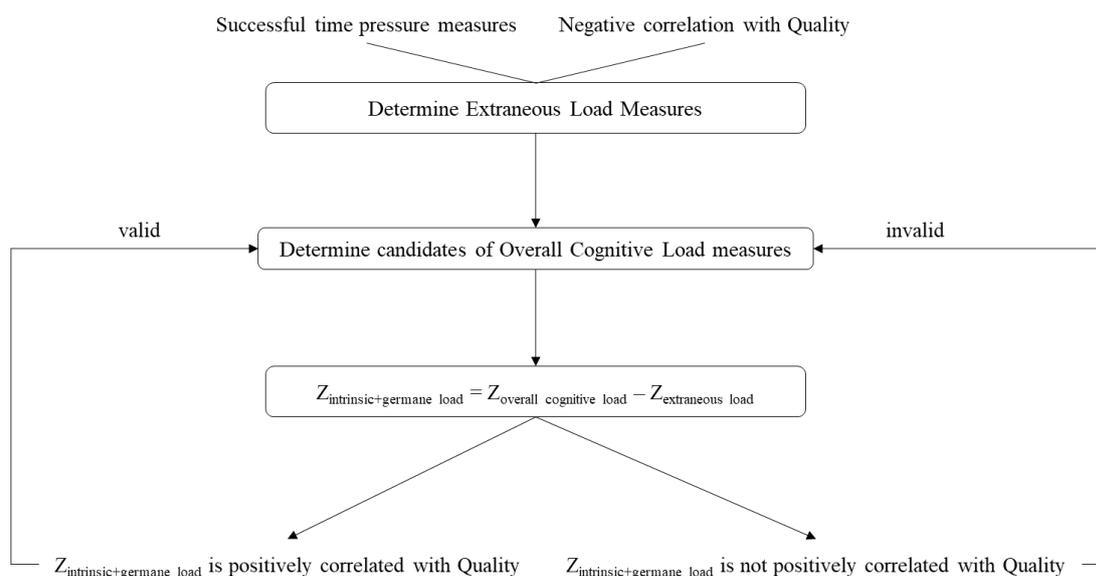


Figure 4-5. Procedures of applying CLT as an analytical method

By exploiting the relationships between different cognitive loads and task performance, overall cognitive load measures can be validated. More importantly, the effects of time pressure and experience on translators' exertion of germane load can be examined using this method.

4.3 Translation quality assessment

Task performance is one aspect of the assessment dimension of cognitive load according to Paas, Tuovinen, et al. (2003). Generally, both primary and secondary task performance can be assessed: primary task performance assessment usually focuses on aspects such as performance speed and accuracy; secondary task performance assessment focuses on the performance of a task that is performed concurrently with the primary task. The typical metrics of secondary task performance include reaction time, accuracy and error rate. The present study examined translation product quality as an overall primary task performance assessment.

Translation quality assessment in the present study was conducted with the strategy of expert assessment. Two expert translators with over 20 years' experience as professional translators (working on the English-Chinese language pair) were invited to assess all the translation products produced in the experiments. Two widely applied approaches for translation quality assessment are the error-based and the scale-based assessment. These two approaches represent the micro-textual and macro-textual perspectives of translation quality assessment respectively, and a combination of the two approaches is called mixed-methods scoring (Han, 2020). In such a mixed-methods scoring system, the final mark is produced by merging the scores from the two components, i.e., the error-based and the scale-based assessment. It is also suggested that a mixed-methods approach could yield more reliable results than that of either system when used alone (e.g., Amini, 2018; Waddington, 2001). The present study uses the Multidimensional Quality Metrics (MQM, Lommel et al., 2013) as an error analysis method and the Functional-Componential Evaluation (FCE, Colina, 2008, 2009) as a scale-based scoring method to produce the final quality scores of the translation. The weight that these two systems took in the final quality score is 50% respectively. These two systems of translation quality assessment are presented in detail in the following sections.

4.3.1 Multidimensional Quality Metrics (MQM)

As an error analysis method for translation quality assessment, MQM (Lommel et al., 2013) defines a total of over 100 error/issue types under eight major dimensions: Accuracy, Fluency, Terminology, Locale convention, Style, Verity, Design and Internationalisation. These dimensions represent the top level in the MQM hierarchy,

and under each dimension, more specific issues constituting subtypes of their parents are defined. In order to simplify its application, MQM also defines a smaller “MQM Core”¹⁰ consisting of 20 error types in seven dimensions that represent the most common issues arising in quality assessment of translated texts (Figure 4-6).

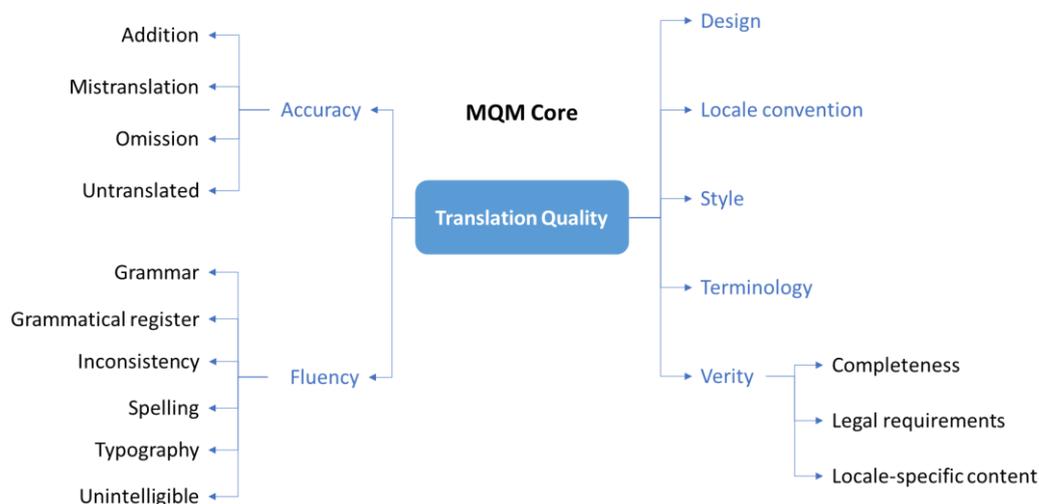


Figure 4-6. Dimensions and error types in MQM Core

MQM is highly flexible and can be customised to fit users’ needs, which constitutes the most important feature or advantage of this system. These core issues/error types were further tailored in the present study with four main dimensions included, namely accuracy, fluency, style and terminology; eleven specific error types under these dimensions which are more relevant to and better suit the scopes and purpose of the present study are presented in Table 4-5 below with their descriptions given by MQM¹¹:

¹⁰ For more details about MQM Core, see MQM website: <http://www.qt21.eu/mqm-definition/definition-2015-12-30.html#mqm-core>.

¹¹ The descriptions of the error types are from MQM website: <http://www.qt21.eu/mqm-definition/issues-list-2015-12-30.html>.

Table 4-5. MQM dimensions and error types applied in the present study

<i>Dimension</i>	<i>Error type</i>	<i>Description</i>
Accuracy	Addition	The target text includes text that is not present in the source.
	Omission	Content is missing from the translation that is present in the source.
	Mistranslation	The target content does not accurately represent the source content.
Fluency	Ambiguity	The text is ambiguous in its meaning.
	Coherence	The text is not coherent: While individual sentences of the text are all perfectly fluent, the text as a whole does not make sense and is inconsistent with itself.
	Cohesion	Portions of the text needed to connect it into an understandable whole (e.g., reference, substitution, ellipsis, conjunction, and lexical cohesion) are missing or incorrect. Example: An English text is missing conjunctions and particles (e.g., “thus”, “therefore”, “but”, and “however”) needed for the logic of the text to be clear.
	Grammar	Issues related to the grammar or syntax of the text, other than spelling and orthography.
	Inconsistency	The text shows internal inconsistency. Example: The text states that bug reports should be submitted to a mailing list in one place and via an online bug tracker tool in another.
	Spelling	Issues related to spelling of words.
Style	Register	The text uses a level of formality higher or lower than required by the specifications or general language conventions. Example: A formal letter uses contractions, colloquialisms, and expressions characteristic of spoken rather than written language, and those comes across as less serious than intended.
Terminology	Inconsistent with domain	A term is used contrary to general (subject) domain expectations.

By default, there are three levels of severity for each error type in the scoring procedures of MQM, namely the minor error, major error and critical error. In the current research, a critical error only referred to incomplete translation of the whole text, i.e., cases that one or more sentences were missing in the target text and one missing sentence was counted as one critical error. In some cases, due to the deadlines, participants might not manage to complete the task and left some sentences untranslated usually at the end. The untranslated ST sentence(s) could make the TT obviously incomplete. This can be easily distinguished from the intentional *omission* of words, as a translation strategy. If part of a sentence was untranslated, a half (0.5) critical penalty applied. The number of minor and major errors were counted against each error/ issue type. The penalty for a critical error (i.e., missing one sentence in the translation) was 20 points; the penalty of one major error was 6 points and the penalty of one minor error was 2 points. The final MQM score was calculated by the formulas below:

$$MQM \text{ score} = 100\% - \text{Penalties}$$

$$\text{Penalties} = \frac{\text{Minor Error} \# \times 2 + \text{Major Error} \# \times 6 + \text{Critical Error} \# \times 20}{\text{ST word count (average 200)}} \times 100\%$$

The raters needed to identify all the errors in a translation that fell into the prescriptive error types and determine their severity level (either minor or major); the critical errors, i.e., number of missing/untranslated sentences, were calculated by the researcher. For example, if five minor errors and five major errors were identified in a TT by a rater, and 0.5 sentences in an ST containing 200 words were untranslated, the final MQM score would be $100\% - (5 \times 2 + 5 \times 6 + 0.5 \times 20) / 200 \times 100\% = 75\%$.

4.3.2 Functional-componential evaluation (FCE)

The Functional-Componential Evaluation (FCE) system is a feasible and efficient translation assessment system proposed by Colina (2008, 2009). The system is based on a theoretical approach of functionalist and textual models of translation. It represents a scale-based or holistic assessment approach. This system has been verified by a series of empirical experiments (of different language pairs including Chinese) with solid inter-rater reliability. It includes four components of quality separately, namely the target language, functional and textual adequacy, non-specialised content and the specialised content and terminology. Each component of this system has four categories/bands with different score values as presented in Table 4-6. Unlike the error-based system, such a scale-based or holistic assessment tool does not rely on a point deduction system, but tries to match the text under evaluation with one of several descriptors provided for each category/component of evaluation (Colina 2009, p. 240).

Table 4-6. Descriptions of components and categories in FCE provided by Colina (2009)

<i>Component</i>	<i>Category</i>	<i>Value</i>	<i>Description</i>
Target language	1.a	5	The translation reveals serious language proficiency issues. Ungrammatical use of the target language; spelling mistakes. The translation is written in some sort of 'third language' (neither the source nor the target). The structure of source language dominates to the extent that it cannot be considered a sample of target language text. The amount of transfer from the source cannot be justified by the purpose of the translation. The text is extremely difficult to read, bordering on being incomprehensible.
	1.b	15	The text contains some unnecessary transfer of elements/structure from the source text. The structure of the source language shows up in the translation and affects its readability. The text is hard to comprehend.
	1.c	25	Although the target text is generally readable, there are problems and awkward expressions resulting, in most cases, from unnecessary transfer from the source text.
	1.d	30	The translated text reads similarly to texts originally written in the target language that respond to the same purpose, audience and text type as those specified for the translation in the brief. Problems/awkward expressions are minimal if exist at all.
Functional and textual adequacy	2.a	5	Disregard for the goals, purpose, function and audience of the text. The text was translated without considering textual units, textual purpose, genre, need of the audience, (cultural, linguistic, etc.). Cannot be repaired with revisions.
	2.b	10	The translated text gives some consideration to the intended purpose and audience for the translation, but misses some important aspect/s of it (e.g., level of formality, some aspect of its function, needs of the audience, cultural considerations, etc.). Repair requires effort.
	2.c	20	The translated text approximates to the goals, purpose (function) and needs of the intended audience, but it is not as efficient as it could be, given the restrictions and instructions for the translation. Can be repaired with suggested edits.
	2.d	25	The translated text accurately accomplishes the goals, purpose (function: informative, expressive, persuasive) set for the translation and intended audience (including level of formality). It also attends to cultural needs and characteristics of the audience. Minor or no edits are needed.
Non-specialised content	3.a	5	The translation reflects or contains important unwarranted deviations from the original. It contains inaccurate renditions and/or important omissions and additions that cannot be justified by the instructions. Very defective comprehension of the original text.
	3.b	10	There have been some changes in meaning, omissions or/and additions that cannot be justified by the translation instructions. Translation shows some misunderstanding of original and/or translation instructions.
	3.c	20	Minor alterations in meaning, additions or omissions.
	3.d	25	The translation accurately reflects the content contained in the original, insofar as it is required by the instructions without unwarranted alterations, omissions or additions. Slight nuances and shades of meaning have been rendered adequately.
Specialised content and terminology	4.a	5	Reveals unawareness/ignorance of special terminology and/or insufficient knowledge of specialised content.
	4.b	10	Serious/frequent mistakes involving terminology and/or specialised content.
	4.c	15	A few terminological errors, but the specialised content is not seriously affected.
	4.d	20	Accurate and appropriate rendition of the terminology. It reflects a good command of terms and content specific to the subject.

The raters needed to select an appropriate category for each of the four components in FCE that best described or matched the given translation. The total FCE score would be the sum of the values corresponding to the category of each component rated by the raters. What should be noted is that the completeness of the translated text was accounted for by the components of functional and textual adequacy and non-specialised content. This means that if a translation was incomplete, these two components should have lower marks rated by the raters.

4.3.3 Inter-rater and internal reliability of the assessment

As mentioned, two raters were recruited for assessing the quality of all the translations in the present study. They received £1 per text as remuneration for their work and each of them assessed 195 (65*3) texts in total. The inter-rater reliability (Krippendorff's alpha) for each translation assessment system was examined and it showed that the two raters were highly consistent in their ratings for both systems (Table 4-7). Also, an internal reliability (Cronbach alpha) for each rater in using the two translation quality assessment systems was examined, and the two systems exhibited a high consistency in the results of the two raters as well. These tests prove that the application of these two translation quality assessment systems can yield highly reliable results on translation quality scores. The final overall translation quality score for each target text was calculated as the average score given by Rater 1 and Rater 2 with each system taking a weight of 50% in each rater's assessment.

Table 4-7. Inter-rater and internal reliability of the translation quality assessment

<i>Inter-rater reliability (Rater 1 vs. Rater 2)</i>	<i>Krippendorff's alpha</i>
MQM	0.958
FCE	0.956
<i>Internal reliability (MQM vs. FCE)</i>	<i>Cronbach alpha</i>
Rater 1	0.947
Rater 2	0.955

4.4 Summary

This chapter examines the overall research methodology adopted by the present study. Firstly, in order to properly induce time pressure, both objective and subjective

manipulation strategies were proposed and applied. These included the externally imposed deadlines, giving pre-task instructions about time, increasing participants' intrinsic motivation for the task and visualising the passing of time. It is considered that since time pressure comes with a perceptual nature, the supplementary subjective strategies could intensify the participants' feelings of time pressure in addition to the objective strategy of imposing external deadlines.

The methods applied in assessing time pressure, effort/cognitive load and translation behaviours included subjective, physiological, behavioural and analytical tools. The subjective or self-report scales included a psychometric instrument, STAI, for measuring participants' state and trait anxiety, and NASA TLX for measuring the level of Effort, Mental Demand, Temporal Demand and Stress experienced by the participants in the tasks. A variety of biomarkers were measured in order to examine the physiological effects of time pressure on translation. Specifically, HR, GSR, HRV and skin temperature were continuously recorded by a none-intrusive device, i.e., an Empatica E4 wristband during the experiment; pupil size was also continuously recorded by an eye tracker, while blood pressure was measured immediately after the tasks. These biomarkers were tested to indicate either time pressure or effort in the translation tasks depending on their relationships to one another and with other subjective and/or behavioural measures in the analyses. The behavioural methods mainly included eye tracking and keystroke logging techniques with which meaningful metrics relating to different aspects (such as automaticity and cognitive rhythm) of translation behaviours can be generated. Based on the inherent relationships between (different types of) cognitive load and task performance, CLT was applied as an analytical method for calculating or inferring different types of cognitive load (especially germane load and overall cognitive load) in the current study. This method offers a new pathway to validating various overall cognitive load indicators, and more importantly, an in-depth scrutiny can be carried out with regard to how germane load, which is supposed to facilitate learning and performance, can be modulated by time pressure in the analysis.

Lastly, the translation product quality was assessed by two different systems, i.e., MQM and FCE, which constitutes a mixed-methods scoring system for translation quality evaluation. While MQM features an error-based assessment system, FCE represents an efficient scale-based or holistic assessment system. It turned out that the combination of these two systems in the present study yielded reliable results. Overall,

the present investigation involved a variety of methods which generated diversified types of data including online and offline as well as process and product data (see Figure 4-7). The methods and data are expected to provide a solid methodological foundation for the study.

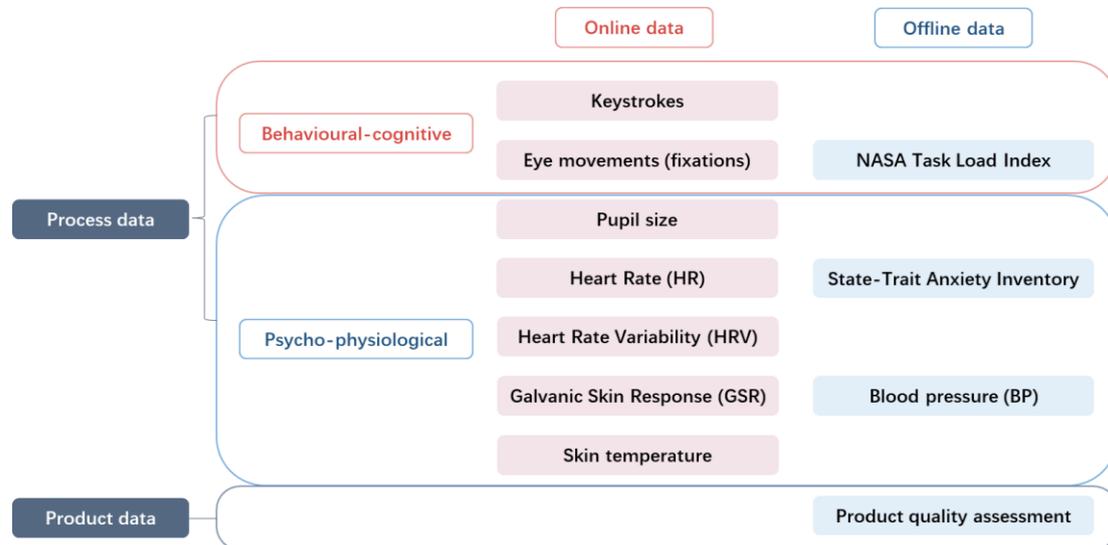


Figure 4-7. Categorisation of data type in the present study

Chapter 5. Experimental Setup and Data Processing

The present study adopted a within-subject design. Each participant translated three comparable texts from English into Chinese, with each translation task rotated across three time conditions, i.e., *Short*, *Standard* and *Free*. A series of experiments were conducted from May to December 2019. This chapter presents the details of the experimental setup including information about the participants (section 5.1), the selection and evaluation procedures of the source materials (section 5.2), and a pre-test for determining the time constraints and several other key issues for the subsequent formal experiment (section 5.3). The procedures and apparatus involved in the experiment are demonstrated in section 5.4. In addition, illustrations about the physiological dataset and data screening criteria for the eye tracking data are given in section 5.5. The statistical analysis methods and tools are discussed in section 5.6. Section 5.7 summarises the whole chapter.

5.1 Participants

Sixty-five participants were recruited in the formal experiment of this study on a voluntary basis. These participants were grouped as *novice translators* and *experienced translators* based on their professional translation work experience in years. The novice group consisted of 35 postgraduate students (30 master's students in the 2018-19 one-year programme and five first- or second-year doctoral candidates) majoring in Translation Studies at Durham University. The master's students had completed most of their courses in the programme at the time the experiment was conducted. The majority of participants in this group had no professional translation work experience by then. There were five males and 30 females in the novice group and their average age was 24.14 years (range=21-34, SD=2.68).

The group of experienced translators consisted of 30 participants: 10 of them were postgraduates (five master's students and five doctoral candidates majoring in Translation Studies) who already had over three years' professional experience in English-Chinese translation; 20 were non-student translators who had three to twenty years of professional experience in English-Chinese translation. There were nine males and 21 females in the group of experienced translators and their average age was 30.73 years (range=24-50, SD=7.80).

The participants were grouped by their professional translation work experience (without work experience vs. more than three years' experience) instead of their identity

(student vs. non-student) because the present study aims to understand how professional experience makes a difference in translating under time pressure. The profiles of the student participants showed that a portion of them already had plenty of professional experience before they took the postgraduate programme. For example, one student had worked for a large newspaper as a news translator for four years and a few of them had been working for translation companies as stable free-lance translators for three to five years. Concerning their experience, it would be more reasonable to group them as experienced translators rather than “students” simply because of their current identity.

As pointed out by Jääskeläinen et al. (2011, p. 146), with a didactic purpose, translation process research often seeks to find out how “what constitutes professional performance (descriptive aim) will help us educate the students into professionals (applied aim)”. Thus, comparing the performance of experienced and non-experienced groups with the assumption that novices can attain expertise has become one of the most researched foci in translation process research. To this end, the notion of *relative expertise* (as opposed to *absolute expertise* proposed by Chi, 2006), which reflects the underlying didactic aims, has been widely adopted in process studies so that the conventional requirement of “expertise” (e.g., with 10 years or 10,000 hours of deliberate practice as mentioned in Shreve, 2006) has been much relaxed in recent years.

Nonetheless, the present study does not adopt terms such as “professional” or “expert” translator to avoid an arbitrary definition of professionalism or expertise; rather, the translators are grouped based only on their actual working experience in translation. The criterion set for labelling a translator as experienced or novice in the present study is whether or not they had over three years’ experience of working as a translator (either working as a free-lance or an in-house translator with a stable income from the translation work). This criterion (three years) also conforms to previous studies (e.g., Hvelplund, 2011).

All the participants in the two groups were native Mandarin Chinese speakers with English as their second language. The student participants (from both the novice and experienced translator group) shared a similar background in language proficiency with an average IELTS score of 7.4 (range=7-8, SD=0.38). All the non-student participants were certified translators with translation related degrees. All the participants were touch-typists and had normal or corrected-to-normal vision.

In addition to the 65 participants in the formal experiment, another 13 master’s students (three males and 10 females; average age was 21.38 years ranging from 19 to

25 years) majoring in Translation Studies (from the 2017-18 programme at Durham University) participated in a pre-test which was designed to determine several key issues for the formal experiment. The procedures of the pre-test are demonstrated in section 5.3.

5.2 Source texts

Three source texts were selected and excerpted from *The Economist*, an international weekly newspaper focusing on current affairs, international business, politics and technology. Each of these texts comprises 11 sentences and on average, 201 words (range=196-207, SD=5.51). They share a similar topic which is about the trade war between China and the US. They are comparable in their difficulty levels regarding both the *objective measures* and *subjective evaluation*. The *objective measures* of text complexity were examined mainly based on Hvelplund (2009), including the main readability indices (the first six items in Figure 5-1) and Word Frequency (the last item in Figure 5-1). Each word in the texts was examined in the British National Corpus, and the words which appeared less than 1000 times (i.e., the frequency value) in the corpus were marked as *Frequency 1000*. The Word Frequency index in this study thus refers to: $[(\text{Frequency } 1000 \div \text{the number of total unrepeated words of the text}) \times 100\%]$. Figure 5-1 shows that the three texts were comparable across all these indices.

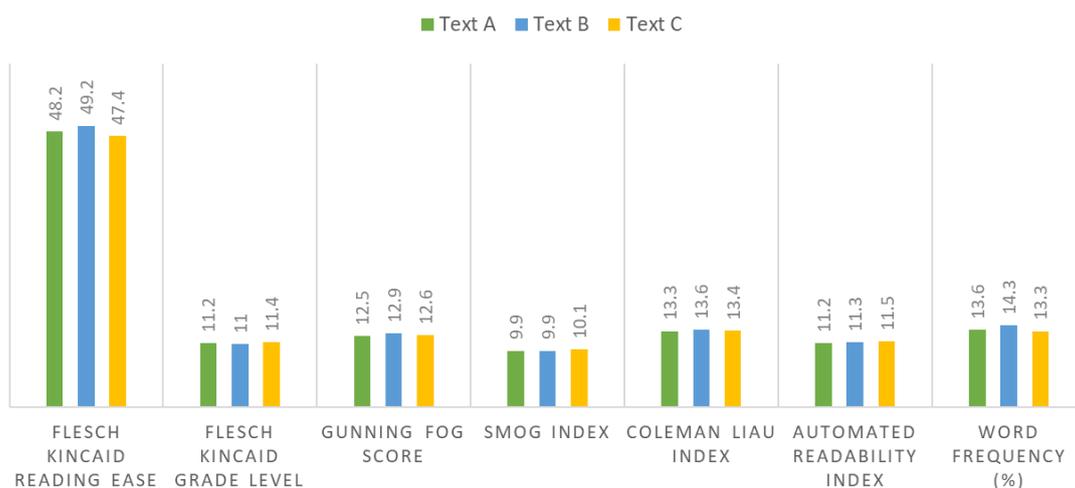


Figure 5-1. Scores of the objective measures of text difficulty

The *subjective evaluation* was conducted in two steps: 1) eight external experts

voluntarily evaluated each text regarding their “Comprehensibility” and “Translatability” and gave a score on a 21-point scale¹², with “very low” and “very high” on the left and right ends; and 2) after translating the texts, 13 participants in the pre-test (see details in section 5.3) retrospectively assessed the difficulty level of translating the texts on the 21-point scale. The repeated measures ANOVA test shows that there was no significant difference among the three texts (Table 5-1). Overall, concerning the *objective measures* and the *subjective evaluation*, the three texts are considered comparable in translation difficulty.

Table 5-1. Statistics on the subjective evaluation of text difficulty level

Texts	Comprehensibility			Translatability			Difficulty of Translation		
	Mean (SD)	Range	N	Mean (SD)	Range	N	Mean (SD)	Range	N
A	50.00 (10.69)	40-60	8	52.50 (10.35)	40-60	8	55.00 (17.08)	30-80	13
B	40.00 (0.00)	40-40	8	45.00 (14.14)	20-60	8	55.38 (14.21)	30-80	13
C	47.50 (21.21)	40-100	8	57.50 (19.82)	40-100	8	58.85 (10.03)	40-70	13
ANOVA	F(2,14)=1.36, <i>p</i> =.289			F(2, 14)=1.43, <i>p</i> =.272			F(2, 24)=0.71, <i>p</i> =.501		
<i>Note:</i> SD=standard deviation; N=number of participants; Range=participants' response range									

5.3 The pre-test and time constraints

A pre-test was carried out at the preparation stage (in August 2018) of this study involving 13 participants. They took part in the study voluntarily. The specific purposes of the pre-test were: 1) to determine appropriate time constraints for time pressure inducement in the formal experiment; 2) to collect subjective evaluation data on the translation difficulty of the three source texts; 3) to identify the most likely unfamiliar words in each text which could be provided to participants in the later experiment as a glossary list.

Technically speaking, there were two stages of the pre-test, i.e., pre-test 1 and pre-test 2, because adjustments were made during the process in order to best fulfil all the purposes. Initially, in pre-test 1, each participant translated all the three texts in a counterbalanced sequence, with the first task having no time limit, the second task having a deadline which is 75% of the time used in the first task and the last task having a deadline which is 50% of the time used in the first task. There were four participants

¹² The difficulty assessment scales designed for both experts (comprehensibility and translatability) and students (difficulty of translation) adopted the NASA TLX scale, and thus its score calculation method (raw scores) was also applied: [(the number of points a participant marked-1)×5].

who took part in this pre-test 1 and it was found that most participants could not complete the last translation task given the tight deadline. Therefore, the design was adjusted. In pre-test 2 which involved nine participants, each participant translated all the three texts in a counterbalanced sequence, with the first two tasks having no time limit and the last task having a deadline which was 75% of the averaged time used in the first two tasks. It was found that all the nine participants managed to complete the last task, and two of them even completed and submitted the task before the deadline. They all retrospectively reported that they felt the urgency of time during the last task.

Thus, there were 24 task sessions from the 13 participants that were used for calculating the average timeframe of translating the three texts: 1) the sessions without a time limit in pre-test 1 (one session * four participants) and pre-test 2 (two sessions * nine participants); 2) the sessions in pre-test 2 that were assumed to have a deadline, but the time was not used up by two participants (one session * two participants). The average (which is also the median) of this dataset, 20 minutes 25 seconds, was determined as the *Standard* time condition for the formal experiment, and quartile one (16 minutes 15 seconds) of the dataset was determined as the *Short* time condition which was slightly shorter than the average time used in the last task in pre-test 2 (16 minutes 36 seconds). The third time condition for the formal experiment has no time limit, which is called a *Free* time condition.

Participants in both pre-test 1 and 2 were asked to answer a retrospective question on the translation difficulty of each text regardless of the time constraint (if there was one). Thus, 13*3 answers were collected regarding the subjective evaluation of the translation difficulty of the texts. All the participants in the pre-tests were also asked to point out the English words in the source texts for which they did not know the meaning. Based on the results, the top seven words for each text were identified as the most likely unfamiliar words which were provided in a glossary list in the formal experiment (see Appendix 2).

5.4 Procedures and apparatus

The study was approved by the research ethics committee of School of Modern Languages and Cultures, Durham University before the experiments started. The experiments with 52 participants (35 novices and 17 experienced translators) were conducted at an eye tracking lab (with Tobii TX300 eye tracker) at the University in

the UK; the experiments with the other 13 experienced translators were conducted either outside the campus in the UK or in China with a portable eye tracker (Tobii Pro X3-120). To minimise any negative influences on data quality, the participants were required not to drink alcohol in the 24 hours before the experiment, and they all reported no known diseases. A consent form (see Appendix 5) ensuring anonymity and confidentiality was signed by each participant, and they were rewarded with a supermarket gift card (£20) for their participation in the experiment.

Figure 5-2 presents a flow chart of the experiment procedures. After the consent form was signed, each participant was asked to fill in a background information form (see Appendix 6); then, the participants had a short rest session which was used for measuring the baseline values of the biomarkers (HR, GSR, HRV, skin temperature and BP). A warm-up session followed to get them familiarized with the interfaces and devices in the experiment. Before each task session started, an *Instruction* was briefed: for the *Short* session, they were told that the time given was less than what was normally required; for the *Standard* session, they were told that the time given was the average time interval that was normally needed; and for the *Free* session, they were told that they could take as much time as they need to complete the task. In addition, the participants were informed that their translation product would be assessed by an expert and they would receive the feedback later.

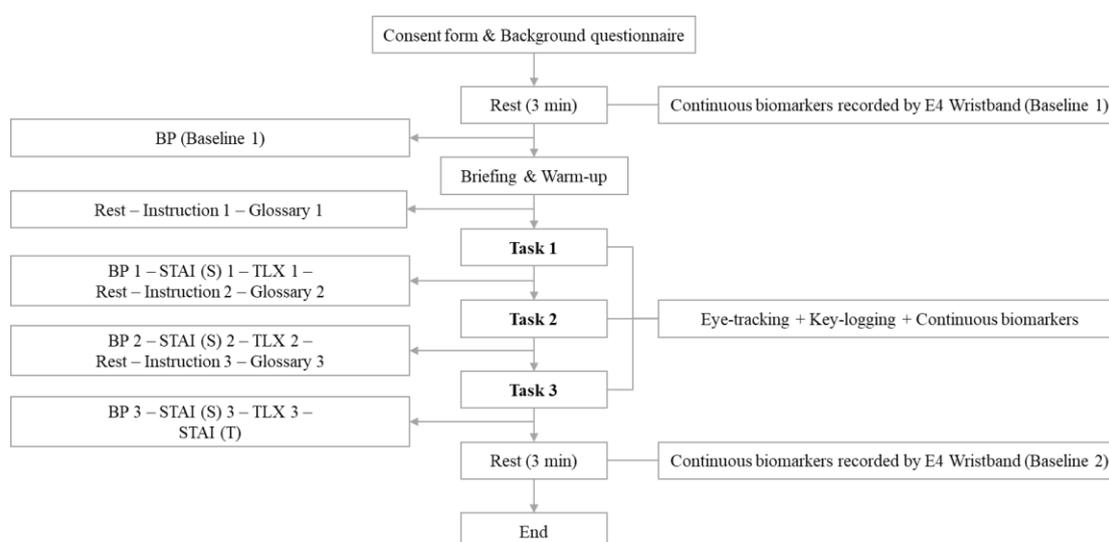


Figure 5-2. Experiment procedures

When the task session started, the source text was presented at the top of the screen

followed by the target text editing box. A countdown timer (for the *Standard* and *Short* sessions) or a stopwatch (for the *Free* session) was displayed at the bottom to remind the participants about the time remaining or time used (see Appendix 8). The purpose of visualising the passing of time and giving the instructions about timing and feedback was to intensify the participants' subjective perception of time pressure during the process of translation. The combinations of task order, text and time condition were substantially counterbalanced (see Appendix 7) to minimise the fatigue or order effects on the testing variable which is time condition.

The participants were not allowed to consult any online or offline resources during the translation tasks, but they were provided with a glossary list of the top seven unfamiliar words in the source text, which were identified in the pre-test; and they were free to consult resources before each task regarding the words on this glossary list. This decision was made due to a concern that when given access to online resources, most participants would opt for using a machine translation (such as Google translate) and perform post-editing instead of translation, especially when working with a tight deadline. This would largely alter the nature of the task and influence the results. Thus, to let the participants concentrate solely on the translation activity, there was no access to consultation or documentation during the translation process.

During each task, the eye tracking and keystroke logging data was recorded by an eye tracker (Tobii TX300 of 300Hz sampling rate or Tobii Pro X3-120 of 120Hz sampling rate) and Translog II respectively. The desktop Tobii TX300 eye tracker used in this study was integrated with a removable 23'' monitor in 1280×1024 pixels; the Tobii Pro X3-120 (portable) eye tracker was used with a ThinkPad laptop with a 14'' monitor in 1920*1080 pixels. Participants were asked to sit in a chair at around 60-65 centimetres from the screen. The eye tracker was calibrated with the regular 5-points-calibration at the beginning of the warm-up exercise and all task sessions. To keep the lighting consistent, the window-less room of the eye tracking lab (with desktop Tobii TX300 eye tracker) was equipped with an overhead fluorescent light during the experiments. Since the experiments with the 13 experienced translators using the portable Tobii Pro X3-120 eye tracker were not conducted in a lab – mostly taking place at the participants' workplace or home, the environment of these places was not as controllable and stable as in the university's eye tracking lab. It should thus be noted that the mixed data collection settings may potentially impact the final results.

HR, skin temperature, HRV and GSR data were continuously recorded by the non-

invasive device Empatica E4 Wristband with a frequency of 4Hz during the tasks. This device was worn by the participants on their wrist just like a normal watch. Therefore, it did not affect the translation or typing process. The software Empatica E4 Realtime offers real-time and accurate data acquisition and visualisation, which is suitable for laboratory setting and academic research purposes.

After each task, blood pressure was measured immediately by an Omron M7 Intelli IT blood pressure monitor. During the measurement, the participants sat in a chair with their feet on the floor and their right arm supported by the desk to make the elbow at about heart level. The inflatable cuff was placed on their upper (right) arm on the bare skin before it was pumped up to obtain the readings. The State Anxiety form was completed after the measurement of blood pressure, followed by the NASA TLX questionnaire. The Trait Anxiety form was completed after the last task session as recommended by the instrument's instructions. In addition, the continuous biomarkers were recorded by the Empatica E4 Wristband in another 3-minute rest session (Baseline 2) before the end of the whole experiment.

Although the participants did not report any specific discomfort feelings regarding the experimental measurements and observations made before, during and after the translation sessions, the experimental setting may have an impact on their performance. This is inevitable, however, concerning the nature of an empirical study which cannot duplicate a one-hundred percent natural workplace in the experiment.

It should also be noted that at the initial stage of the formal experiment, there was still a trial-and-error process or a pilot phase in which the Baseline 2 session (for the measurement of the continuous biomarkers) was not included. This stage involved 16 participants (eight novices and eight experienced translators). This means that these 16 participants only took the Baseline 1 session at the beginning of the experiment. Later, the Baseline 2 session at the end of the experiment was added to the formal procedures for the rest of the 49 participants in order to have a more comprehensive examination of their physiological baseline data.

5.5 Data acquisition and quality assessment

The three task sessions (i.e., Short, Standard and Free) did not observe a data loss problem for the continuously recorded physiological data obtained via the Empatica E4 Wristband (i.e., HR, skin temperature, GSR and HRV), since the task sessions lasted

long enough to allow the device to record the data for at least parts of the whole session; however, the 3-minute baseline (1&2) measurements had a problem of data loss for several sessions due to the short recording time and malfunction of the device (e.g., the wristband was not worn firmly enough and sometimes the sensors were not fully in contact with the participant's skin when they were moving). Section 5.5.1 illustrates the physiological data acquisition and processing procedures for the datasets. As mentioned by Hvelplund (2014), poor quality eye tracking data have the risk that they may not realistically reflect the translator's eye movements. The eye tracking data thus went through a screening process (section 5.5.2) to ensure that the results were not skewed by flawed data.

5.5.1 Physiological data

The GSR indices of SCR Amplitude and SCR Frequency were obtained via a MATLAB based software, Ledalab (version 3.4.9; <http://www.ledalab.de/>). The continuous decomposition analysis (Benedek & Kaernbach, 2010) was used for the extraction of SCRs, and the amplitude threshold of 0.01 μS was adopted. The index of GSR Mean was obtained by averaging the raw GSR data collected in each baseline and task session by Empatica E4 Wristband. The indices of HRV (i.e., RMSSD, SDNN and PNN50) were extracted through the software Kubios (Tarvainen et al., 2014) based on the raw inter-beat interval data acquired from Empatica E4 Wristband. The HR and skin temperature data were obtained from Empatica E4 Wristband directly.

Two components were involved in the analyses of all the physiological data: 1) a general comparison between the baseline session(s) and the task sessions, and 2) a more comprehensive comparison of the (three) task sessions. As presented in section 5.4, while all the physiological data were recorded in the three task sessions, pupil size was not measured in the baseline sessions and blood pressure was measured once at the beginning of the experiment as a baseline value (i.e., Baseline 1) for all the 65 participants. The rest of the biomarkers, i.e., HR, HRV, GSR and skin temperature, were measured twice in two baseline sessions (1&2) at the beginning and end of the experiment respectively for 49 participants, and were measured once only at the beginning of the experiment for 16 participants. To keep a higher conformity, the analyses of HR, HRV, GSR and skin temperature for the first component (comparing the baseline and task session values) were based on data from the 49 participants who

took both baseline sessions; the analyses of blood pressure for the first component and the analyses of all the biomarkers for the second component (comparing the values of the three tasks sessions) were based on the full dataset of 65 participants.

In addition, one (out of 49) participant's GSR data was not recorded in Baseline 1 due to a malfunction of the device; thus, this participant was eliminated from the GSR baseline dataset (n=48) for the physiological data analysis of the first component. Further, another 15 participants' HRV data from either Baseline 1 or Baseline 2 were not properly recorded as well due to a malfunction of the device; thus, only 34 (49-15) participants' HRV data were valid for the analysis of the first component (see Appendix 10 for more details). The dataset for HR and skin temperature was complete for the first component with all the 49 participants included. Table 5-2 below summarises the number of participants included in each dataset for the analyses of the physiological measures.

Table 5-2. Datasets for physiological data analyses

<i>Components of analysing the physiological data</i>	<i>Biomarkers</i>	<i>Number of participants in dataset</i>
Component 1: comparison between the baseline session(s) and the task sessions	HR and skin temperature	49 (27 novice & 22 experienced)
	GSR	48 (27 novice & 21 experienced)
	HRV	34 (21 novice & 13 experienced)
	BP	65 (35 novice & 30 experienced)
Component 2: comparison of the (three) task sessions	All biomarkers (HR, BP, GSR, HRV, pupil size, skin temperature)	65 (35 novice & 30 experienced)

5.5.2 Eye tracking data

Processed by Tobii Pro Studio or Tobii Pro Lab, the pupil size data (left and right eye) with a validity labelled “high confidence”¹³ for both eyes were extracted, and the data on both eyes were then averaged. Due to the need for “standardisation of the different filter settings” in experiments (Alves et al., 2009, p. 274), all experiments in this study used the same Velocity-Threshold Identification Filter, which had its velocity threshold set at 30 degrees/second. The maximum time between fixations was set at 75 milliseconds; the maximum angle between fixations was set at 0.5 degrees; and the fixation filter was set to include fixation samples that fell within a time window of at

¹³ Tobii Pro Studio and Tobii Pro Lab label the confidence level that each eye has been correctly identified. The values range from 0 (high confidence) to 4 (eye not found).

least 60 milliseconds.

Hvelplund (2011, 2014) put forward three criteria for eye tracking data quality assessment which has been fundamental to eye tracking studies in translation process research. Many studies, e.g., Sjørup (2013) and Schmaltz et al. (2016), have applied the quality assessment criteria on eye tracking data in their experiments based on Hvelplund's (2011) approach. The three criteria are:

1) Mean Fixation Duration (MFD). With the reference that the mean fixation duration during silent reading is around 225 milliseconds (Rayner, 1998, p. 373), Hvelplund (2011) used a mean fixation duration threshold of 200 milliseconds to discriminate acceptable data from non-acceptable data, while Sjørup (2013) applied a threshold of 180 milliseconds.

2) Gaze Time on Screen (GTS). GTS is the percentage of time spent gazing at the text in relation to the total time of translation production, i.e., $[\text{total fixation duration} \div \text{total task time} \times 100\%]$. GTS is an indication of either how much time the participant spent looking at the screen, or the quality of the eye tracking data. Hvelplund (2011), Sjørup (2013) and Schmaltz et al. (2016) took a threshold of GTS at around 30%.

3) Gaze Sample to Fixation percentage (GSF). The GSF percentage is calculated by comparing the total number of gaze samples with the total number of gaze samples that formed part of a fixation. Hvelplund (2011) adopted a practical threshold of GSF at 75%.

In addition to these criteria, the software Tobii Pro Studio and Tobii Pro Lab can give the value of Gaze Sample Percentage (GSP), which is calculated by dividing the number of eye tracking samples with usable gaze data that are correctly identified, by the number of attempts. This value can be taken as a convenient indicator of the overall quality of the gaze samples in an eye tracking file.

As suggested by Hvelplund (2014), while obtaining the GSF percentage is a potentially labour-intensive process, some analysis software can be useful in determining a GTS percentage. The GSP value given by Tobii Pro Studio/Tobii Pro Lab thus provides a reliable substitute for GSF. The present study adopted MFD, GTS and GSP as three criteria for eye tracking data assessment. The thresholds for MFD and

GTS were set at 180 milliseconds and 30% respectively based on previous research (e.g., Hvelplund, 2011; Sjørup, 2013; Schmaltz et al., 2016), and the threshold for GSP was set at 50% to ensure that the eye tracker identified at least half of its sampling attempts as valid gaze samples during each task. An eye tracking data file should comply with at least two of the three criteria to ensure an acceptable data quality. In addition, if one eye tracking file from a participant failed to pass the screening process, all the three eye tracking files from the same participant were discarded.

As a result, 27 (out of 195) eye tracking data files were discarded with a discard rate of 13.85%. These 27 files were from nine participants, including eight experienced translators who performed the tasks with the portable eye tracker and one novice translator who performed the tasks with the desktop eye tracker (see details in Appendix 9). This relatively high discard rate of eye tracking data for the experienced translator group might be owing to the generally less stable conditions in the experiments conducted with the portable eye tracker and also the relatively low sampling rate of the portable eye tracker itself. Overall, the trimmed eye tracking dataset included data files from 56 participants (34 novice and 22 experienced translators) and all the metrics related to the gaze data were analysed based on this dataset.

5.6 Statistical analysis methods and tools

The statistical analysis was conducted using the Linear Mixed-Effects Regression (LMER) modelling provided in the `lmer4` package (Bates et al., 2015) of the statistical software, R (version 3.6.3; <https://www.R-project.org/>). Separate LMER models were built for each target measure with the fixed effects of time condition (coded as *Condition* with three levels: *Short*, *Standard* and *Free*) and experience (coded as *Group* with two levels: *Novice* and *Experienced*). In the analysis of State Anxiety, Trait Anxiety was considered as an additional predictor in the model. Two levels (*High* and *Low*) of Trait Anxiety of the participants were categorised based on whether it was higher than ($>$) or lower than/equal to (\leq) the average value of the sampling group. In the analysis of attention allocation, an additional predictor of AOI (including two levels: *ST* and *TT*) was involved. Individual difference was considered by taking each participant as the random effect in all the models.

P-values of the fixed effects in each model were obtained via the Satterthwaite's approximation (with the `ANOVA ()` function) in the `lmerTest` package (Kuznetsova et

al., 2017). The post-hoc tests were conducted with the conservative Bonferroni¹⁴ corrections using the emmeans package (Russell, 2020) in order to address the problem of the increased probability of a Type I error in multiple comparisons performed simultaneously (Rasinger, 2008). The repeated measures correlation tests were conducted via the rmcrr package (Bakdash & Marusich, 2017). Compared with simple regression/correlation, rmcrr does not violate the assumption of the independence of observations and tends to have greater statistical power because neither averaging nor aggregation is necessary for an intra-individual research question (*ibid.*, p. 1).

The use of LMER models assumes that the distribution of residuals (i.e., the differences between the data points and the sample mean) in a given dataset is normally distributed. To account for the problem of skewness in several models (i.e., GSR Mean, SCR Amplitude, MP and LP length, SG length measured by TU count and time), the data were transformed using a logarithmic function (Baayen, 2009, p. 31). By doing so, the distribution of the data points could be more symmetrical, and the risk of observing significant effects that were driven by random outliers could be substantially reduced (Hvelplund, 2011, p. 120).

5.7 Summary

This chapter provides details of the experimental setups and data processing procedures. Firstly, 65 participants took part in the formal experiment, and they were grouped as novices (n=35) and experienced translators (n=30) based on whether or not they had more than three years' professional work experience as a translator in the past. This is because this study aims to understand how professional experience makes a difference in translating under time pressure. Secondly, three English source texts of around 200 words were excerpted from *The Economist* which showed comparable difficulty levels in terms of objective measures and subjective evaluation scores (assessed by experts and students). Thirdly, before the formal experiment was carried out, a pre-test involving another 13 master's students were conducted to determine the time constraints, the source text difficulty and the glossary list for the subsequent formal experiment. The time constraints for the formal experiment were determined as 16 minutes 15 seconds for the *Short* session and 20 minutes 25 seconds for the *Standard*

¹⁴ This thesis reports the automatically Bonferroni-adjusted p-values for pairwise comparisons with alpha set at 0.05 (i.e., the unadjusted p-value had been multiplied by the number of comparisons).

session based on this pre-test.

The research adopted a within-subject design with which each participant performed three tasks under different time conditions (i.e., Short, Standard and Free). During the experiment, each participant firstly signed the consent form and became familiar with the settings of the experiment in a warm-up session; what followed was the three task sessions and two 3-minute baseline sessions at the beginning and the end of the experiment in which the biomarkers of HR, HRV, GSR and skin temperature were measured by an Empatica E4 Wristband. The baseline value of blood pressure was measured after the first baseline session at the beginning of the experiment. The participants were not allowed to consult any online or offline resources during the task sessions, but they were provided with a glossary list identified in the pre-test before each task.

Since the 3-minute rest sessions for measuring baseline values of the continuous biomarkers were short in time, there was a problem of data loss (i.e., no data were recorded) for GSR and HRV indices in several sessions; thus, only those participants whose baseline sessions were recorded with valid data were included in the physiological data analysis of comparison between the baseline and task sessions. The in-depth physiological data analysis of comparing the three task sessions was not affected by the problem of data loss, since all task sessions had valid data recorded due to the longer time duration of the task sessions. In addition, the eye tracking data went through a data screening process to ensure reliable results. The three criteria were MFD, GTS and GSP with thresholds of 180 milliseconds, 30% and 50% respectively. An eye tracking data file should comply with at least two of the three criteria to be considered as acceptable, and once a file was discarded, the other two files from the same participant were discarded as well. As a result, 27 (out of 195) eye tracking data files from nine participants were discarded with a discard rate of 13.85%.

Lastly, all the statistical analyses were conducted using the LMER models which provides more reliable inferential results with random effects being taken into consideration. The post-hoc comparisons were carried out with Bonferroni corrections to avoid increased probability of a Type I error; instead of simple regression/correlation, repeated measures correlation (rmcorr) tests were conducted where applicable since it does not violate the assumption of the independence of observations and tends to have greater statistical power.

Chapter 6. Results

This chapter reports all the results in response to the five research questions and the corresponding hypotheses. Section 6.1 focuses on the psycho-physiological effects of time condition and experience corresponding to RQ1 (H1-H3); section 6.2 examines the behavioural-cognitive effects aiming to answer RQ2 (H4-H5) and RQ3 (H6-H9); section 6.3 focuses on how time condition and experience affect the translation quantity and quality which answers RQ4 (H10-H11), and section 6.4 reports the interplay among time condition, experience, cognitive load and translation quality in response to RQ5 (H12-H14).

6.1 The psycho-physiological effects

In the preliminary study (Weng et al., in press), it was found that several psychological and physiological measures (self-reported Stress and State Anxiety, HR, BP and pupil size) demonstrated a statistically significant effect of time stringency on student participants (n=45), showing a linear (positive) relation between time stringency and the detected arousal level indexed by these measures. The exact physiological and psychological effect of time condition was expected in the present analysis when the sampling size was further enlarged (n=65) and more importantly when the participants were grouped based on their professional experience (see section 5.1). As mentioned in Chapter 1, the research question and hypotheses for the psycho-physiological effects of time condition and experience were formulated as follows:

RQ1: What are the effects of time condition and experience on participants' psycho-physiological responses during translation?

H1: The psycho-physiological responses would show a significant effect of time condition and would vary in accordance with the stringency of time condition for both groups of participants.

H2: With the time condition getting more stringent, novice translators would have stronger psycho-physiological responses (with higher increments in the tested measures) than experienced translators.

H3: Those measures that corroborate H1 would be correlated with each other.

The measures under study were the psychological measures, including State

Anxiety (from STAI), Stress and Temporal Demand (from NASA TLX), and the physiological measures, including HR, BP, GSR, HRV, skin temperature and pupil size. The following sections present the results regarding these measures and indices. Separate LMER models were built for each target measure/index with the fixed effects of Condition (*Short, Standard and Free*) and Group (*Novice and Experienced*) unless otherwise stated.

6.1.1 The psychological effects

The fixed effects of Condition and Group on the self-reported State Anxiety, Stress and Temporal Demand levels reported in the retrospective questionnaires are presented in Table 6-1 and visualised in Figure 6-1.

Table 6-1. Type III ANOVA table of the LMER models of *State Anxiety, Stress and Temporal Demand*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
State Anxiety	Condition	2059.27	1029.63	2	126	21.46	<.001	***
	Group	0.95	0.95	1	63	0.02	.889	
	Condition:Group	40.13	20.06	2	126	0.42	.659	
Stress	Condition	13065.9	6532.9	2	126	35.63	<.001	***
	Group	191.8	191.8	1	63	1.05	.310	
	Condition:Group	466.4	233.2	2	126	1.27	.284	
Temporal Demand	Condition	46960	23480.1	2	126	127.91	<.001	***
	Group	370	370.2	1	63	2.02	.161	
	Condition:Group	1102	550.9	2	126	3.00	.053	

Note: Sum Sq. = Sum of squares; Mean Sq. = Mean square; NumDF = Numerator degrees of freedom; DenDF = Denominator degrees of freedom; Sig. = (Statistical) Significance. Statistical significance codes: $p \leq 0.001$ ***; $p \leq 0.01$ **; $p \leq 0.05$ *; $p > 0.05$: not significant. The same applies to all the tables henceforth.

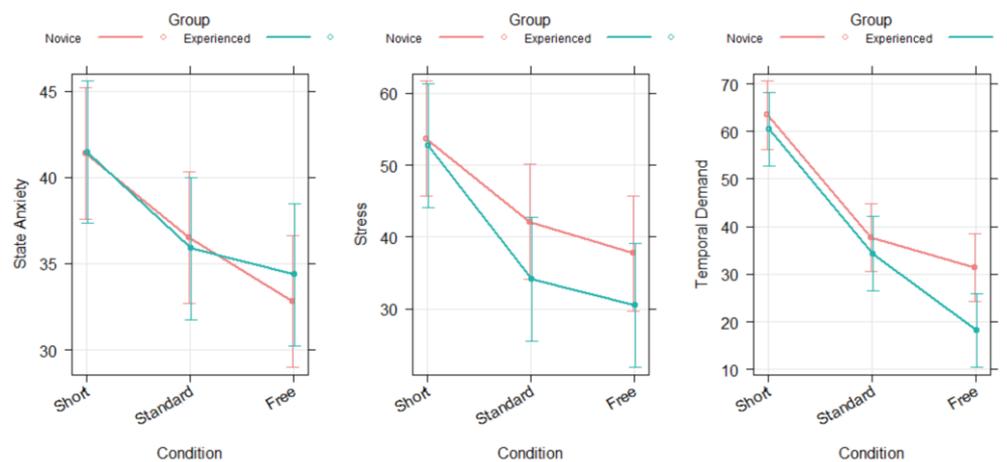


Figure 6-1. Effects of Condition and Group on *State Anxiety*, *Stress* and *Temporal Demand*

Table 6-1 shows that the effect of Condition was statistically significant in affecting all the three measures, while the effect of Group and the interaction between Condition and Group were not statistically significant. This meant that there was no noticeable group difference between novices and experienced translators, and their State Anxiety, Stress and Temporal Demand levels might vary in a similar tendency across different time conditions.

The exhibited tendencies confirmed that participants' self-reported State Anxiety, Stress and Temporal Demand levels decreased significantly with the time constraint becoming increasingly relaxed. This pattern corroborated H1 with a statistical significance of the main effect of Condition. The measure of Temporal Demand showed a marginally significant¹⁵ interaction effect between Condition and Group ($p=.053$). However, there was no sign of novice translators having a higher increment of Temporal Demand than experienced translators when the time condition was getting more stringent. In fact, when the time constraint became more stringent from the Free to the Standard condition, experienced translators showed a higher increment in Temporal Demand. Thus, H2 was not corroborated by these three psychological measures.

Additional analysis for State Anxiety was conducted involving the Trait Anxiety

¹⁵ An effect is reported as “marginally significant” in this thesis when the p-value falls between 0.05 to 0.1, which means that the effect does not reach the threshold of statistical significance (i.e., $p < 0.05$) but is close to it.

level (*High* and *Low*) as another predictor. Since the above analysis showed no obvious group difference for State Anxiety, the predictor of Group was dropped, retaining only Condition and Trait Anxiety as two fixed effects in the model. The results showed that both Condition ($F(2,126)=21.79$, $p<.001$) and Trait Anxiety ($F(1,63)=6.74$, $p=.002$) significantly influenced the variation of State Anxiety, but there was no interaction effect ($F(2,126)=0.55$, $p=.576$) between these two predictors. It was found that irrespective of the time conditions, participants with a higher Trait Anxiety tended to give higher State Anxiety scores than those with a lower Trait Anxiety. However, Trait Anxiety did not modulate the effect of Condition on State Anxiety.

In summary, H1 was corroborated by all the three psychological measures of State Anxiety, Stress and Temporal Demand. They displayed a positive relationship with time stringency in both groups. However, H2 is rejected by these measures since there was no sign of novice translators having stronger psychological responses (with higher increments in these tested measures) than experienced translators when the time condition was getting more stringent.

6.1.2 The physiological effects

This section examines how Condition and Group modulate the physiological responses during translation. Firstly, a general picture of the comparison between the baseline session(s) and the task sessions is presented in 6.1.2.1. A more in-depth analysis of how the physiological measures vary across task sessions with different time conditions is presented in 6.1.2.2.

6.1.2.1 Comparison between baseline session(s) and task sessions

As mentioned in section 5.5.1, the baseline values of the blood pressure indices (SBP, DBP and MAP; $n=65$) were measured once at the beginning of each experiment (Baseline 1), while the baseline values of HR ($n=49$), GSR ($n=48$), HRV ($n=34$) and skin temperature ($n=49$) were measured twice at the beginning (Baseline 1) and the end (Baseline 2) of each experiment. The following box plots (Figure 6-2) demonstrate the comparisons of these biomarkers between the baseline session(s) and the three task sessions.

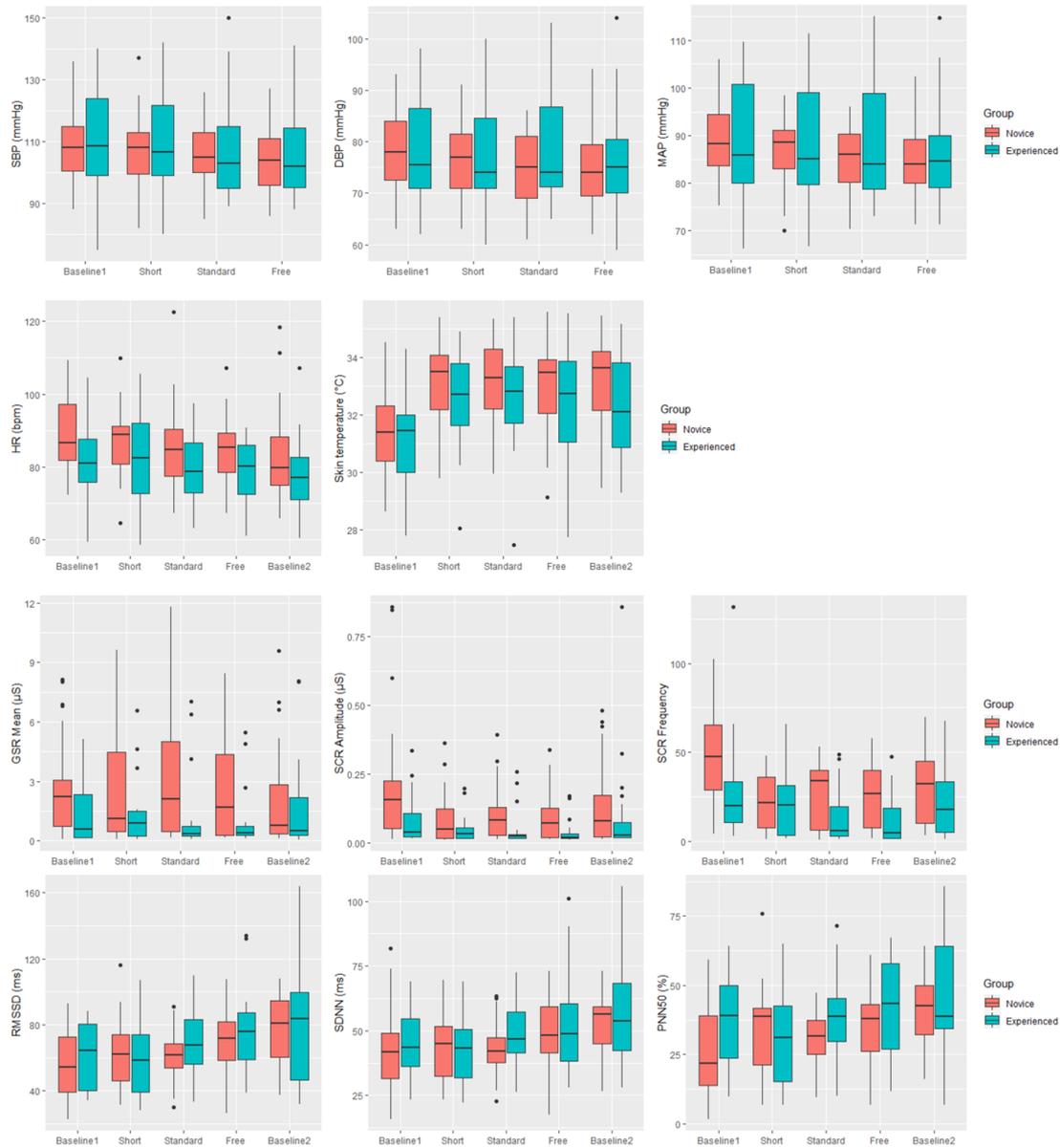


Figure 6-2. Box plots of the physiological measures in baseline and task sessions

The blood pressure values (SBP, DBP and MAP) of Baseline 1 were generally slightly higher than the values of the three task sessions for both groups, which might be caused by the effect of nervousness and feelings of uncertainty at the beginning of the experiment. Also, experienced translators showed a greater variability than novice translators, which could be caused by their wider age range (i.e., 24-50 years).

For novice translators, their HR values in Baseline 1 were also higher than that in the task sessions, but their HR values in Baseline 2 were lowered considerably compared with those in Baseline 1 and in the three task sessions. This signalled a recovery state at the end of the whole experiment, during which the nervousness and

feelings of uncertainty diminished. A similar tendency was observed in experienced translators as well, except that their HR values in Baseline 1 were much lower than that of the novices, and the median value approximated the value of the Short task session (Figure 6-2). It was found that overall, novice translators had significantly higher HR values than experienced translators in the two baseline sessions ($F(1, 47)=6.56, p=.014$), which might be caused by the age difference of the two groups – novice translators were younger than experienced translators in average.

The tendency of the skin temperature values was similar for the two groups: the values in Baseline 1 were obviously lower than those in the task sessions and Baseline 2, indicating a highly aroused state at the beginning of the experiment, while the values in Baseline 2 were similar to those in the three task sessions (Figure 6-2).

The GSR indices had considerably more outliers than other measures, especially for GSR Mean and SCR Amplitude. The median values and variability of these indices for novice translators were higher and greater than experienced translators. For novice translators, the median values of the three GSR indices in Baseline 1 were generally higher than those in Baseline 2. Although it appears that for novices, the GSR Mean value of Baseline 2 was slightly lower than that of the task sessions, there was no noticeable difference in their SCR Frequency and SCR Amplitude values between Baseline 2 and the task sessions. For experienced translators, there was little difference between Baseline 1 and Baseline 2 of all the three GSR indices; the contrast of the median values of GSR Mean and SCR Amplitude between the baseline sessions and the task sessions was also not evident. However, the median value of SCR Frequency for experienced translators in the two baseline sessions was similar to that of the Short task session and higher than the other two (Standard and Free) task sessions (Figure 6-2). Overall, the GSR indices showed more divergent results in manifesting the contrast between the baseline and task sessions for the two groups.

The HRV indices showed that for both novice and experienced translators, Baseline 1 had lower values, indicating a more aroused state, than Baseline 2. For novices, the HRV values in Baseline 1 approximated that in the Short task session while the values in Baseline 2 was higher than those in all the task sessions, indicating a more relaxed state in Baseline 2. A similar tendency was observed in experienced translators, except that their values of PNN50 in Baseline 1 were higher than those in the Short task session. These results regarding the comparison between the two baseline sessions and between the baseline and task sessions were generally consistent with most other

physiological measures (e.g., HR, BP and skin temperature).

Overall, a consistent feature revealed by the comparison of these physiological measures between the baseline sessions and the task sessions was that the participants tended to be in a highly aroused state at the beginning of the experiment (Baseline 1), even though they were told to have a rest in this session. Primarily, it showed that their arousal level in Baseline 1 came close to or went higher than that in the task with the most stringent time condition, as manifested in several biomarkers (e.g., HR, skin temperature and HRV). In contrast, they recovered to reach a much more relaxed state at the end of the experiment (Baseline 2) when they completed all the tasks, as indicated by the change of most biomarkers under study.

6.1.2.2 Comparison of the task sessions

The LMER models were built for each physiological measure to examine the effects of Condition and Group. The fixed effects of Condition, Group and their interaction on the physiological measures (indices) of HR, BP, pupil size, skin temperature, GSR and HRV are reported in Table 6-2 below. Figure 6-3 displays the variation of these physiological measures (and indices) across different time conditions and groups.

Table 6-2. Type III ANOVA table of the LMER models of the physiological measures and indices

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
HR	Condition	174.23	87.12	2	126	3.87	0.023	*
	Group	44.76	44.76	1	63	1.99	0.163	
	Condition:Group	10.64	5.32	2	126	0.24	0.790	
SBP	Condition	434.16	217.08	2	126	6.63	0.002	**
	Group	22.19	22.20	1	63	0.68	0.413	
	Condition:Group	6.47	3.23	2	126	0.10	0.906	
DBP	Condition	116.16	58.08	2	126	2.79	0.065	
	Group	27.74	27.74	1	63	1.33	0.253	
	Condition:Group	52.16	26.08	2	126	1.25	0.290	
MAP	Condition	247.83	123.91	2	126	5.36	0.006	**
	Group	22.15	22.15	1	63	0.96	0.331	
	Condition:Group	37.08	18.54	2	126	0.80	0.450	
Pupil size	Condition	0.16	0.08	2	126	5.85	0.004	**
	Group	0.01	0.01	1	63	1.03	0.313	
	Condition:Group	0.04	0.02	2	126	1.26	0.286	
Skin Temperature	Condition	0.11	0.05	2	126	0.14	0.870	
	Group	0.29	0.29	1	63	0.75	0.390	
	Condition:Group	0.87	0.44	2	126	1.12	0.330	
GSR Mean (log)	Condition	0.60	0.30	2	126	1.22	0.300	
	Group	0.57	0.57	1	63	2.29	0.135	
	Condition:Group	1.88	0.94	2	126	3.80	0.025	*
SCR Amplitude (log)	Condition	0.35	0.17	2	126	0.95	0.391	
	Group	0.81	0.81	1	63	4.45	0.039	*
	Condition:Group	0.85	0.42	2	126	2.32	0.102	
SCR Frequency	Condition	78.46	39.23	2	126	0.56	0.575	
	Group	161.46	161.46	1	63	2.29	0.135	
	Condition:Group	693.45	346.73	2	126	4.91	0.009	**
RMSSD	Condition	314.88	157.44	2	126	0.78	0.460	
	Group	105.29	105.29	1	63	0.52	0.472	
	Condition:Group	409.25	204.62	2	126	1.02	0.365	
SDNN	Condition	183.19	91.59	2	126	1.13	0.325	
	Group	62.79	62.79	1	63	0.78	0.382	
	Condition:Group	250.00	125.00	2	126	1.55	0.217	
PNN50	Condition	166.01	83.00	2	126	0.74	0.478	
	Group	141.35	141.36	1	63	1.26	0.265	
	Condition:Group	288.35	144.18	2	126	1.29	0.279	

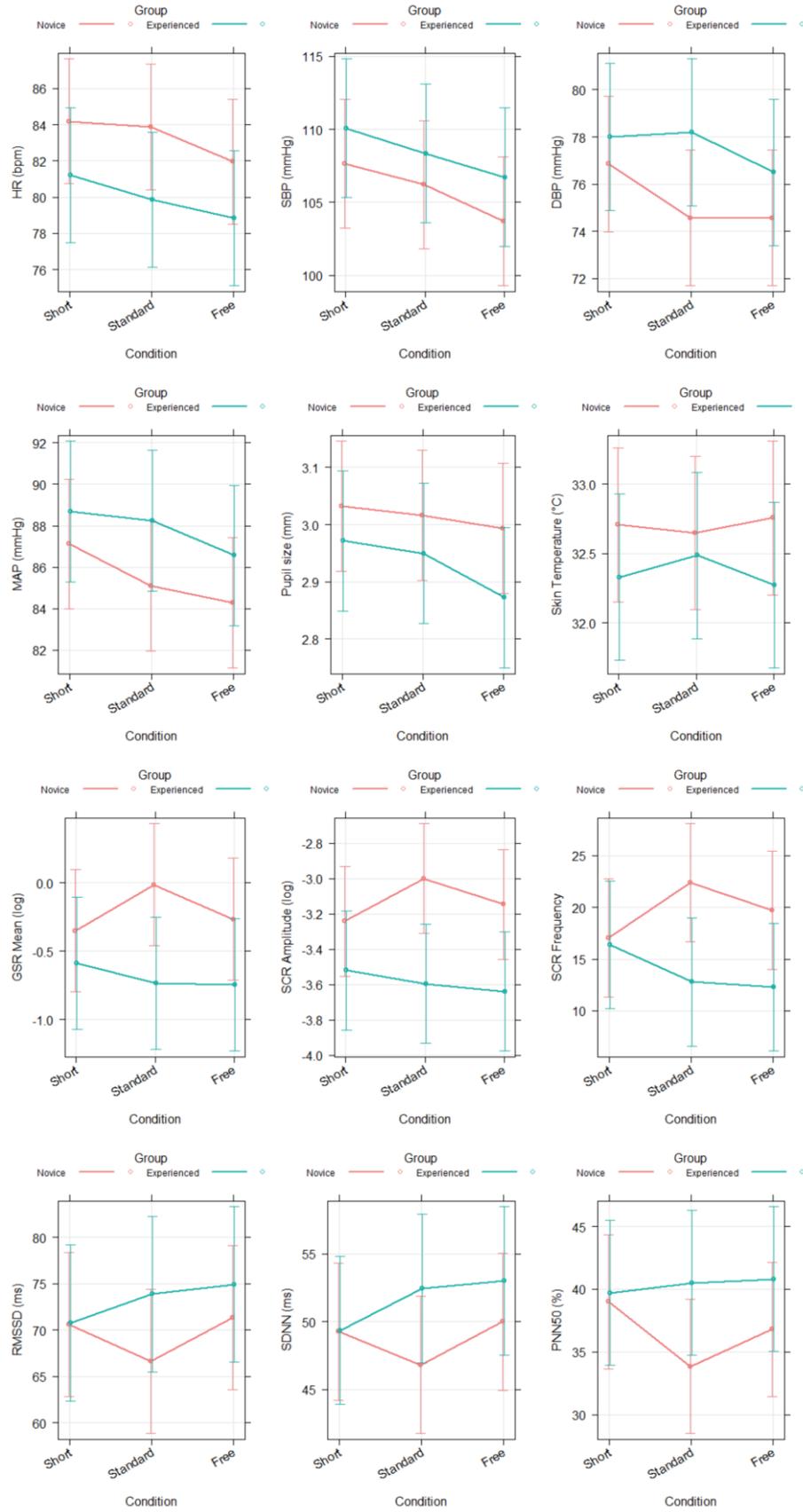


Figure 6-3. Effects of Condition and Group on the physiological measures and indices

Figure 6-3 shows that the biomarkers of HR, BP and pupil size of both groups displayed a positive relationship with the stringency of time condition. When the time constraint was getting more relaxed, the values of these measures decreased accordingly. This pattern corroborated H1. As shown in Table 6-2, Condition was statistically significant (in HR, SBP, MAP and pupil size) or marginally significant (in DBP) in these models. The measure of skin temperature displayed discrepant patterns for novice and experienced translators in terms of its variation with time condition (Figure 6-3). The effects of Condition, Group and their interaction were not statistically significant on this measure (Table 6-2). Thus, H1 was not corroborated by skin temperature.

In terms of the difference between the two groups, it showed that novice translators tended to have higher values of HR and pupil size (indicating a higher arousal level) than experienced translators regardless of the time conditions. However, experienced translators tended to have higher BP (SBP, DBP and MAP) and lower skin temperature values (indicating a higher arousal level) than novice translators regardless of the time conditions. It should be noticed that although discrepancies existed for these measures regarding the main effect of Group (shown in Figure 6-3), the difference between the two groups was not statistically significant – none of the measures mentioned above showed a significant effect of Group or the interaction effect between Group and Condition. No evidence was found that novice translators had a higher increment in arousal levels than experienced translators when the time condition was getting more stringent. Thus, H2 was not corroborated by the measures of HR, BP, pupil size and skin temperature.

As shown in Table 6-2, GSR Mean and SCR Frequency showed a statistically significant interaction between Condition and Group. This meant that the variation pattern of these indices modulated by time conditions was evidently different for the two groups of participants. The fixed effects of Condition and Group and their interaction effect on the HRV indices were not statistically significant; however, all the HRV indices showed a tendency consistent with that of the GSR indices (Figure 6-3).

Specifically, the GSR and HRV indices displayed different patterns for the two groups of participants. For experienced translators, the variation of these GSR and HRV indices across different time conditions was not evident – none of the post hoc comparisons between time conditions for the six indices was statistically significant for experienced translators. The tendency of the novice group, however, displayed a

remarkably different pattern across the three time conditions in that the arousal level peaked at the Standard session and lowered in the Short and Free sessions. The post hoc comparisons showed that the difference of GSR Mean and SCR Frequency between the Short and Standard conditions were statistically significant ($t(126)=-2.83, p=.016$ for GSR Mean; $t(126)=-2.64, p=.028$ for SCR Frequency). The difference between the Standard and Free conditions was also prominent for novices in GSR Mean ($t(126)=2.76, p=.020$). This tendency displayed by novice translators in which the Standard condition characterised a higher arousal level than the other two conditions was consistent in all the indices of GSR and HRV. This pattern echoes an inverted “U-shaped” relationship between time stringency and arousal. Thus, both H1 and H2 were not corroborated by GSR and HRV.

Figure 6-3 also shows that novice translators tended to have higher arousal levels indexed by GSR and HRV than experienced translators (consistent with HR and pupil size) regardless of the time conditions. However, only SCR Amplitude showed a statistically significant main effect of Group.

To sum up, H1 was corroborated by the biomarkers of HR, BP and pupil size for both novice and experienced translators, while H2 was not corroborated by any of the tested biomarkers. The indices of GSR and HRV displayed a consistent tendency in which experienced translators were not affected much by time condition, but novice translators displayed an inverted-U relation between time stringency and the arousal level indexed by these two measures. Since GSR and HRV were often considered indicative of (mental/cognitive) effort (see section 4.2.2.1), these two measures were inspected further together with other indicators of effort in section 6.4.

6.1.3 Correlation tests

In order to test H3, the repeated measures correlation tests for the measures that corroborated H1 were conducted. Table 6-3 below presents the results.

Table 6-3. Repeated measures correlation coefficients among the psycho-physiological measures that display a positive relationship with time stringency

	1.	2.	3.	4.	5.	6.	7.	8.
1. Stress	1							
2. State Anxiety	$r=.$.49***	1						
3. Temporal Demand	$r=.$.63***	$r=.$.50***	1					
4. HR	$r=.$ 03	$r=.$ 04	$r=.$.25**	1				
5. SBP	$r=.$.20*	$r=.$.23**	$r=.$.31***	$r=.$.29***	1			
6. DBP	$r=.$ 07	$r=.$ 02	$r=.$.21*	$r=.$ 03	$r=.$ 51***	1		
7. MAP	$r=.$ 15	$r=.$ 15	$r=.$.30***	$r=.$ 12	$r=.$ 80***	$r=.$ 85***	1	
8. Pupil size	$r=.$ 11	$r=.$ 06	$r=.$.26**	$r=.$.38***	$r=.$.22*	$r=.$ 08	$r=.$ 14	1

All the correlations among these measures were positive ($r > 0$). The three psychological measures (State Anxiety, Stress and Temporal Demand) were relatively strongly correlated with each other. In addition, Temporal Demand was moderately correlated with SBP and MAP, and weakly correlated with DBP, HR and pupil size. SBP was weakly correlated with Stress and State Anxiety. Within the group of the physiological measures, HR, SBP and pupil size were either moderately or weakly correlated with each other. The strong correlations among the three indices of BP are not discussed because they are inherently correlated.

Overall, H3 was partially, but to a great extent corroborated. Not all the tested physiological measures displayed a sufficiently significant correlation with one another and with the psychological measures; however, considerable (positive and significant) correlations were captured among these psycho-physiological measures.

6.1.4 Summary of RQ1

In summary, the results showed that the variations of the self-reported State Anxiety, Stress and Temporal Demand levels as well as the biomarkers of HR, BP and pupil size modulated by Condition displayed a consistent tendency with a statistical significance in which the values of these measures decreased accordingly with the time condition becoming more relaxed for both groups. The pattern shown by these measures corroborated H1. The measures of skin temperature, GSR and HRV did not corroborate H1 as the variations of these measures across time conditions were not in accordance with the stringency of the time constraints. In addition, H2 was not corroborated by any of the tested psycho-physiological measures. This meant that novice translators did not show stronger psycho-physiological responses or a higher increment in these measures

than experienced translators when the time condition was getting more stringent. H3 was partially but to a great extent corroborated since most measures that corroborated H1 (i.e., State Anxiety, Stress, Temporal Demand, HR, BP and pupil size) were correlated.

A different pattern concerning the effect of Condition was observed on the GSR and HRV indices. Although only two GSR indices (GSR Mean and SCR Frequency) showed a statistically significant interaction effect between Condition and Group, all the indices of GSR and HRV displayed the same pattern. The significant interaction effect indicated that the effect of Condition (on GSR Mean and SCR Frequency) was markedly different in the two groups of participants. Specifically, it was found that while an inverted-U relation between time stringency and arousal indexed by GSR and HRV was shown in novice translators, the GSR and HRV of experienced translators did not vary significantly across time conditions. Since these two measures displayed a consistent pattern but did not corroborate H1 in representing time pressure, possible implications of these measures were further explored in relation to effort measures in the following analysis.

Concerning the main effect of Group, only SCR Amplitude (one of the indices of GSR) demonstrated a significant main effect of Group showing that novices had higher values than experienced translators regardless of the time conditions. Further, despite the effects of Condition and Group in the translation task sessions, it showed that all the participants tended to be in a highly aroused state at the beginning of the experiment (Baseline 1), possibly due to feelings of nervousness and uncertainty, even though they were told to have a rest in this session. However, they recovered to reach a much more relaxed state at the end of the experiment (Baseline 2) when all the tasks were completed, as indicated by the change of most biomarkers under study.

6.2 The behavioural-cognitive effects

This section examines the effects of Condition and Group on the behavioural metrics (e.g., eye movements and keystroke activities) from the perspectives of automaticity (6.2.1) and cognitive rhythm (6.2.2). Apart from the two predictors of Condition and Group, another fixed effect, Area of Interest (AOI), is introduced in the analysis where relevant.

6.2.1 Automaticity

The degree of automaticity at the micro-level can be manifested primarily in the process of producing the TT products, i.e., in the typing activities. Therefore, automaticity at the micro-level was gauged based on the minimum unit of TT production – the TUs. Three aspects, namely processing speed, parallel processing degree and average TU effort were inspected. At the macro-level, the proportion of attention distributed to the ST/TT area of the screen was inspected. This section examines these aspects to answer RQ2 and to test the corresponding hypotheses below:

RQ2: What are the effects of time condition and experience on participants' level of automaticity during translation as indicated by the behavioural metrics?

H4. A more stringent time condition gives rise to more automated processing, evidenced by faster processing speed, deeper parallel processing, less effort in each processing unit, as well as an increased proportion of attention on TT production and a decreased proportion of attention on ST comprehension.

H5. Experienced translators have a higher automaticity level than novice translators, evidenced by faster processing speed, deeper parallel processing, less effort in each processing unit, as well as a higher proportion of attention on TT production than on ST comprehension.

Each aspect was examined based on one or more specific metric(s). Table 6-4 details these metrics and their operational explanations, which are also presented in detail in Chapter 4 (section 4.2.3.1 – 4.2.3.3). The following sections report the results regarding these aspects, respectively.

Table 6-4. Summary of the metrics related to automaticity in translation

Metric name	Explanation and calculation formula	Aspects
Keystroke production speed	total number of keystrokes ÷ drafting and end revision duration (min) of a task	Processing speed
Average IwTU	Average duration of Intervals within TUs	
Average IbtTU	Average duration of Intervals between TUs	
ST fixation duration % in TU	fixation duration on ST when producing a TU ÷ fixation duration on ST and TT when producing a TU × 100%	Parallel processing degree
ST-TT fixation ratio in TU	fixation duration on ST when producing a TU ÷ fixation duration on TT when producing a TU	
TU time	average time duration spent on a TU	Average TU effort
TU fixation	average fixation duration on ST and TT when producing a TU	
Fixation duration % on ST/TT in the drafting phase	fixation duration on ST/TT as a percentage of the total fixation duration on ST and TT during the drafting phase	Attention allocation

6.2.1.1 Processing speed

Table 6-5 shows that the metrics of keystroke production speed, average IwTU and average IbtTU demonstrated a significant effect of Condition. Figure 6-4 visualises the tendencies of these metrics across time conditions and groups.

Table 6-5. Type III ANOVA table of the LMER models of *Keystroke production speed, Average IwTU and Average IbtTU*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	p value	Sig.
Keystroke production speed	Condition	10996.9	5498.5	2	126	34.51	<.001	***
	Group	6.7	6.7	1	63	0.04	0.838	
	Condition:Group	946.1	473.1	2	126	2.97	0.055	
Average IwTU	Condition	13196.7	6598.4	2	126	17.59	<.001	***
	Group	0	0	1	63	0.00	0.994	
	Condition:Group	1706.9	853.4	2	126	2.28	0.107	
Average IbtTU	Condition	2.12E+07	1.06E+07	2	126	23.24	<.001	***
	Group	6.25E+05	6.25E+05	1	63	1.37	0.247	
	Condition:Group	1.32E+06	6.62E+05	2	126	1.45	0.239	

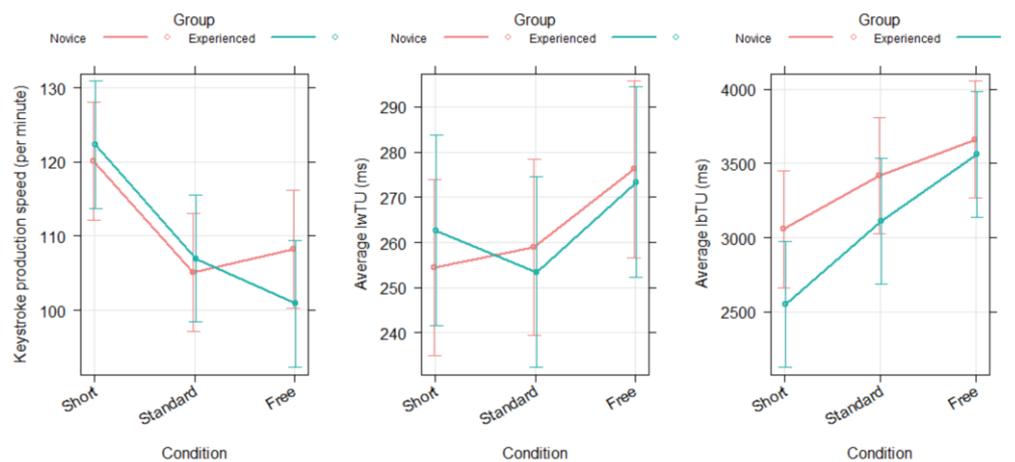


Figure 6-4. Effects of Condition and Group on *Keystroke production speed*, *Average IwTU* and *Average IbTU*

A general tendency for the three metrics was that when the time condition was getting more relaxed, both novice and experienced translators slowed down their processing speed with fewer keystrokes produced per minute and longer between-TU and within-TU intervals. It should be noted that for the metric of keystroke production speed, it was the most stringent time condition that made a difference (i.e., there was no statistically significant difference captured between the Standard and Free conditions for both groups). For the metric of IwTU, it was the most relaxed time condition that made the difference (i.e., there was no statistically significant difference captured between the Short and Standard conditions for both groups). For the metric of IbTU, however, both groups gradually prolonged the intervals between TUs when the time condition was getting more relaxed. Despite these slight divergencies, their overall tendency across time conditions corroborated H4, meaning that a stringent time condition could increase the processing speed of both groups.

None of the three metrics showed a statistically significant effect of Group, but the interaction between Condition and Group was marginally significant on the metric of keystroke production speed. This meant that the two groups of participants exhibited slightly different patterns on this metric with the change of time condition. It shows in Figure 6-4 that experienced translators decreased the number of keystrokes per minute gradually when the time condition changed from the most stringent (Short) to the most relaxed (Free) one; novice translators, however, slightly increased, rather than decreased their production speed from the Standard to the Free condition; but as

mentioned, this slight increase was not statistically significant.

Generally, experienced translators maintained faster processing speed than novice translators, as manifested by all the three metrics, with exceptions only for keystroke production speed in the Free condition and IwTU in the Short condition. However, the effect of Group was not statistically significant in all the models, which meant that experienced translators did not process significantly faster than novice translators during translation regardless of the time conditions. Hence, H5 was not corroborated by these three metrics.

6.2.1.2 Parallel processing degree

Table 6-6 presents the results for the two metrics reflecting the parallel processing degree, and Figure 6-5 visualises how the metrics vary across time conditions and groups.

Table 6-6. Type III ANOVA table of the LMER models of *ST fixation duration % in TU* and *ST-TT fixation ratio in TU*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
ST fixation duration % in TU	Condition	657.87	328.94	2	108	6.27	0.003	**
	Group	9.37	9.37	1	54	0.18	0.674	
	Condition:Group	36.84	18.42	2	108	0.35	0.705	
ST-TT fixation ratio in TU	Condition	0.43	0.21	2	108	7.07	0.001	**
	Group	0.02	0.02	1	54	0.58	0.448	
	Condition:Group	0.05	0.02	2	108	0.77	0.466	

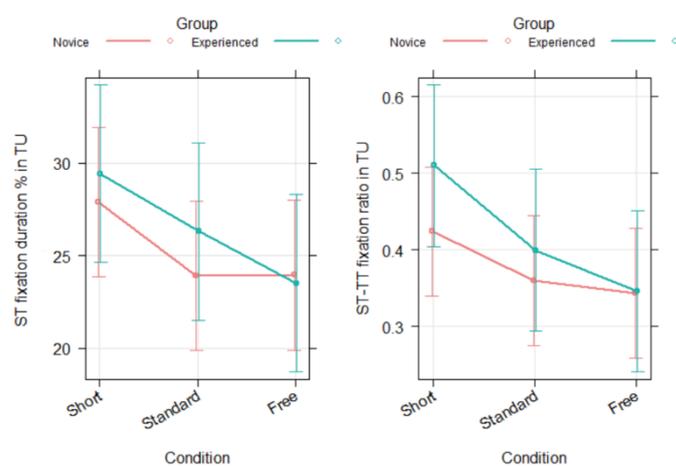


Figure 6-5. Effects of Condition and Group on *ST fixation duration % in TU* and *ST-TT fixation ratio in TU*

Table 6-6 reveals that the effect of Condition was significant on both metrics. Generally, a more stringent time condition led to a higher proportion of attention on ST when producing a TU and a higher ST-TT fixation ratio when producing a TU, indicating deeper parallel processing for both groups of participants. Specifically, for experienced translators, the estimated marginal means of *ST fixation duration % in TU* and *ST-TT fixation ratio in TU* were 29.4% (SE¹⁶=2.40) and 0.51 (SE=0.05) respectively in the Short condition; the values decreased to 26.3% (SE=2.40) and 0.399 (SE=0.05) respectively in the Standard condition and further reduced to 23.5% (SE=2.40) and 0.346 (SE=0.05) respectively in the Free condition. Only the contrasts between the Short and Free conditions were statistically significant ($t(108)=3.138$, $p=.007$ and $t(108)=2.703$, $p=.024$, respectively). This tendency featured a gradual decrease in parallel processing degree with the alleviation of time stringency for experienced translators. Overall, this result corroborated H4.

For novice translators, there was an evident decrease from the Short condition to the Standard condition with the estimated marginal means of *ST fixation duration % in TU* and *ST-TT fixation ratio in TU* reducing from 27.7% (SE=1.93) and 0.424 (SE=0.04) to 24.5% (SE=1.93) and 0.359 (SE=0.04), respectively. There was a slight decrease from the Standard condition to the Free condition with the estimated marginal means of the two metrics maintaining at 23.8% (SE=1.93) and 0.343 (SE=0.04) respectively which were close to that of the Standard condition. Also, none of the pairwise comparisons of these values between time conditions was significant for novice translators. Thus, novice translators demonstrated a less prominent effect of time condition on these metrics than experienced translators.

Although experienced translators generally showed a higher degree of parallel processing, especially in the time-constrained conditions, than novice translators, as shown in both metrics (Figure 6-5), H5 was not corroborated by these two metrics since there was no statistically significant effect of Group on the two metrics.

6.2.1.3 Average TU effort

Before examining the two metrics of TU effort measured by time and fixation duration, the average number of keystrokes in a TU was inspected with regard to the effects of

¹⁶ SE=Standard Error

Condition and Group. Table 6-7 below reports the results for TU length and the two metrics of TU effort measured by time and fixation duration.

Table 6-7. Type III ANOVA table of the LMER models of *TU length (keystrokes)*, *TU time*, and *TU fixation*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
TU length (keystrokes)	Condition	0.18	0.09	2	126	0.76	0.469	
	Group	0.35	0.35	1	63	2.88	0.095	.
	Condition:Group	0.08	0.04	2	126	0.34	0.713	
TU time	Condition	248290	124145	2	126	7.75	<.001	***
	Group	20245	20245	1	63	1.26	0.265	
	Condition:Group	34269	17134	2	126	1.07	0.346	
TU fixation	Condition	67019	33510	2	108	1.67	0.193	
	Group	3337	3337	1	54	0.17	0.685	
	Condition:Group	87462	43731	2	108	2.18	0.118	

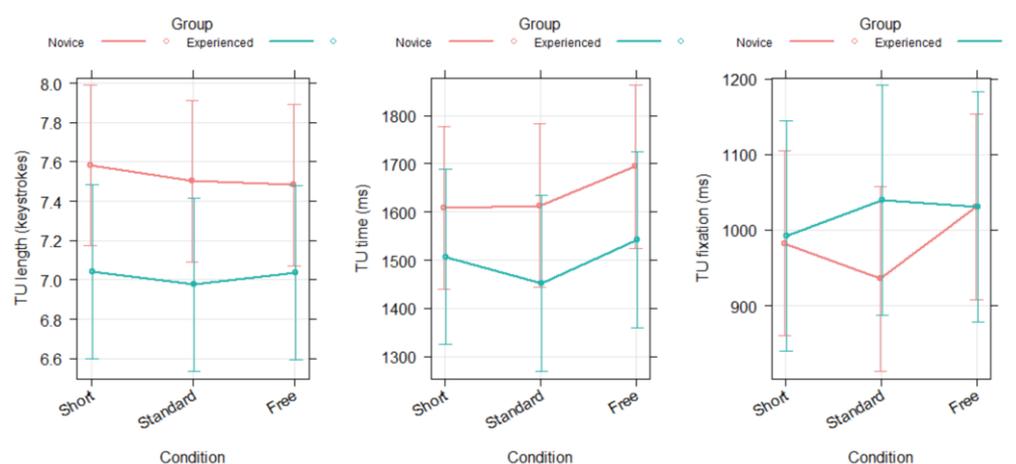


Figure 6-6. Effects of Condition and Group on *TU length (keystrokes)*, *TU time*, and *TU fixation*

The results showed that the average number of keystrokes in a TU was not affected by Condition or Group, although novices tended to have slightly longer TUs with more keystrokes than experienced translators (i.e., novices prefer to type slightly more Chinese characters in a TU) regardless of the time conditions. This indicated that such habitual typing behaviours were not easily affected by time condition.

The metric of TU time was significantly affected by Condition, and the metric of TU fixation was not affected by either Condition or Group or their interaction. The tendency in Figure 6-6 reveals that the amount of time spent per TU by novices in the

Short session (estimate=1609, SE=85.7) approximated that of the Standard session (estimate=1614, SE=85.7) and was lower than that of the Free session (estimate=1694, SE=85.7). Both the contrasts between Short and Free ($t(126)=-2.829$, $p=.016$) and between Standard and Free condition ($t(126)=-2.664$, $p=.026$) were statistically significant. The amount of time per TU spent by experienced translators, however, plummeted in the Standard session (estimate=1452, SE=92.5) from the Short session (estimate=1507, SE=92.5) and peaked at the Free session (estimate=1542, SE=92.5). Only the contrast between the Standard and the Free condition was statistically significant ($t(126)=2.755$, $p=.020$).

This result indicated that the effect of Condition arose mainly from the presence/absence of a deadline. That is, for both groups, the time spent on each TU significantly increased when the deadline was removed (i.e., in the Free session), but the contrast between conditions with different levels of time stringency (i.e., between the Short and the Standard session) was not evident. As such, H4 was corroborated by the metric of TU time in terms of the presence/absence of a deadline. This meant that a deadline was likely to cut down the effort (measured by time) spent on each TU. Although experienced translators tended to spend less time (but longer fixation duration) on each TU than novice translators, the group difference was not statistically significant. Therefore, H5 was not corroborated by TU time or TU fixation.

6.2.1.4 Attention allocation

At the macro-level, automaticity can be gauged by how the participants allocate their visual attention to the two processing areas, i.e., ST and TT, during the drafting phase. Therefore, the metric of the proportion of fixation duration that was allocated to ST/TT was examined. The main effect of AOI, which contained two levels, ST and TT, was introduced to the model to exhibit how the participants allocated their attention. To test H4, the model with AOI and Condition as fixed effects was built firstly. Table 6-8 displays the results, and Figure 6-7 visualises the effects.

Table 6-8. Type III ANOVA table of the LMER models of *Fixation duration % on ST/TT in drafting phase* with AOI and Condition as fixed effects

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	p value	Sig.
Fixation duration % on ST/TT in drafting phase	AOI	1606.75	1606.75	1	330	10.44	0.001	**
	Condition	0	0	2	330	0	1	
	AOI:Condition	272.91	136.46	2	330	0.89	0.413	

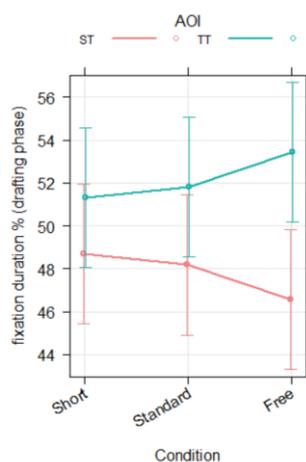


Figure 6-7. Effects of Condition and AOI on *Fixation duration % on ST/TT in the drafting phase*

The results in Table 6-8 show that Condition did not affect the amount of attention allocated to ST/TT during the drafting phase, and nor did it interact with the effect of AOI. Thus, the stringency of time condition could not increase the proportion of attention invested in TT production or reduce the proportion of attention invested in ST comprehension. Consequently, H4 was not corroborated by this metric. Figure 6-7 shows that although the effect of Condition was not statistically significant, participants tended to slightly increase the proportion of attention on TT production and reduce the proportion of attention on ST comprehension when the stringency of time condition was alleviated. Only the main effect of AOI made a significant difference, which meant that the participants tended to invest more attention in TT production than in ST comprehension regardless of the time conditions (Figure 6-7).

To test H5, a model with the fixed effects of AOI and Group was built. Table 6-9 presents the results.

Table 6-9. Type III ANOVA table of the LMER model of *Fixation duration % on ST/TT in drafting phase* with AOI and Group as fixed effects

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
Fixation duration % on ST/TT in drafting phase	AOI	2323.7	2323.7	1	332	15.66	<.001	***
	Group	0	0	1	332	0	1	
	AOI:Group	1784.2	1784.2	1	332	12.02	<.001	***

Table 6-9 shows that the main effect of AOI and its interaction effect with Group were statistically significant. This indicated that novice and experienced translators had different attention allocation patterns during the drafting phase. Figure 6-8 below displays the patterns.

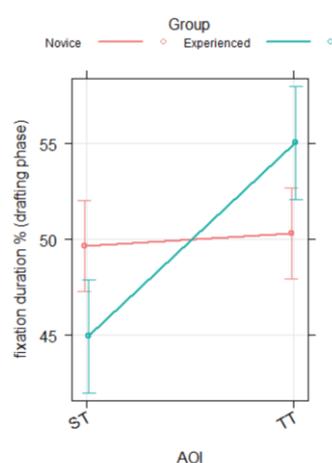


Figure 6-8. Effects of Group and AOI on *Fixation duration % on ST/TT in the drafting phase*

It is evident that experienced translators tended to have a considerably lower proportion of attention on ST (estimate=44.9, SE=1.50) than on TT (estimate=55.1, SE=1.50) during the drafting phase ($t(278) = -4.764$, $p < .001$). Novice translators, however, tended to have comparable proportions of attention on ST (estimate=49.7, SE=1.21) and on TT (estimate=50.3, SE=1.21) during the drafting phase ($t(278) = -0.391$, $p = .696$). In addition, novices spent a significantly higher proportion of fixation than experienced translators on the ST area ($t(181) = 2.452$, $p = .015$) and correspondingly, a significantly lower proportion of fixation than experienced translators on the TT area ($t(181) = -2.452$, $p = .015$).

This result corroborated H5 in that experienced translators spent much more attention on TT production than on ST comprehension during the drafting phase, and

this difference was more salient than that of novice translators.

Attention distributed to the passing of time

The number of fixations distributed to the timing AOI (i.e., TM) of the screen (*Fixation count on TM*) was examined to check if time condition and experience made a difference in the attention paid to the passing of time. The results showed that although both novice and experienced translators reduced their fixations on the AOI of TM in the Free condition (Figure 6-9), there was no statistical significance captured for Condition or Group. This result indicated that novice and experienced translators behaved similarly regarding the amount of attention paid to how much time was remaining or used during the translation process. They tended to pay slightly more attention to this timing area when there was a deadline (in the Short or Standard condition), which meant that they attended more to how much time was remaining. When the deadline was removed, i.e., in the Free condition, they paid less attention to this timing area, which meant that they cared less about how much time they had consumed in this task.

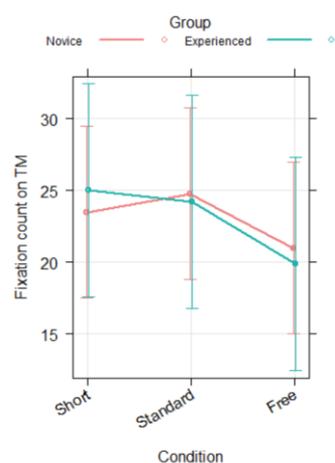


Figure 6-9. Effects of Condition and Group on *Fixation count on TM*

6.2.1.5 Summary of RQ2

In summary, in response to RQ2 (*What are the effects of time condition and experience on participants' level of automaticity during translation as indicated by the behavioural metrics?*), this section examined four aspects of automaticity in translation, namely, processing speed, parallel processing degree, average TU effort and the (proportion of)

attention allocated to ST/TT. Specific metric(s) corresponding to these aspects were tested against H4 and H5, respectively. Table 6-10 below summarises the results. Basically, time condition mainly affected automaticity from the perspectives of processing speed, parallel processing degree and the ease/effort of processing each TU (measured by time). At the same time, experience came with a salient effect on the (proportion of) attention allocated to ST/TT. Although variation patterns for some metrics (e.g., Keystroke production speed, Average IwTU, and TU time) did not conform to a strict linear positive or negative relationship with the stringency of time condition, the overall effect of Condition was significant in those models of metrics that corroborated H4. This means that at least some comparisons between time conditions, if not all the pairwise comparisons, were prominent enough to elicit the significant changes (e.g., for TU time, it is the existence of deadline that makes a difference).

Table 6-10. Summary of the results for H4 and H5

Aspects	Metric name	H4	H5
Processing speed	Keystroke production speed	✓	✗
	Average IwTU	✓	✗
	Average IbTU	✓	✗
Parallel processing degree	ST fixation duration % in TU	✓	✗
	ST-TT fixation ratio in TU	✓	✗
Average TU effort	TU time	✓	✗
	TU fixation	✗	✗
Attention allocation	Fixation duration % on ST/TT in the drafting phase	✗	✓

6.2.2 Cognitive rhythm

Cognitive rhythm can be manifested in the way translators coordinate the subtasks during translation. At the micro-level, cognitive rhythm was examined from the aspects of the pauses and segments' features. At the macro-level, cognitive rhythm was examined by the amount of time distributed to different translation phases and the revision preference patterns observed in different time conditions. This section reports the results in response to RQ3 by testing the corresponding hypotheses as listed below:

RQ3: What are the effects of time condition and experience on participants' cognitive rhythm during translation as indicated by the behavioural metrics?

H6. A more stringent time condition would shorten and reduce the number of all types of pauses.

H7. Time condition would affect the segment length differently for novice and experienced translators.

H8. When the time condition becomes more stringent, the orientation and end revision phases are more likely to be trimmed than the drafting phase.

H9. Time condition would affect the pattern of preference for online/end revision differently for novice and experienced translators.

Table 6-11 presents the detailed metrics related to these hypotheses, which are examined in this section.

Table 6-11. Summary of the metrics related to cognitive rhythm in translation

Metric name	Explanation and calculation formula	Aspects
SP count & length	total number and average duration of SPs	Pauses
MP count & length	total number and average duration of MPs	
LP count & length	total number and average duration of LPs	
Segment length (time)	average duration of a segment	Segments
Segment length (TU number)	average number of TUs in a segment	
Orientation duration & percentage	time duration of the orientation phase and time duration of orientation phase as a percentage of the total task time	Translation phases
Drafting duration & percentage	time duration of the drafting phase and time duration of the drafting phase as a percentage of the total task time	
End revision duration & percentage	time duration of the end revision phase and time duration of the end revision phase as a percentage of the total task time	
Online deletions	number of deletion keystrokes in the drafting phase	Types of revision
End revision deletions	number of deletion keystrokes in the end revision phase	behaviour

6.2.2.1 Pauses

The number and length of three types of pauses, i.e., SP, MP and LP, were examined regarding the effects of Condition and Group. Table 6-12 presents the effects of Condition and Group as well as their interaction on the six metrics; Figure 6-10 demonstrates the effects.

Table 6-12. Type III ANOVA table of the LMER models of SP, MP and LP count and length

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	p value	Sig.
SP count	Condition	340055	170027	2	126	36.37	<.001	***
	Group	829	829	1	63	0.18	0.675	
	Condition:Group	2307	1153	2	126	0.25	0.782	
MP count	Condition	113374	56687	2	126	18.12	<.001	***
	Group	889	889	1	63	0.28	0.596	
	Condition:Group	5140	2570	2	126	0.82	0.442	
LP count	Condition	6789.2	3394.6	2	126	18.42	<.001	***
	Group	227.8	227.8	1	63	1.24	0.271	
	Condition:Group	603.6	301.8	2	126	1.64	0.199	
SP length	Condition	579.85	289.923	2	126	3.81	0.025	*
	Group	48.96	48.956	1	63	0.64	0.425	
	Condition:Group	88.46	44.23	2	126	0.58	0.560	
MP length	Condition	133208	66604	2	126	7.62	<.001	***
	Group	7556	7556	1	63	0.86	0.356	
	Condition:Group	2312	1156	2	126	0.13	0.876	
LP length	Condition	6.96E+07	3.48E+07	2	126	5.55	0.005	**
	Group	9.00E+06	9.00E+06	1	63	1.43	0.236	
	Condition:Group	4.44E+06	2.22E+06	2	126	0.35	0.703	

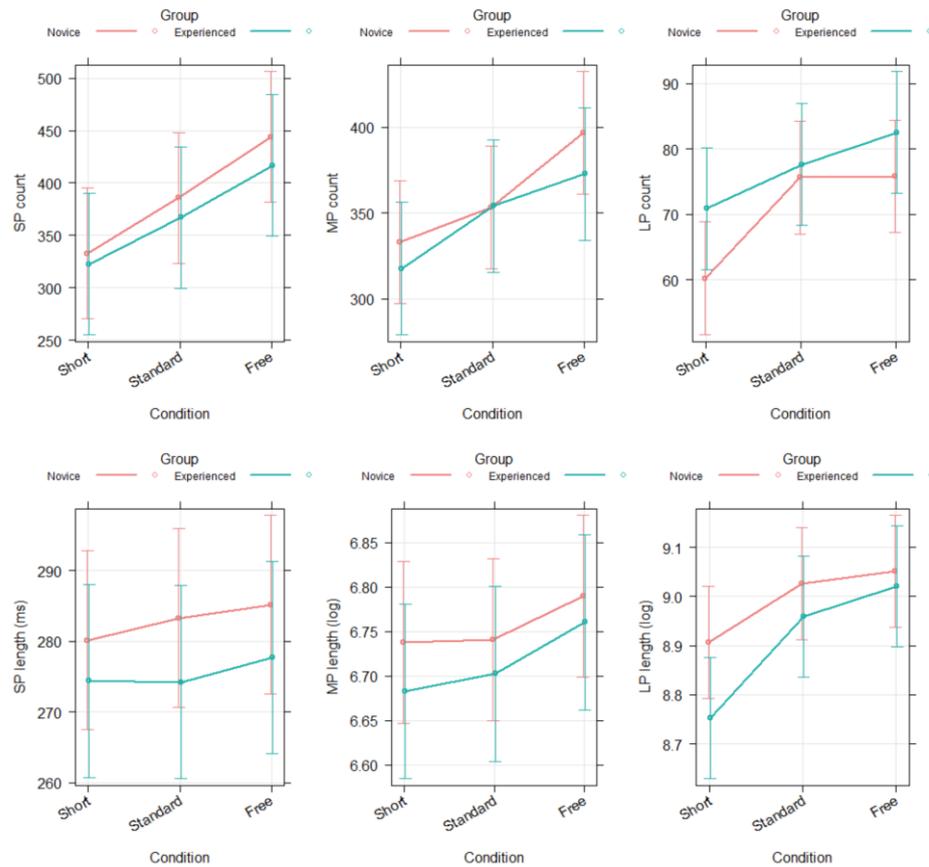


Figure 6-10. Effects of Condition and Group on SP, MP and LP count and length

The results in Table 6-12 show that the effect of Condition was significant in affecting the number and length of all the three types of pauses; the effects of Group and the interaction of Condition and Group were not evident enough to reach a statistical significance on these metrics. Their variations across time conditions were generally consistent for the two groups: with the time condition being increasingly stringent, the number and length of all the three types of pauses were reduced and shortened. Thus, H6 was corroborated.

It is revealed by Figure 6-10 that novices tended to have slightly longer and more SPs and MPs than experienced translators regardless of the time conditions; they also tended to have longer but fewer LPs than experienced translators. In addition, despite the failure of capturing a significant interaction effect between Condition and Group, the tendency shown by LP count is worth noticing. Novice translators made almost the same number of LPs in the Standard condition as in the Free condition, while experienced translators managed to keep increasing the number of LPs when the time stringency was alleviated from the Standard condition to the Free condition.

It is also of interest to examine the actual values of the three types of pauses. The estimated means of SP length were 280, 283 and 285 milliseconds for novice translators in the Short, Standard and Free conditions, respectively, and 274, 274 and 278 milliseconds for experienced translators in the three conditions, respectively. The estimated means of MP length were 892, 884 and 947 milliseconds for novice translators, and 822, 829 and 878 milliseconds for experienced translators in the three conditions, respectively. The estimated means of LP length were 8202, 8762 and 9387 milliseconds for novice translators, and 6791, 8065 and 8508 milliseconds for experienced translators in the three time conditions respectively. These values of SP, MP and LP lengths generally fell within the pause ranges mentioned in Muñoz Martín and Martín de Leon's (2018) study.

To further explore whether Condition and Group affected the distribution of MPs and LPs in different translation phases, the metrics of MP and LP count in the drafting and end revision phases were examined (Table 6-13 and Figure 6-11).

Table 6-13. Type III ANOVA table of the LMER models of MP and LP count in the drafting and end revision phase

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
MP count (drafting)	Condition	19799	9899.6	2	126	6.05	0.003	**
	Group	418	418	1	63	0.26	0.615	
	Condition:Group	10453	5226.4	2	126	3.20	0.044	*
LP count (drafting)	Condition	2341.35	1170.68	2	126	6.63	0.002	**
	Group	140.76	140.76	1	63	0.80	0.375	
	Condition:Group	1023.57	511.78	2	126	2.90	0.059	
MP count (end revision)	Condition	38468	19234	2	126	12.18	<.001	***
	Group	118	117.6	1	63	0.07	0.786	
	Condition:Group	1312	656.1	2	126	0.42	0.661	
LP count (end revision)	Condition	1193.43	596.72	2	126	10.66	<.001	***
	Group	28.38	28.38	1	63	0.51	0.479	
	Condition:Group	439.87	219.94	2	126	3.93	0.022	*

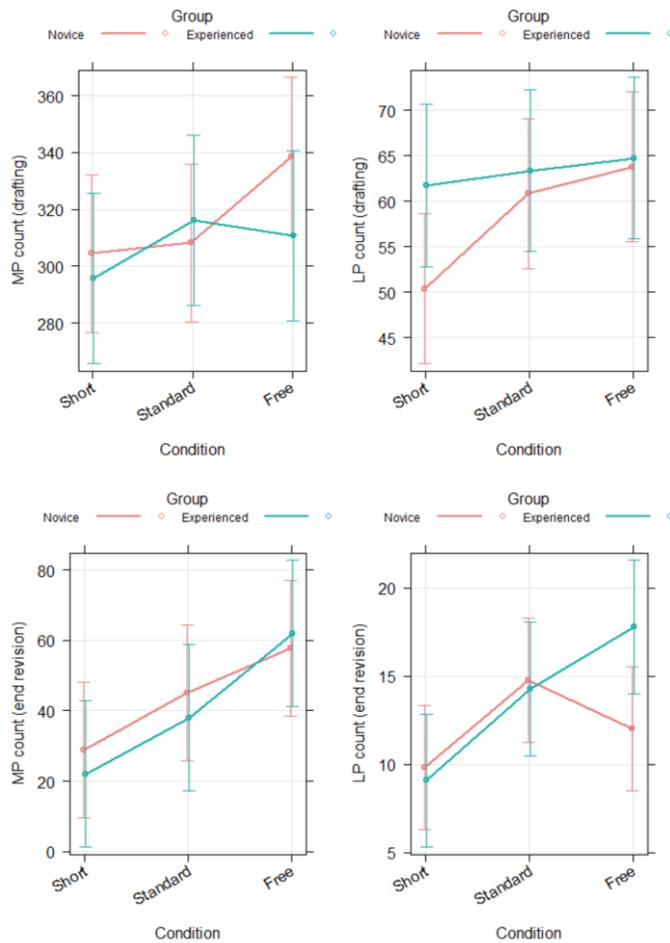


Figure 6-11. Effects of Condition and Group on MP and LP count in the drafting and end revision phase

The results indicated that the alleviation of time stringency constantly increased

MPs in the end revision phase for both groups. However, the interaction effect between Condition and Group was statistically significant or marginally significant in MP and LP count in the drafting phase and LP count in the end revision phase. As shown in Figure 6-11, novice translators tended to make more MPs and LPs in the drafting phase with the alleviation of time stringency, which signaled increased occurrences of monitoring and problem-solving/planning activities in the drafting phase. However, the number of LPs in the end revision phase which flag the problem-solving/planning activities faltered in the Free condition (compared with that in the Standard condition). Experienced translators tended to have a relatively stable number of MPs and LPs in the drafting phase across the three time conditions, which signaled a steady rhythm in terms of the occurrences of monitoring and problem-solving/planning activities in the drafting phase. However, they continuously and remarkably increased the number of LPs in the end revision phase with the alleviation of time stringency. Thus, time stringency had a more evident effect on novice translators than on experienced translators in the drafting phase regarding the occurrences of monitoring and problem-solving/planning activities. Time stringency also affected the problem-solving/planning activities in the end revision phase of both novice and experienced translators but in a different manner.

In summary, the effect of time condition on the number and length of the three types of pauses was consistent for the two groups of participants: the more time was given, the more and the longer the pauses they tended to make. Thus, H6 was corroborated. In addition, the average length of all the three types of pauses was longer for novices than for experienced translators, although this difference was not statistically significant. Overall, novices tended to have more SPs and MPs, while experienced translators tended to have more LPs regardless of the time conditions. Further, experienced translators continuously increased the LPs (especially in the end revision phase) when time pressure was alleviated. Novices made almost the same number of LPs in the Free condition as in the Standard condition. Further exploration of the distribution of MPs and LPs in the drafting and end revision phases showed that novice and experienced translators displayed different patterns in these two metrics. Novices tended to increase LPs in the drafting phase, while experienced translators tended to increase LPs in the end revision phase when time pressure was alleviated.

6.2.2.2 Segments

The length of segments was measured by their average time duration and average number of TUs, respectively. Table 6-14 shows that neither Condition nor Group had an evident effect on segment length measured by time; in contrast, Condition significantly affected the average number of TUs per segment. Figure 6-12 visualises the effects.

Table 6-14. Type III ANOVA table of LMER models of *Segment length (time)* and *Segment length (TU count)*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
Segment length (time)	Condition	4.29E+08	2.15E+08	2	126	1.62	.201	
	Group	2.08E+08	2.08E+08	1	63	1.57	.214	
	Condition:Group	2.63E+08	1.32E+08	2	126	0.99	.373	
Segment length (TU number)	Condition	12.25	6.12	2	126	8.12	<.001	***
	Group	0.71	0.71	1	63	0.94	.336	
	Condition:Group	2.55	1.27	2	126	1.69	.189	

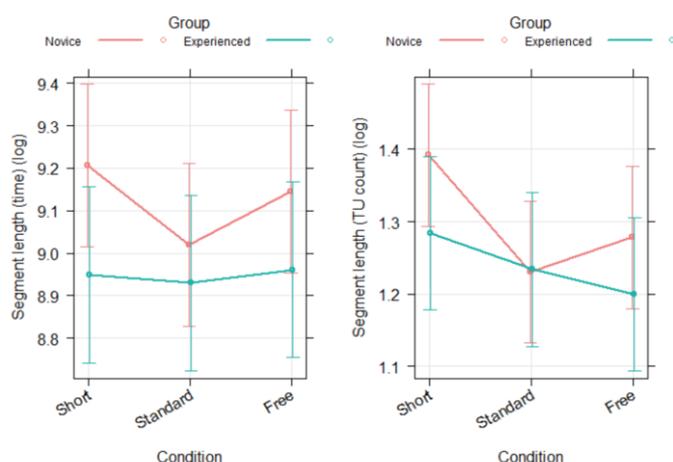


Figure 6-12. Effects of Condition and Group on *Segment length (time)* and *Segment length (TU count)*

It showed a pronounced reduction in segment length (measured by both time and TU number) for novice translators from the Short to the Standard and from the Short to the Free condition. Experienced translators, however, showed little variation throughout the three conditions in segment length. In addition, regardless of the time conditions, novice translators seemed to have longer segments than experienced translators, especially in the Short condition, although the overall difference was not statistically significant.

A follow-up pairwise comparison was conducted for the statistically significant effect of Condition on segment length measured by time and TU numbers. As expected, for novices, the significant change in segment length (time) occurred at the Short-Standard ($t(126)=3.39, p=.003$) contrast. For segment length (TU number), the significant changes occurred at the Short-Standard ($t(126)=4.38, p<.001$) and Short-Free ($t(126)=3.08, p=.008$) contrasts. None of the contrasts between time conditions for experienced translators was significant. Thus, H7 (*Time condition would affect the segment length differently for novice and experienced translators.*) was corroborated by the metric of segment length (TU numbers) since a stringent time condition (Short condition) prominently increased the TU number in a segment for novice translators, whilst the TU number of a segment produced by experienced translators was affected little by the stringency of time condition.

6.2.2.3 Translation phases

Before inspecting the effects of Condition and Group on translation phases, the amount of time used in the Free condition by the two groups of participants was examined. The average time duration of the Free session was 22.77 minutes (SD=5.50) for novice translators and 22.62 minutes (SD=4.16) for experienced translators. There were 23 out of 35 (65.71%) novice translators, and 22 out of 30 (73.33%) experienced translators who spent more time in the Free session than in the Standard session. Overall, most participants consumed more time in the condition without a deadline than in the Standard condition with a deadline of 20 minutes 25 seconds.

Time duration of each translation phase

The amount of time used in each translation phase was examined, and it showed that Condition prominently affected the length of the drafting and the end revision phases while Group affected the length of the orientation phase (Table 6-15). As shown in Figure 6-13, the tendencies tell that experienced translators had a statistically significantly longer orientation phase than novice translators regardless of the time conditions. The drafting phase was prolonged with the alleviation of time stringency for both groups. A discrepancy emerged between the two groups for the duration of the end revision phase in the Free condition: while experienced translators increased their end revision time further in the Free condition compared with that in the Standard

condition, novice translators reduced their end revision time in the Free condition.

Table 6-15. Type III ANOVA table of the LMER models of Orientation, Drafting and End revision duration

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
Orientation duration (log)	Condition	0.80	0.40	2	126	1.00	0.371	
	Group	1.97	1.97	1	63	4.91	0.030	*
	Condition:Group	1.71	0.85	2	126	2.13	0.124	
Drafting duration	Condition	1.64E+12	8.20E+11	2	126	49.95	<.001	***
	Group	3.71E+10	3.71E+10	1	63	2.26	0.138	
	Condition:Group	3.60E+10	1.80E+10	2	126	1.10	0.338	
End revision duration	Condition	8.51E+11	4.25E+11	2	126	22.58	<.001	***
	Group	6.40E+09	6.40E+09	1	63	0.34	0.562	
	Condition:Group	7.55E+10	3.78E+10	2	126	2.00	0.139	

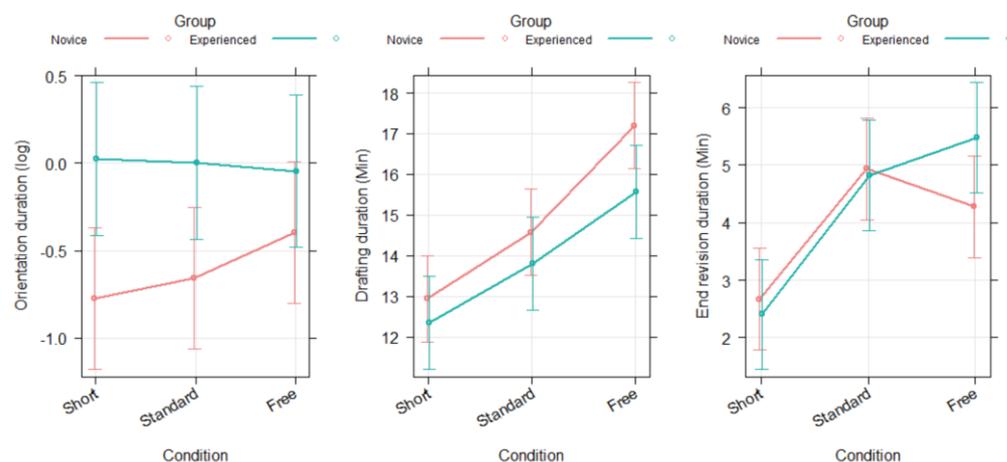


Figure 6-13. Effects of Condition and Group on Orientation, Drafting and End revision duration

These results indicated that the stringent time condition significantly shortened the time duration of the drafting phase and the end revision phase (except for the end revision phase of novice translators in the comparison between the Standard and Free condition), while the length of the orientation phase was not prominently affected by time condition. Therefore, H8 (*When the time condition becomes more stringent, the orientation and end revision phases are more likely to be trimmed than the drafting phase.*) was not corroborated.

The proportion of time consumed by each translation phase

The proportion of time distributed to the three translation phases were also examined. The results showed that Group pronounced a statistically significant effect on the percentage of time spent on the orientation phase, while Condition mainly influenced the percentage of time spent on the drafting and end revision phases. There was a marginal interaction effect between Condition and Group on the percentage of the end revision time (Table 6-16 and Figure 6-14).

Table 6-16. Type III ANOVA table of the LMER models of Orientation, Drafting and End revision percentage

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
Orientation percentage	Condition	21.47	10.74	2	126	0.98	0.379	
	Group	47.68	47.68	1	63	4.34	0.041	*
	Condition:Group	44.80	22.40	2	126	2.04	0.134	
Drafting percentage	Condition	1926.22	963.11	2	126	11.77	<.001	***
	Group	217.07	217.07	1	63	2.65	0.108	
	Condition:Group	236.34	118.17	2	126	1.44	0.240	
End revision percentage	Condition	2340.46	1170.23	2	126	13.43	<.001	***
	Group	23.27	23.27	1	63	0.27	0.607	
	Condition:Group	463.46	231.73	2	126	2.66	0.074	

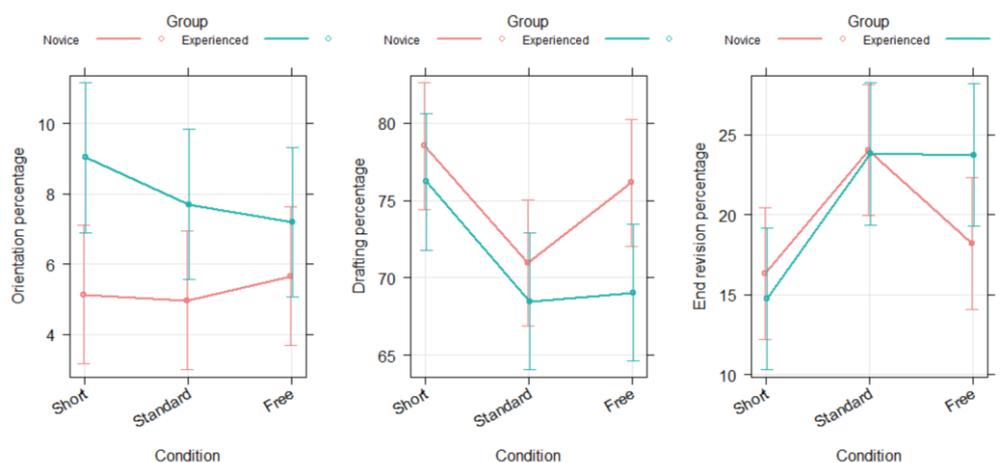


Figure 6-14. Effects of Condition and Group on Orientation, Drafting and End revision percentage

Specifically, compared with novice translators, experienced translators tended to invest a higher proportion of time in the orientation phase regardless of the time conditions, and the contrasts between the two groups were significant in the Short condition ($t(97.5)=-2.651, p=.009$) and marginally significant in the Standard condition

($t(97.5)=-1.859, p=.066$). Both groups were not significantly affected by time condition regarding the percentage of the orientation time.

Both the percentages of the drafting time and end revision time were prominently affected by time condition. The two groups showed a very similar tendency in these two metrics when the time condition changed from the Short to the Standard condition: the drafting percentage decreased, and the end revision percentage increased evidently in the Standard condition. When the stringency of time condition was further alleviated, i.e., comparing the Standard with the Free condition, there was a discrepant tendency for the two groups: novice translators tended to increase the drafting percentage and reduce the end revision percentage in the Free condition; experienced translators tended to maintain the similar drafting and end revision percentages in the Standard and Free conditions.

This meant that novice translators' pattern of time distribution to the drafting and end revision phases in the Free condition resembled that in the Short condition (high percentage in drafting and low percentage in end revision), whereas experienced translators' pattern of time distribution to these two phases in the Free condition resembled that in the Standard condition (high percentage in end revision and low percentage in drafting). The follow-up comparison of this group difference (Novice-Experienced) in the Free condition showed that the contrast for the drafting percentage ($t(133)=2.323, p=.022$) was statistically significant, and the contrast for the end revision percentage ($t(138)=-1.808, p=.073$) was marginally significant.

Thus, experienced translators managed to maintain a certain proportion of end revision when the deadline was removed (i.e., in the Free condition). However, novice translators chose to increase the proportion of time taken by the drafting phase and, accordingly, remarkably decrease the proportion of time taken by the end revision phase when the deadline was removed. This tells one of the critical behavioural features about how the two groups of participants diverged in response to the free-of-time-pressure situation during the translation process.

Overall, both groups tended to maintain a relatively stable proportion of time on the orientation phase regardless of the time conditions. However, the proportions of time spent on the drafting and end revision phases varied more evidently across time conditions and showed divergent tendencies between the two groups.

6.2.2.4 Types of revision behaviour

Online and end revision behaviours were examined with the indicator of deletion count during the drafting and the end revision phase respectively. It showed that the effect of Condition was statistically significant in affecting both the number of online revision activities and end revision activities (Table 6-17). Generally, the more relaxed the timeframe, the more online and end revision activities were carried out (Figure 6-15). There was also an interaction effect between Condition and Group on the number of end revision deletions. This indicates that the two groups of participants tended to display different patterns for end revision deletions modulated by time condition. As shown in Figure 6-15, experienced translators managed to continuously increase their end revision activities with the alleviation of time pressure, while novice translators maintained the number of end revision activities in the Free condition at a level similar to the Standard condition.

Table 6-17. Type III ANOVA table of the LMER models of *Online deletions* and *End revision deletions*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
Online deletions	Condition	5275.60	2637.80	2	126	5.52	0.005	**
	Group	51.60	51.60	1	63	0.11	0.744	
	Condition:Group	1144.80	572.42	2	126	1.20	0.305	
End revision deletions	Condition	5621.60	2810.82	2	126	13.71	<.001	***
	Group	226.40	226.43	1	63	1.10	0.297	
	Condition:Group	1616.00	808.02	2	126	3.94	0.022	*

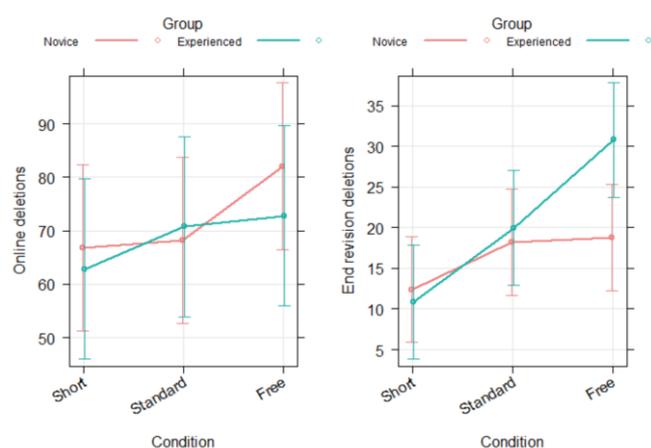


Figure 6-15. Effects of Condition and Group on *Online deletions* and *End revision deletions*

The follow-up pairwise comparisons showed that the increase of online deletions

for the novice group from the Short to the Free condition ($t(126)=-2.93, p=.012$) and from the Standard to the Free condition ($t(126)=-2.66, p=.027$) were statistically significant, while the contrasts among the three time conditions for experienced translators were not prominent. For the end revision deletions, on the other hand, no statistical significance was observed for the three contrasts for the novice group while all the three contrasts were statistically significant for the group of experienced translators ($t(126)=-2.46, p=.046$ for the Short-Standard comparison; $t(126)=-5.39, p<.001$ for the Short-Free comparison; and $t(126)=-2.93, p=.012$ for the Standard-Free comparison).

These results indicated that novices increased their online revision activities remarkably with the removal of the deadline. However, the removal of the deadline did not stimulate them to increase their end revision activities in a similar way. On the contrary, the Free condition provoked end revision activities for experienced translators rather than the online revision activities. In other words, when the deadline was absent, novices opted for more online revisions while experienced translators opted for more end revisions. With regard to these differences in the revision behaviours of the two groups, H9 (*Time condition would affect the pattern of preference for online/end revision differently for novice and experienced translators.*) was corroborated, especially for the contrast between the Standard and Free conditions for the two groups.

6.2.2.5 Summary of RQ3

Overall, it was found that a more stringent time condition could shorten and reduce the number of all types of pauses for both groups of participants. Thus, H6 was corroborated. The stringency of time condition influenced the segment length for novice translators more significantly than for experienced translators. Novice translators' segmentation was evidently prolonged, especially in the most stringent time condition. In contrast, time condition did not affect the segment length for experienced translators, which indicated a relatively stable rhythm in terms of the segmentation of this group. Therefore, H7 was corroborated. A stringent time condition also significantly shortened the drafting and the end revision phase (with an exception for novice translators who had a shorter end revision phase in the Free condition than in the Standard condition). The length of an orientation phase was not affected by time condition. Thus, H8 was not corroborated.

As novice translators preferred to have a longer drafting phase and a shorter end revision phase in the Free condition compared with the Standard condition, they carried out more online revisions and fewer end revisions when the deadline was removed. Experienced translators, on the contrary, preferred to extend their end revision phase when the deadline was removed so that their end revision activities boosted in the Free condition. Although their drafting phase was prolonged in the Free condition as well, the online revision activities did not increase evidently in this condition. Hence, when the deadline was absent, novices opted for more online revisions while experienced translators opted for more end revisions. As such, the pattern of preference for end/online revision changed across time conditions differently for novice and experienced translators. Thus, H9 was corroborated.

6.3 The performance effects

This section reports how time condition and experience affect translation performance with regards to the translation product quantity (6.3.1) and product quality (6.3.2). Section 6.3.2, which focuses on the translation product quality, addresses RQ4 and the corresponding hypotheses as presented below:

RQ4: What is the relationship between time condition and translation quality for the two groups of participants?

H10. The stringency of time condition would negatively affect translation quality for both groups.

H11. Experienced translators can achieve better translation quality than novice translators regardless of the time conditions.

6.3.1 Product quantity

There were 8 out of 35 (22.86%) and 2 out of 35 (5.71%) novice translators who did not manage to complete the task under the Short and the Standard condition respectively. All the experienced translators managed to complete the three tasks.

Product quantity was examined from two aspects: the number of Chinese characters produced in the final translation product (TT word count) and the total number of TUs produced during the whole translation process. The results of how

Condition and Group modulated the two metrics are presented in Table 6-18 and Figure 6-16 below.

Table 6-18. Type III ANOVA table of the LMER models of *TT word count* and *TU count*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
TT word count	Condition	20132	10066	2	126	13.10	<.001	***
	Group	354.5	354.5	1	63	0.46	0.500	
	Condition:Group	2849.8	1424.9	2	126	1.85	0.161	
TU count	Condition	25310.7	12655.3	2	126	21.81	<.001	***
	Group	208.5	208.5	1	63	0.36	0.551	
	Condition:Group	854.7	427.4	2	126	0.74	0.481	

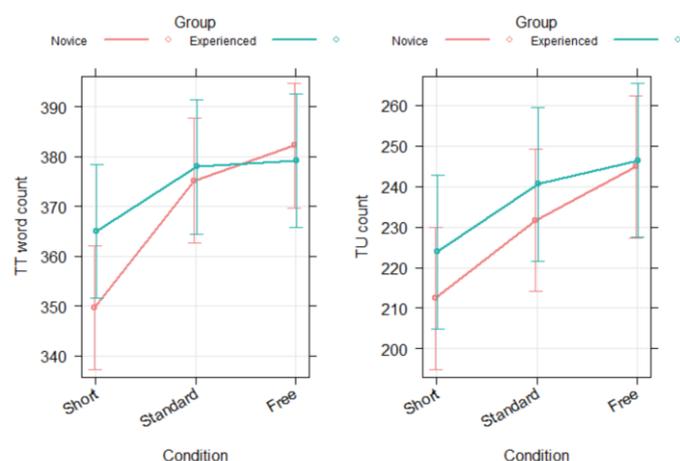


Figure 6-16. Effects of Condition and Group on *TT word count* and *TU count*

It is not surprising that the effect of Condition was statistically significant for both metrics, indicating that the stringent time condition restrained the quantity of translation products. The follow-up pairwise test showed that novice translators had a statistically significant increase in TT word count from the Short to the Free condition ($t(126)=-4.92$, $p<.001$); but there were no statistically significant contrasts among the comparisons of the three conditions for experienced translators in TT word count. A general tendency displayed for both groups in TU count was that more TUs were produced during the translation process with the alleviation of time stringency. The follow-up pairwise test for the novice group showed that the Short-Standard ($t(126)=-3.35$, $p=.003$) and Short-Free ($t(126)=-5.65$, $p<.001$) contrasts were statistically significant, and the Standard-Free contrast ($t(126)=-2.29$, $p=.071$) was marginally

significant. The follow-up pairwise test for experienced translators showed that only the Short-Standard ($t(126)=-2.69$, $p=.024$) and Short-Free ($t(126)=-3.63$, $p=.001$) contrasts were statistically significant.

Thus, it can be considered that there was a prominent spur for novice translators in their online and final product quantity from the Short to the other two conditions. While experienced translators' online product quantity (TU count) increased obviously from the Short to the other two conditions, they tended to have a stabilised final product quantity (TT word count) in the three conditions.

6.3.2 Product quality

The results for translation quality scores (*Quality*) showed that both Condition and Group were statistically significant in affecting the translation quality (Table 6-19). However, there was no interaction effect between these two predictors. Figure 6-17 shows that overall, experienced translators achieved higher scores than novices regardless of the time conditions. Also, experienced translators managed to improve their translation quality when time pressure was gradually alleviated, while novice translators' translation quality was improved substantially when the time condition was changed from the most stringent to the two less stringent conditions.

Table 6-19. Type III ANOVA table of the LMER model of *Quality*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
Quality	Condition	2508	1254	2	126	13.22	<.001	***
	Group	1817.5	1817.5	1	63	19.17	<.001	***
	Condition:Group	338.04	169.02	2	126	1.78	0.173	

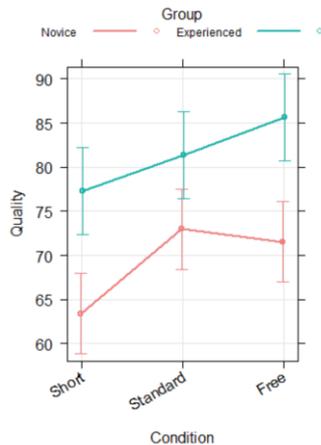


Figure 6-17. Effects of Condition and Group on *Quality*

The follow-up pairwise tests showed that for novice translators, the improvements in translation quality score from the Short condition to the other two conditions ($t(126)=-4.11$, $p<.001$ for the Short-Standard contrast and $t(126)=-3.50$, $p=.002$ for the Short-Free contrast) were statistically significant. For experienced translators, only the increase from the Short to the Free condition was statistically significant ($t(127)=-3.32$, $p=.004$). Overall, H10 (*The stringency of time condition would negatively affect translation quality for both groups.*) was partially corroborated by these tendencies, with a violation for novice translators in the Standard and Free condition. In addition, all the contrasts between the two groups in each time condition were statistically significant, which meant that experienced translators achieved significantly higher scores than novice translators in each time condition. Thus, H11 (*Experienced translators can achieve better translation quality than novice translators regardless of the time conditions.*) was corroborated.

6.3.3 Summary of RQ4

The results of translation quality modulated by time condition and experience corroborated H11 and partially corroborated H10. Firstly, the quality scores of experienced translators showed a clear tendency that conformed with a negative relation with time stringency. A feature shown by the novice translator group was that while they exhibited an evident improvement in translation quality from the Short to the Standard condition, they failed to make further improvement in the Free condition. Although the interaction effect between Condition and Group on translation quality

score was not statistically significant, this difference in tendencies that emerged for the two groups was compelling. Thus, H10 was especially corroborated by the group of experienced translators and was corroborated by novice translators when comparing the most time-constrained (Short) and the less time-constrained (Standard and Free) conditions. Secondly, experienced translators achieved significantly higher translation quality in each time condition than novice translators. Thus, H11 was fully corroborated.

6.4 The role of cognitive load

This section sets out to address RQ5 and the corresponding hypotheses as listed below. It looks into the indicators of extraneous load (section 6.4.1) firstly and then determines the indicators of the overall cognitive load based on CLT (section 6.4.2). The effects of time condition and experience on germane load are examined in section 6.4.3. Based on these results, other behavioural metrics are explored in relation to overall cognitive load, germane load and translation quality (section 6.4.4).

RQ5: Within the framework of Cognitive Load Theory, how does time condition affect participants' overall cognitive load, germane load and the way in which germane load influences translation performance?

H12. The stringency of time condition would proportionately exacerbate the overall cognitive load.

H13. The effect of time condition on germane load would be similar to its effect on translation quality for the two groups of participants.

H14: The stringency of time condition could modulate the way germane load influences translation quality.

6.4.1 Intrinsic and extraneous load

As mentioned in section 5.2, the difficulty levels of the three source materials are comparable. This was confirmed by the results obtained from the dimension of Mental Demand in the NASA TLX questionnaire (Table 6-20). Both Condition and Group were not statistically significant in affecting participants' perception of Mental Demand, which indicated a comparable complexity of the tasks. Thus, it can be assumed that the

intrinsic load of the three tasks was consistent for the two groups of participants.

Table 6-20. Type III ANOVA table of the LMER model of *Mental Demand*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
Mental Demand	Condition	137.08	68.54	2	126	0.49	0.612	
	Group	6.18	6.18	1	63	0.04	0.834	
	Condition:Group	2.20	1.10	2	126	0.01	0.992	

In the present study, the variation of extraneous load mainly originated from the variation of time pressure induced by the three manipulated time conditions because all the other experimental conditions in each task were constant. Therefore, the psycho-physiological measures which corroborated H1 (see section 6.1) could be taken as time pressure indicators and thus may represent the extraneous load. Another essential property of extraneous load is its destructive effect on, or negative relation with performance (translation quality), based on Cognitive Load Theory. With these two preconditions, i.e., successful time pressure indicators and of negative correlation with translation quality, the indicators of extraneous load can be determined.

The repeated measures correlations between translation quality and each of the psycho-physiological measures which indicated time pressure (i.e., State Anxiety, Stress, Temporal Demand, HR, BP, and pupil size) were examined. The results showed that although all the time pressure indicators showed negative correlations with translation quality, the two measures of Stress ($r=-0.27$, $p=.002$) and Temporal Demand ($r=-0.33$, $p<.001$) demonstrated the most prominent negative correlations with translation quality compared with other measures. Thus, self-reported Stress and Temporal Demand were considered as indicators of extraneous load in the following analysis.

6.4.2 Overall cognitive load

As mentioned in Chapter 3, possible indicators of overall cognitive load could be filtered from the indicators of *effort* (which refers to the problem-based coping and active regulatory activity of changing the cognitive state). These indicators include self-reported Effort in the NASA TLX questionnaire, the behaviour-based measures of PWR and MFD, and the physiological measures of GSR and HRV. The results of how Condition and Group affected the self-reported and behaviour-based measures are

presented in Table 6-21.

Table 6-21. Type III ANOVA table of the LMER models of *Effort*, *PWR* and *MFD*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
Effort	Condition	695.28	347.64	2	126	3.90	0.023	*
	Group	110.17	110.17	1	63	1.24	0.270	
	Condition:Group	533.74	266.87	2	126	2.99	0.054	
PWR	Condition	2.60E-02	1.30E-02	2	126	5.60	0.005	**
	Group	2.47E-03	2.47E-03	1	63	1.07	0.306	
	Condition:Group	8.88E-03	4.44E-03	2	126	1.91	0.152	
MFD	Condition	5.43E-05	2.72E-05	2	108	0.16	0.852	
	Group	2.89E-06	2.89E-06	1	54	0.02	0.897	
	Condition:Group	4.93E-04	2.46E-04	2	108	1.46	0.237	

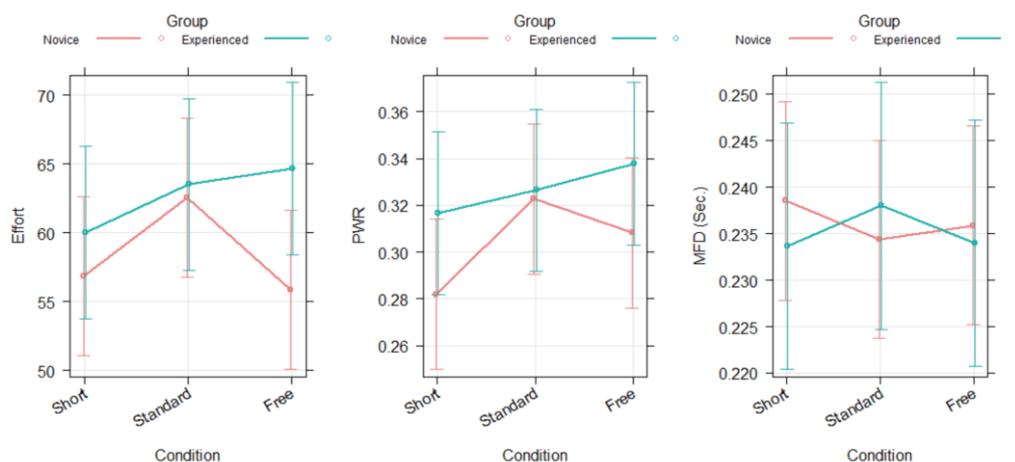


Figure 6-18. Effects of Condition and Group on *Effort*, *PWR* and *MFD*

The effects of Condition and Group on MFD were not prominent, as shown by the results. However, Condition significantly affected the measures of Effort and PWR. In the follow-up pairwise tests, the Short-Standard ($t(126)=-2.53, p=.038$) and Standard-Free ($t(126)=2.98, p=.011$) contrasts in Effort were statistically significant for novice translators, while no contrast was statistically significant for experienced translators. Similarly, the Short-Standard contrast in PWR was prominent for novice translators ($t(126)=-3.54, p=.002$), whereas no contrast was prominent for experienced translators. This meant that the effect of Condition was more remarkable on novices than on experienced translators in self-reported Effort and PWR.

The variation tendencies of Effort and PWR modulated by Condition and Group

were similar (Figure 6-18), and such a pattern of variation across the three time conditions displayed by the two groups conformed with the pattern exhibited by the indices of GSR and HRV (see section 6.1). In this consistent pattern displayed by these measures (i.e., Effort, PWR, GSR and HRV), novice translators tended to have a remarkably higher effort level in the Standard condition than in the other two conditions, while experienced translators were affected little by time condition in these measures.

In order to identify concrete measures of overall cognitive load, the repeated measures correlation tests among these measures of effort were carried out (Table 6-22). It showed that GSR Mean was significantly (but weakly) correlated with PWR and RMSSD, respectively¹⁷. Consequently, these three measures/indices were taken as possible indicators of overall cognitive load and were further testified in the following analysis based on the assumptions of CLT.

Table 6-22. Repeated measures correlation coefficients among the effort indicators

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Effort	1								
2. PWR	0.11	1							
3. MFD	-0.01	0.16.	1						
4. GSR Mean	0.05	0.22*	-0.09	1					
5. SCR Amplitude	0.12	0.13	-0.15	0.76***	1				
6. SCR Frequency	0.01	0.17.	-0.01	0.45***	0.60***	1			
7. RMSSD	-0.11	-0.14	-0.11	-0.21*	-0.18.	-0.06	1		
8. SDNN	-0.11	-0.11	-0.11	-0.19.	-0.14	-0.05	0.93***	1	
9. PNN50	-0.17.	-0.17.	-0.07	-0.19.	-0.15	-0.03	0.89***	0.78***	1

The sum of the intrinsic and germane load can be determined by subtracting the z score of extraneous load (represented by one of the extraneous load indicators, i.e., Stress and Temporal Demand) from the z score of overall cognitive load (represented by one of the candidates identified above, i.e., PWR, GSR Mean and RMSSD), as shown in the following formula:

$$Z_{\text{intrinsic+germane load}} = Z_{\text{overall cognitive load}} - Z_{\text{extraneous load}}$$

If $Z_{\text{intrinsic+germane load}}$ is positively correlated with translation quality, it can be considered that the germane load derived from this approach positively correlates with translation quality, as the intrinsic load is consistent for the three tasks. In this way, the overall cognitive load indicators can be further determined by assessing the correlation

¹⁷ Indices of GSR or HRV (e.g., RMSSD and SDNN) are inherently correlated. Thus, those strong and significant correlations between GSR indices and between HRV indices are not taken into consideration.

between $Z_{\text{intrinsic+germane load}}$ and translation quality. The results of the repeated measures correlation tests (with different combinations of extraneous load and effort indicators) are presented in Table 6-23.

Table 6-23. Repeated measures correlation coefficients between $Z_{\text{intrinsic+germane load}}$ and Quality with $Z_{\text{intrinsic+germane load}}$ being calculated by different combinations of overall cognitive load and extraneous load indicators

Overall cognitive load	Extraneous load	Correlation coefficients between $Z_{\text{intrinsic+germane load}}$ and Quality
PWR	Temporal Demand	0.27**
	Stress	0.23*
GSR Mean	Temporal Demand	0.22*
	Stress	0.14
RMSSD	Temporal Demand	0.24**
	Stress	0.17*

With the overall cognitive load represented by the three candidates, i.e., PWR, GSR Mean and RMSSD, $Z_{\text{intrinsic+germane load}}$ was positively and significantly correlated with translation quality for at least one of the combinations when either Temporal Demand or Stress represented extraneous load. This meant that the three tested indicators of overall cognitive load were valid based on the assumption of CLT (i.e., positive relation between germane load and performance). Particularly, when extraneous load was represented by Temporal Demand, its combinations with all the three overall cognitive load indicators demonstrated significant correlations between $Z_{\text{intrinsic+germane load}}$ and translation quality. Thus, to present the results more efficiently, Temporal Demand was taken as the representative measure for extraneous load in the following analysis and discussion.

Valid measures of overall cognitive load (PWR, GSR Mean and RMSSD) are thus determined now. As shown by the results obtained from the measures of PWR, GSR Mean and RMSSD regarding the effects of Condition and Group, H12 (*The stringency of time condition would proportionately exacerbate the overall cognitive load.*) was not corroborated because the effect of Condition on these measures did not show a linear pattern for novice translators and was not statistically significant for experienced translators.

6.4.3 Germane load

The effects of Condition and Group on germane load can thus be examined based on

their influence on $Z_{\text{intrinsic+germane load}}$ (IGL). Since Temporal Demand (TD) has been determined as a representative measure for extraneous load, there are three combinations of overall cognitive load and extraneous load to calculate or denote $Z_{\text{intrinsic+germane load}}$, i.e., $Z_{\text{PWR}} - Z_{\text{TD}}$ (IGL_PWR TD), $Z_{\text{GSR Mean}} - Z_{\text{TD}}$ (IGL_GSR TD) and $Z_{\text{RMSSD}} - Z_{\text{TD}}$ (IGL_RMSSD TD). Table 6-24 below presents the results. It shows that there was an interaction effect between Condition and Group on $Z_{\text{intrinsic+germane load}}$ when the overall cognitive load was represented by PWR and GSR Mean, indicating that the variation tendency of these measures across time conditions can be different for the two groups.

Table 6-24. Type III ANOVA table of the LMER models of *IGL_PWR TD*, *IGL_GSR TD* and *IGL_RMSSD TD*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	p value	Sig.
IGL_PWR TD	Condition	94.624	47.31	2	126	91.24	<.001	***
	Group	1.518	1.52	1	63	2.93	0.092	
	Condition:Group	3.303	1.65	2	126	3.19	0.045	*
IGL_GSR TD	Condition	69.039	34.52	2	126	78.83	<.001	***
	Group	0.43	0.43	1	63	0.98	0.326	
	Condition:Group	3.103	1.55	2	126	3.54	0.032	*
IGL_RMSSD TD	Condition	59.212	29.61	2	126	43.35	<.001	***
	Group	0.081	0.08	1	63	0.12	0.732	
	Condition:Group	2.289	1.14	2	126	1.68	0.191	

Figure 6-19 below visualises the variations of $Z_{\text{intrinsic+germane load}}$ calculated with different indicators of overall cognitive load. Generally, the variations of germane load showed a consistent pattern in these measures: experienced translators managed to increase their germane load with the time stringency being alleviated; novice translators could increase their germane load from the Short condition to the Standard condition, but not from the Standard to the Free condition. The gradual increments of $Z_{\text{intrinsic+germane load}}$ from the Short to the Standard condition and from the Standard to the Free condition for experienced translators were statistically significant in all the three measures. However, for novice translators, only the Short-Standard contrast was prominent in the three measures, and the difference between the Standard and Free conditions was minor.

Such a pattern exhibited by the two groups of participants was consistent with the pattern shown by translation quality regarding the effect of Condition on the two groups of participants (in section 6.3.2). Thus, H13 (*The effect of time condition on germane*

load would be similar to its effect on translation quality for the two groups of participants.) was corroborated.

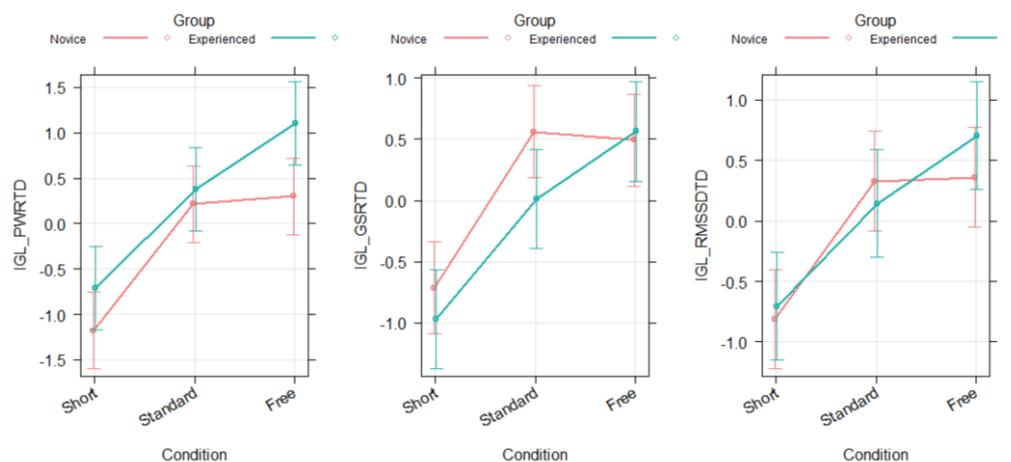


Figure 6-19. Effects of Condition and Group on *IGL_PWRTD*, *IGL_GSRTD* and *IGL_RMSSDTD*

To examine whether time condition modulated the way in which germane load influenced translation quality (H14), a three-way analysis was carried out to examine the effects of three predictors, Condition, Group and $Z_{\text{intrinsic+germane load}}$, on the dependent variable, which is translation quality. Three possible $Z_{\text{intrinsic+germane load}}$, i.e., *IGL_PWRTD*, *IGL_GSRTD* and *IGL_RMSSDTD*, were tested in the models (Table 6-25). If time condition modulates the effect of germane load on translation quality, an interaction effect between Condition and $Z_{\text{intrinsic+germane load}}$, or among Condition, Group and $Z_{\text{intrinsic+germane load}}$ should be significant in the models.

It was found that the measure of *IGL_PWRTD* demonstrated a significant interaction effect with Condition. This meant that time condition modulated the way in which germane load (represented by *IGL_PWRTD*) affected translation quality. Figure 6-20 illustrates the effects.

Table 6-25. Type III ANOVA table of the LMER models of *Quality* with fixed effects of Condition, Group and $Z_{\text{Intrinsic+Germane}}$ represented by *IGL_PWRTD*, *IGL_GSRTD* and *IGL_RMSSDTD*

Measure	Fixed effects	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	<i>p</i> value	Sig.
Quality	Condition	868.86	434.43	2	135.14	5.05	0.008	**
	Group	1178.92	1178.92	1	62.75	13.70	<.001	***
	IGL_PWRTD	5.12	5.12	1	164.11	0.06	0.807	
	Condition:Group	170.76	85.38	2	135.14	0.99	0.373	
	Condition:IGL_PWRTD	960.96	480.48	2	120.75	5.59	0.005	**
	Group:IGL_PWRTD	40.33	40.33	1	164.11	0.47	0.495	
	Condition:Group:IGL_PWRTD	342.14	171.07	2	120.75	1.99	0.141	
Quality	Condition	1285.16	642.58	2	145.56	7.03	0.001	**
	Group	1829.47	1829.47	1	72.84	20.03	<.001	***
	IGL_GSRTD	53.41	53.41	1	168.27	0.58	0.446	
	Condition:Group	790.61	395.3	2	145.56	4.33	0.015	*
	Condition:IGL_GSRTD	122.1	61.05	2	133.69	0.67	0.514	
	Group:IGL_GSRTD	178.8	178.8	1	168.27	1.96	0.164	
	Condition:Group:IGL_GSRTD	185.64	92.82	2	133.69	1.02	0.365	
Quality	Condition	1249.6	624.8	2	133.23	6.75	0.002	**
	Group	1380.69	1380.69	1	67.61	14.91	<.001	***
	IGL_RMSSDTD	0.71	0.71	1	181.78	0.01	0.930	
	Condition:Group	253.27	126.64	2	133.23	1.37	0.258	
	Condition:IGL_RMSSDTD	478.45	239.23	2	131.99	2.58	0.079	
	Group:IGL_RMSSDTD	91.79	91.79	1	181.78	0.99	0.321	
	Condition:Group:IGL_RMSSDTD	216.6	108.3	2	131.99	1.17	0.314	

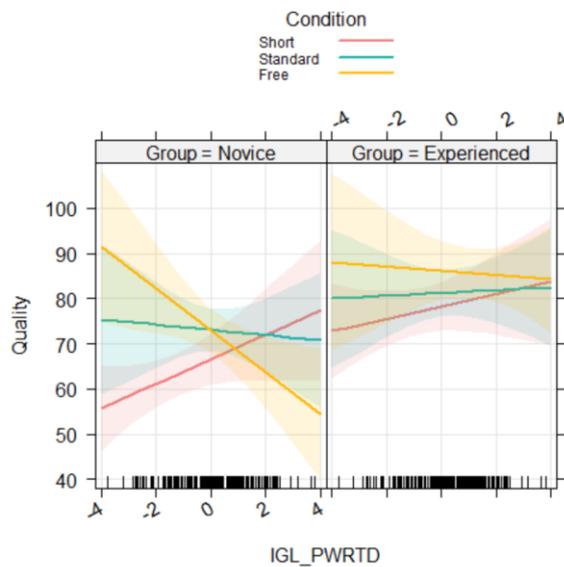


Figure 6-20. Effects of Condition, Group and *IGL_PWRTD* on *Quality*

Generally, the effect of germane load on translation quality for experienced translators was not modulated much by time condition. However, germane load showed different effects on translation quality in different time conditions for novice translators:

in the Short condition, germane load positively fostered performance; in the Standard condition, its function was not apparent; but in the Free condition, it acted destructively on performance. This result provided evidence that the free-of-deadline condition may undermine the facilitative effect of germane load on the performance of novice translators. Thus, H14 (*The stringency of time condition could modulate the way germane load influences translation quality.*) was corroborated with the measure of IGL_PWRTD and for the group of novice translators.

6.4.4 Behavioural metrics, cognitive load, and translation quality

To further explore whether translation behaviours had specific correlations with overall cognitive load, germane load, or translation quality, five behavioural metrics (total LP count, LP count in the end revision phase, SG length measured by TU number, end revision duration, and deletions in the end revision phase) from the previous analyses were examined. These metrics shared an effect of time condition (showing either the same or opposite tendency) similar to the pattern displayed by either the overall cognitive load or germane load.

Table 6-26 below presents the repeated measures correlation coefficients of these behavioural metrics with the overall cognitive load indicators, the germane load measures ($Z_{\text{intrinsic+germane load}}$), and translation quality. The correlations between PWR related measures and total LP count/SG length (measured by TU count) were considered not applicable since they were inherently correlated based on their calculation formulas ($\text{PWR} = \text{LP count} \div \text{TU count}$; $\text{SG length} = \text{TU count} \div \text{SG count}$; $\text{LP count} \approx \text{SG count}$). The results showed that the metrics of total LP count and LP count in the end revision phase were weakly correlated with GSR Mean. More importantly, all the five behavioural metrics under examination demonstrated moderate or mild correlations with the germane load measures. This indicated that these behavioural metrics were possibly associated with germane load in translation. That is, more LPs, shorter SGs, a longer end revision phase, and more end revision behaviours might signal a higher germane load in translation. In addition, all the behavioural metrics were statistically significantly but mildly correlated with translation quality except for the number of deletions in the end revision phase. Hence, these metrics might be of particular interest in gauging the germane load levels in translation and thus might be indicative of translation quality based on these results.

Table 6-26. Repeated measures correlation coefficients between the behavioural metrics and the overall cognitive load measures, $Z_{\text{intrinsic+germane load}}$ indicators and *Quality*

Measure		Total LP count	LP count (end revision)	SG length (TU count)	End revision duration	End revision deletions
Overall cognitive load	PWR	N/A	N/A	N/A	0.07	0.09
	GSR Mean	0.17*	0.27**	0.02	0.13	0.07
	RMSSD	-0.05	0.00	0.10	0.06	0.07
$Z_{\text{intrinsic+Germane}}$	IGL_PWRTD	N/A	N/A	N/A	0.40***	0.34***
	IGL_GSRTD	0.43***	0.47***	-0.27**	0.47***	0.37***
	IGL_RMSSDTD	0.35***	0.31***	-0.32***	0.35***	0.28**
Quality		0.19*	0.21*	-0.21*	0.30***	0.14

6.4.6 Summary of RQ5

This section determines the measures of overall cognitive load and germane load based on the framework of CLT. The effects of time condition and experience on participants' overall cognitive load and germane load are thus examined. Firstly, it was found, based on the verified measures (PWR, GSR Mean and RMSSD) of overall cognitive load, that the stringency of time condition did not constantly exacerbate the overall cognitive load, and thus H12 was not corroborated. This might be caused by the trade-off effect between extraneous load and germane load in different time conditions. Secondly, the stringency of time condition did influence germane load in the same way as it influenced translation quality for both groups, which corroborated H13. Thirdly, time condition modulated the effect of germane load (represented by IGL_PWRTD) on translation quality for novice translators in that the facilitative function of germane load was crippled in the Free condition. This result corroborated H14, but only partially, since such an effect of time condition was not visible on experienced translators.

Generally, this section identified relatively reliable indicators of extraneous load and overall cognitive load, based on which the indicators of germane load ($Z_{\text{intrinsic+germane load}}$) were generated. In addition, it was found that several behavioural metrics, i.e., LP count, LP count in the end revision phase, SG length measured by TU number, end revision duration and deletions in the end revision phase, were prominently related to the level of germane load exerted during translation and might be indicative of translation quality.

Chapter 7. Discussion

This chapter provides a thorough discussion of all the results obtained from the translation experiments under the three time conditions completed by the two groups of participants. The effects of time pressure and translators' professional experience on the translation process, the translation product and the psycho-physiological responses are examined wherever possible to reveal the mechanisms underpinning the observable phenomena. The implications of the findings from the psycho-physiological (section 7.1), behavioural-cognitive (section 7.2 – 7.5) and performance (section 7.6) aspects are discussed in the context of the aims and research questions proposed in the investigation.

7.1 Psycho-physiological changes associated with time pressure

When the participants were translating with more stringent time constraints, they reported significantly higher perceived time pressure (Temporal Demand), Stress and State Anxiety. Both groups of participants exhibited the same trend in terms of these psychological measures, with little group difference captured in each time condition. It was also found that Trait Anxiety did not modulate the role time condition played in affecting State Anxiety. In general terms, this implies that participants with both lower and higher Trait Anxiety tended to experience significantly elevated State Anxiety when they were confronted with a stringent time constraint during a translation task.

These results provide empirical evidence that tight deadlines (along with the subjective time pressure manipulation strategies applied in the study) can lead to a series of changes in translators' psychological experience, no matter whether they are novice or experienced translators. The strong correlations among the three psychological measures (shown in section 6.1.3) further prove that the (perceived) shortage of time in handling the task demands could elicit subjective feelings of time pressure, stress and anxiety.

Furthermore, the results suggest that there was little difference between novice and experienced translators in regulating the relatively negative emotions (as reflected in the self-reported feelings of time pressure, stress and anxiety) under the most stringent time condition. However, experienced translators reported slightly lower perceived time pressure (Temporal Demand) and Stress than novice translators in the less time-pressured conditions. As such, it is possible that the supposed superior capability of emotion regulation and stress management of experienced translators (e.g., Hubscher-

Davidson, 2016) diminished under the extremely stressed conditions.

Meanwhile, these results appear to support the argument that an intrinsic part of the emotional experience is the subjective feeling associated with it (Minkel & Phillips, 2015). The self-reported question on Stress, Temporal Demand and the State-Trait Anxiety Instrument have been proven to be valid psychometric instruments for time pressure measurement in translation. However, despite their adequacy together in signalling time pressure in the present study, Slavich et al. (2019, p. 409) suggested the inherent limitation of the subjective measures: they can easily be affected by a “cognitive bias and social desirability”. Thus, it is understandable that participants tended to focus their consciousness on perceptible changes in either their stress level or anxiety in response to the overtly manipulated time constraints.

As indicated by most biomarkers under study, it is not surprising that participants were more nervous at the beginning of each experiment (Baseline 1), but they recovered to reach a much more relaxed state at the end of each experiment (Baseline 2) after the completion of all the tasks. Some of the biomarkers suggested that participants’ arousal level in Baseline 1 was approximate to or higher than that in the most stringently time-constrained task, although they were told to be relaxed during the baseline session. Thus, the Baseline 1 session may not reflect a real “baseline” state with regard to the participants’ physiological responses.

The variations in the biomarkers across different time conditions indicate that participants tended to have significantly elevated blood pressure, faster heartbeat and dilated pupils when time pressure was getting more intense. The findings on HR and BP are partly in line with Korpál’s (2016) study in which HR reflected stress levels in interpreting activities but systolic and diastolic blood pressure did not. Korpál (2016) indicated that this may be caused by the fact that hypertension has been recognised as a marker of chronic stress rather than the momentary experience of stress. The present study demonstrated a statistically significant effect of time condition on HR, SBP, DBP and MAP, all of which were moderately or mildly correlated with Temporal Demand. SBP was also mildly correlated with Stress and State Anxiety. These findings suggest that both HR and BP reflected the situational time pressure or stress level to a considerable degree in written translation.

The group difference in HR and BP was not prominent in the task sessions. However, an obvious contrast in HR between the two groups was observed in the baseline sessions in which the HR values of experienced translators were statistically

significantly lower than that of the novice translators. This difference could be caused by their age difference: the novice translators were around six years younger than the experienced translators on average. The reduced group difference in the task sessions indicated a tendency of convergence of the two groups in this biomarker. That is, the translation task demands, regulated by different levels of time pressure, made their physiological response in HR converge to a comparable level, varying in a similar trend across time conditions, with their original group difference (i.e., shown in the baseline values) receding.

Pupil size was also found to be significantly affected by time condition and moderately or weakly correlated with HR, SBP and Temporal Demand. Since pupil dilation has been frequently mentioned and used as an indicator of effort/cognitive load during a cognitively demanding task (Hvelplund, 2011; Paas, Tuovinen, et al., 2003; Seeber, 2013), this visible conformity with the pattern signifying time pressure and stress may entail some caution for future studies to interpret the underlying origin of dilated pupils. To be specific, stress could be a confounder when pupil dilation is taken as an indicator of cognitive load because the stress-related factors (such as time pressure in this case) may prevail over the load representing cognitive processing in this measure. Skin temperature failed to capture any significant effect of time condition or experience in this study. A possible reason could be that this measure is easily influenced by the physical environment in direct contact with the skin surface, which makes it not sufficiently sensitive to detect the variation of arousal innervated by the change of time pressure.

The main effect of time condition on the indices of GSR and HRV was not prominent, while a group difference was captured in SCR Amplitude, with novices having significantly higher values than experienced translators. Moreover, an interaction effect of time condition and experience was shown in GSR Mean and SCR Frequency. Novice translators exhibited an inverted “U-shaped” relation between time stringency and the arousal state represented by GSR and HRV, while experienced translators displayed a tendency without a wide variation across time conditions in these measures. Interestingly, although only two indices of GSR pronounced such a significant interaction effect, all the other GSR and HRV indices conformed with this pattern, which was also visible in other indicators of *effort* such as PWR and self-reported Effort. It was also found that the indices of GSR and HRV were mostly inter-correlated and some of them (e.g., GSR Mean, SCR Frequency and PNN50) were

weakly correlated with PWR and self-reported Effort.

This evidence points to the connection between such physiological measures and *effort*. That is, GSR and HRV might mirror more of the intensity of cognitive processing under different time conditions, and they were thus taken as physiological indicators of *effort/overall* cognitive load in the present study. In this regard, these results confirm that, apart from the stress-related origin, changes in cognitive functioning are also reflected by the physiological arousal, as has been verified in a variety of behavioural studies (e.g., Luque-Casado et al., 2016; Nourbakhsh et al., 2017).

To sum up, the biomarkers of HR, BP and pupil size, together with the psychological measures of Temporal Demand, Stress and State Anxiety, were verified as time pressure measures. These measures corroborated H1 in that they varied in accordance with the stringency of time condition showing a consistent tendency for the two groups of participants. These results were also in line with the preliminary analysis (Weng et al., in press) of the psycho-physiological responses of all the student participants. This means that when the sample size was enlarged (from 45 to 65 participants) and when participants' experience was taken into account in the present analysis, there was no obvious difference. Most of these measures also correlated with each other which further consolidated their convergence in reflecting the time pressure levels.

In addition, GSR and HRV featured a different pattern which conformed to that shown by the behavioural and self-reported indicators of *effort/overall* cognitive load. Thus, GSR and HRV were identified as physiological indicators of *effort/overall* cognitive load. The main effect of Group was not prominent in most of these psycho-physiological responses, despite the fact that the pattern of GSR and HRV modulated by time condition exhibited by novices was different from that exhibited by experienced translators. Also, novices did not have stronger psycho-physiological responses than experienced translators when the time condition was becoming increasingly stringent. Based on these findings, the first aim of this study, which is about determining the measures of time pressure in translation and revealing group difference in the psycho-physiological effects of time pressure on translators has been achieved.

7.2 Translation-related behavioural effects

7.2.1 Automaticity at the micro- and macro-level

Automatic processes are fast, effortless and could occur in parallel with other processes (Logan, 1997; Segalowitz & Hulstijn, 2005). The present study examined these aspects at the micro-level of the translation process using a range of elaborate behavioural metrics. Most metrics from the aspects of processing speed, parallel processing degree and average TU effort supported the hypothesis (H4) that time pressure can enhance automaticity during the translation process. With time condition getting more stringent, participants generally accelerated the processing with more keystrokes produced per minute, and shorter intervals between and within TUs (IbTUs and IwTUs); the participants, especially experienced translators, also increased their parallel processing degree with a higher proportion of attention devoted to ST comprehension during the course of producing a TU; the effort measured by time devoted to each individual TU was reduced in the time-constrained tasks.

However, not all the behavioural metrics under study displayed a fully uniform tendency as predicted in H4. For example, while the production speed of novice translators reached a relatively high level in the Short condition, they had a comparable speed in keystroke production in the Standard and Free conditions. This indicates that only the most time-pressured condition significantly accelerated the production as a whole; the moderate time constraint did not function as expected to speed up the production compared with the condition without a deadline in terms of keystrokes per minute during the whole translation process. It is possible that novice translators had a high processing speed at the beginning part of the task in which they worked towards the deadline of the Standard condition; but as the deadline was not tight, they had a sufficiently long end revision phase (which is confirmed by the results of the end revision duration) before the deadline; this renders a slow production speed in the later part of the whole process. A trade-off effect may thus be elicited leading to a considerably slowed-down production speed as a whole in this condition. Such an effect, however, did not appear in experienced translators as they had a gradually reduced production speed with the time stringency getting more relaxed.

Another slight discrepancy to the tendency predicted in H4 appears at the Short and Standard condition for experienced translators in the metric of IwTU: the value of this metric in the Short session was higher than that in the Standard session. Although this discrepancy was not prominent, it is worth noticing. A possible reason for this could be that under the most stringent time condition, more intensive mental processing (e.g., problem solving activities) may occur when the experienced translators were typing the

TUs, resulting from more intensive parallel processing. This could essentially slow down the typing speed inside the TUs, leading to slightly longer within-TU intervals. In contrast, in the Standard condition, when time pressure was much alleviated, the intensive mental processing (e.g., problem solving activities) was more likely to occur at the intervals in between the TUs leaving the typing process less intervened. Such a sequential processing could essentially shorten the within-TU intervals leading to an accelerated typing process inside the TUs. In the Free condition, both the within-TU and between-TU intervals were lengthened, and such a lengthened within-TU interval was clearly not associated with increased parallel processing as that observed in the Short condition. Rather, it was a result of the slowdown of the overall processing speed.

This interpretation certainly comes with the prerequisite that experienced translators' mental processing in the Short condition was intensive to the extent that not only the pausing, but also the typing process had been filled with such mental processing activities. In this regard, the slightly prolonged within-TU intervals for experienced translators in the most time-pressured condition could be a signal for their enhanced parallel processing as well. Novice translators, on the other hand, may not be able to reach the high degree of parallel processing as the experienced translators do under the most time-pressured condition. Evidence from the metrics that were used to depict the degree of parallel processing in the analysis, i.e., *ST fixation duration percentage in TU* and *ST-TT fixation ratio in TU* (section 6.2.1.2), echoes this finding in that experienced translators possessed a higher proportion of ST fixation when typing the TUs than novice translators in the Short condition.

The results of parallel processing degree also suggest that time pressure especially improved experienced translators' rather than novices' level of parallel processing. This provides an indication that novice translators generally possessed a lower capability of parallel processing than experienced translators; thus, novice translators had a lower flexibility in leveraging such a capability, making it less affected by the externally imposed time pressure.

In addition to faster processing speed and higher degree of parallel processing, another aspect of automaticity at the micro-level is effortlessness, or lower effort spent on each individual processing unit (i.e., TU). This feature was examined in two relevant metrics, i.e., TU time and TU fixation. It was found that time pressure and experience did not affect TU effort measured by fixation. This means that experienced and novice translators generally allocated a comparable amount of visual attention to each

individual TU across the three time conditions. However, time pressure made a difference in the average time consumed by each TU, and it was the presence/absence of a deadline that made this difference. Both groups of participants evidently prolonged the average time spent on an individual TU when the task changed from the one with a deadline (i.e., Short or Standard condition) to the one without a deadline (i.e., Free condition). Furthermore, time condition did not affect the number of keystrokes in an individual TU for both groups which indicates that such habitual typing behaviours (i.e., the number of Chinese characters they are used to producing in a TU) cannot be easily affected by time condition. These results thus reflect a behavioural effect of deadline in translation. That is, a deadline could significantly reduce the time spent on the production of a TU (on the condition that the keystroke count in a TU is not reduced), but how tight the deadline is appears to be nonessential.

At the macro-level of automaticity, concerning the basic processes of ST comprehension and TT production in translation activities, the results from the metrics of ST/TT fixation distribution showed that experienced translators tended to pay significantly more attention to TT production than to ST comprehension. In contrast, novices allocated relatively comparable proportions of attention to ST comprehension and TT production. Time pressure did not modulate such a pattern of attention distribution which made it a marked feature as a group difference. Such a finding echoes Hvelplund's (2016, p. 11) study which also found that experienced translators tended to economise on time and effort on ST processing and only allocate as many resources to ST processing as was necessary to arrive at a reasonable understanding of the text.

Based on the automaticity theory in reading fluency (Samuels, 2006, 2013), reading is a two-step decoding-comprehension process. Since comprehension can only be achieved on the basis of decoding, reading can become more fluent when the decoding process becomes automatic which requires little attention, leaving more attention to the comprehension process. Similarly, in a translation activity where the process of TT production is fundamentally based on ST comprehension, if ST comprehension requires considerably fewer resources as a more automatic process, then more resources can be devoted to TT production. The results from the present study suggest that experienced translators invested many fewer resources to ST comprehension than to TT production, which signals a more automatic ST comprehension process in translation. Novice translators, however, devoted substantial amounts of resources to ST comprehension which almost approximated that devoted to

TT production. Correspondingly, this signals a less automatic ST comprehension process, leaving fewer resources for TT production as a result. In this regard, experienced translators demonstrated a higher automaticity level than novice translators reflected by their ST/TT attention allocation patterns.

Since time pressure did not make a difference in the participants' pattern of attention allocation, the prediction that time pressure can enhance automaticity from the aspect of attention allocation was rejected. The results obtained in the present study are also not fully in line with the findings in Sharmin et al. (2008) and Hvelplund (2011) which discovered that time pressure mainly affected translators' attention distributed to ST comprehension but did not affect their attention distributed to TT production. Although not statistically significant, the present study showed that participants tended to increase the proportion of attention on ST comprehension and decrease the proportion of attention on TT production when time pressure became more intense. This tendency generally means that it is easier for them to adapt the production process of the target text than to adapt their reading-for-comprehension under time pressure. As mentioned, since the process of TT production is essentially based on ST comprehension in translation, the results indicate that the amounts of resources required by ST comprehension could be less compressible under intense time pressure. In addition, this inconformity of the current results with the findings in previous studies also implies a difference in processing the English-Chinese translation from other language pairs. Nevertheless, as the effect of time condition on attention allocation demonstrated in the current study was not statistically significant, further validations in the future would be needed.

There is a wide consensus that repeated practice and experience are central to the automatization process and automaticity is central to the acquisition of expertise (e.g., Anderson, 2009). As practice continues, accuracy increases and the need for attention decreases (Samuels & Flor, 1997). Thus, it is not surprising to find that experienced translators displayed a higher automaticity level in ST comprehension leading to an overall more production-oriented attention allocation pattern at the macro-level as a result of their repeated practice. A highly automated (ST comprehension) process during translation has been embedded in their translation competence over time. Deckert (2017) highlighted the difficulty to overcome automaticity in terms of its effect on translation output. Similarly, time pressure failed to affect automaticity at the macro-level in the present study; this supplies further evidence for the robustness of

automaticity from the perspective of attention allocation in translation.

On the other hand, when considering each of the properties of automaticity separately, i.e., at the micro-level, time pressure played a part. Time pressure could accelerate the processing of mental activities and may engender filtering of low-priority categories of information (Miller, 1960; Maule & Hockey, 1993). Acceleration, or fast speed features one of the most important properties of automaticity. Generally, acceleration to a great extent gave rise to a higher parallel processing degree (for experienced translators) given the speeded-up working condition. There was also a reduction of time spent on each individual processing unit on average, possibly with the use of the filtering strategy (Miller, 1960) when processing under intense time pressure. Such micro-level evidence asserts the potential of time pressure as a facilitative factor for automaticity.

Overall, in response to RQ2, the behavioural metrics show that time pressure generally gave rise to more automated translation processing as evidenced from three aspects: faster processing speed, deeper parallel processing and less effort (measured by time) in each processing unit. These aspects represent the properties of automaticity at the micro-level. It was found that translators' professional experience mainly affected automaticity at the macro-level – experienced translators devoted a higher proportion of attention to TT production than to ST comprehension, as a result of their more automated ST comprehension process. Thus, it turns out that time pressure could enhance the automaticity level of the translators at the micro-level rather than the macro-level. This might result from the fact that automaticity at the macro-level is a part of the translators' core competence which can only be improved through repeated practice, but can rarely be affected by externally imposed cues.

7.2.2 Patterns of macro- and micro-strategy behaviours

The analyses of pausing behaviours suggest that novice translators took longer pauses (i.e., SPs, MPs and LPs) than experienced translators, although this group difference was not statistically significant. In addition, novice translators had more SPs and MPs but fewer LPs than experienced translators regardless of the time conditions. Based on previous studies (e.g., Hansen & Hönic, 2000; Hansen, 2002b; Jensen & Jakobsen, 2000), long pauses in translation can basically be taken as signals of macro-strategy behaviours such as problem-solving or planning activities, while SPs/MPs mainly

signify micro-strategy behaviours such as mechanical and monitoring activities. The results thus reveal that novices tended to maintain a rhythm that featured a slower pace of production with longer pauses and more micro-strategy behaviours (i.e., more SP/MPs) during the process. Experienced translators, on the other hand, were more likely to have a rhythm that featured a faster pace of production with shorter pauses and more intensive problem-solving/planning activities or macro-strategy behaviours (i.e., more LPs) during the process. As indicated by Hansen and Hönig (2000), the use of more macro-strategy behaviours (such as taking a long orientation pause) could be conducive to developing the ability to render the product more spontaneously and automatically. Such an ability was observed in experienced translators who took more LPs and meanwhile, maintained faster production speed in the present study.

For the effect of time pressure, it is not a surprise that the more time that was allocated to a task, the more and the longer the pauses the participants tended to make, which corroborated H6. In other words, both novice and experienced translators reduced the number and length of the three types of pauses to respond to intense time pressure. Still, a noticeable difference was captured between novice and experienced translators on the tendency of LP count modulated by time condition. It showed that experienced translators managed to adjust their processing style with continuously increased number of LPs (i.e., problem-solving/planning activities) when time pressure diminished. However, novice translators had almost the same number of LPs in the Free condition as in the Standard condition. Considering that the average time consumed in the Free condition by novice translators (22.77 minutes) was close to that of the experienced translators (22.62 minutes), such a stagnation of LP increase for novices in the Free condition could not result from the shortened timeframe they had taken.

Further analysis on the distribution of LPs in the drafting and end revision phases revealed that while experienced translators maintained almost comparable LPs in the drafting phase across the three time conditions, they gradually increased LPs in the end revision phase with the alleviation of time pressure. Novice translators increased LPs in both the drafting and revision phases when the time condition changed from the Short to the Standard condition, but they reduced LPs in the end revision phase in the Free condition. As such, it appears that unlike the experienced translators, the Free condition did not provoke novice translators to make use of time on more macro-strategy behaviours especially in the end revision phase – the amount of their macro-strategy behaviours in the end revision phase reached the maximum in the Standard condition

and reduced in the Free condition. Instead, they stayed with the processing style with more micro-strategy behaviours in the end revision phase when time pressure diminished. Such a difference in patterns of macro- and micro-strategy behaviours displayed by the two groups of participants could also be associated with automaticity. As mentioned before, macro-strategy behaviours are conducive to more automatic processing (Hansen & Hönig, 2000), which echoes the argument that characteristic processes such as planning could facilitate automaticity (Gass et al., 2013). Deficient macro-strategy behaviours (e.g., long pauses for planning) could be a contributing factor to the lower level of automaticity for novice translators.

The analyses on segment length suggest that experienced translators tended to have shorter segments than novice translators no matter what the time condition was. Also, the length of segments (measured by time and TU number) was affected little by time pressure for experienced translators. Shorter segments indicate more frequent LPs. This depicts a processing style of experienced translators characterised by a steady rhythm of work influenced little by time pressure with brief and brisk production activities (segments) flanked by macro-strategy activities (LPs). Furthermore, as mentioned earlier, experienced translators maintained a stable number of long pauses in the drafting phase across different time conditions. The relatively steady and fixed working rhythm especially in the drafting phase is a signal of the Knowledge Telling strategy of text production, with which the translator makes “maximum use of already existing cognitive structures and minimise(s) the extent of novel problems that must be solved” (Bereiter & Scardamalia, 1987, p. 5). In this model, the translation task is completed based on fixed and conventional schemas stored in the long-term memory, which functions as “a serviceable strategy for much routine writing” (Schilperoord, 1996, p. 30). Since the translation process proceeds mostly with “fixed cognitive structures and discourse schemas” (*ibid.*, p. 32) which can be called upon whenever they are needed, the working rhythm is more resistant to the influence of time pressure. This result is generally in line with the findings of Jensen (2000) which found that experts could apply the Knowledge Telling strategy in a more consistent way.

Relating to the results about the effect of time pressure on LP distribution to the drafting and end revision phases, experienced translators’ processing style fits well the picture described by Jensen and Jakobsen (2000, p. 112): their rhythm in raw TT production is stable and “corresponds to the time factor but is fairly independent of it”. According to Hansen (2002b), this relatively fixed working rhythm demonstrated by

experienced translators' pausing and segmentation behaviours to a great extent reflects their automaticity and routines.

However, such a relatively fixed processing style did not apply to novice translators for whom the segment length measured by TU number was more markedly affected by time pressure. Specifically, it was the change from the most stringent time condition to the other two conditions that made the difference – the segment length was shortened significantly. The alleviation of time pressure from the Short to the Standard condition made their rhythm of work approximate that of the experienced translators, i.e., with shorter segments and more LPs. In the Free condition, however, the length of segment became slightly longer, deviating from the shortest segment shown in the Standard condition, which indicated that more micro-strategy behaviours filled in the segments. The different effects of time condition on the segmentation behaviours of experienced and novice translators corroborate H7.

In addition, the increase of both micro- and macro-strategy behaviours in the drafting phase with the alleviation of time pressure could also be associated with their preference of online revisions rather than end revisions, which is discussed in detail in the following section. Such a behavioural pattern signals that with the alleviation of time pressure, novice translators tended to switch from using the Knowledge Telling strategy to the use of Knowledge Transforming strategy, with which the translator needs to constantly reprocess and transform existing knowledge in order to proceed (Bereiter & Scardamalia, 1987). As suggested by Schilperoord (1996), sustained efforts put into knowledge transforming writing in the past could result in knowledge telling writing since a rich repertoire of discourse schemas can be accrued and stored in the long-term memory through repetitive performance. The account about Knowledge Telling/Transforming strategy in writing process converges to the development of automaticity in which practice plays a central role. Drawing upon the current results, it turns out that novice translators who were of a lower automaticity level in translation tended to adopt a Knowledge Transforming strategy in a natural setting without time pressure; however, time pressure served as an expediting force for them to take on the Knowledge Telling strategy.

7.2.3 Time distribution and revision type

It was found that experienced translators took a significantly longer time on orientation

than novice translators and time pressure did not affect the length of their orientation phase. The same effect also applied to the time of orientation phase as a percentage of the total task time. However, time pressure significantly reduced participants' drafting time and their end revision time, with an exception on novice translators' end revision time in the Free condition which was shorter than that in the Standard condition. Such tendencies generally disprove H8, which predicts that intense time pressure is more likely to trim the orientation and revision phases, leaving the drafting phase relatively less affected. Thus, these results do not fully conform to the findings found in previous studies (e.g., Jensen, 2000; Kourouni, 2012) as well, which suggested that the orientation and end revision phases are more likely to be shortened under time pressure (see section 2.1.2.3).

On the other hand, time pressure generally enhanced the percentage of time allocated to drafting but reduced the percentage of time allocated to end revision for experienced translators. This effect was remarkable for the contrast between the most stringent time condition and the other two less stringent time conditions, whereas the contrast between the two less stringent time condition (i.e., Standard and Free condition) was minor. Novice translators also reduced the proportion of time allocated to drafting and increased the proportion of time allocated to end revision when time pressure was alleviated from the Short to the Standard condition. However, a prominent feature which distinguished novice translators from experienced translators was that they tended to increase the percentage of time spent on drafting and to reduce the percentage of time spent on end revision in the Free condition compared with that in the Standard condition. This results in a pattern of time allocation in the drafting and end revision phase shared by experienced and novice translators in the Short and Standard condition, but an evidently different pattern of time allocation in these two phases in the Free condition for the two groups of participants.

These results specifically echo the results on the pattern of LP numbers and preference for online/end revision behaviours modulated by time condition. It appears to be a chain reaction that novice translators increased their online revision but reduced their end revision activities remarkably from the Standard condition to the Free condition, leading to a longer drafting phase and a shorter end revision phase in the Free condition. Correspondingly, their macro-strategy behaviours increased in the drafting phase but decreased in the end revision phase in the Free condition compared with the Standard condition. Therefore, when the deadline was absent, novices opted

for more online revision behaviours while experienced translators opted for more end revision behaviours; but when the deadline was present, they tended to share a similar (online/end) revision pattern. Their choice of revision pattern in the Standard and Free condition seemed to be associated with their use of macro-strategy behaviours in these conditions as well: the moderate deadline, which was compulsory but not extremely demanding, gave the impetus for novice translators to have substantial end revision time and provoked maximum macro-strategy behaviours in this phase; however, the time constraint severely restricted experienced translators from increasing their macro-strategy behaviours. As a result, H9 was corroborated. This is also the major difference between the two groups of participants regarding their behavioural patterns in response to the presence/absence of a deadline.

In summary, time pressure unsurprisingly shortened and reduced the number of all types of pauses for both novice and experienced translators; but it functioned differently on several other aspects of behaviours (i.e., LP count, segment length, end revision time and online/end revision preference) for the two groups of participants. While they tended to share a similar behavioural pattern in the change from the Short to the Standard time condition, prominent divergences between the two groups in these behavioural aspects appeared at the change from the Standard to the Free condition. Alleviation of time pressure prolonged the end revision phase and increased the macro-strategy behaviours in this phase in a constant manner for experienced translators. Also, experienced translators' processing style was characterised by a steady rhythm of work influenced little by time pressure, with brief and brisk production activities (segments) flanked by macro-strategy activities (LPs). The Standard condition also prompted novice translators to have substantial end revision time, to shorten the segments and to provoke maximum macro-strategy behaviours, which made their behavioural pattern the most similar to that of the experienced translators. Unlike experienced translators, the Free condition did not motivate novice translators to further prolong their end revision phase and increase their macro-strategy behaviours; rather, they opted for a prolonged drafting phase with more online revision behaviours in this condition.

In addition, the distribution of LPs in the drafting and end revision phases as well as their preference for online/end revision suggest that experienced translators adopted the Knowledge Telling strategy consistently across different time conditions. In contrast, time pressure facilitated the process of switching from using the Knowledge Transforming strategy, which was the default state for novice translators when there

was no time pressure, to the use of Knowledge Telling strategy. Overall, the findings about automaticity and cognitive rhythm can serve to address the second aim of the present study which is about the effects of time pressure on translators' behavioural patterns based on their eye-key activities and specific features underpinning the experienced translators' behavioural patterns.

7.3 The stress effect on coping effort

Drawing upon the transactional model of stress (Lazarus & Folkman, 1984), the results on most psycho-physiological responses (e.g., self-reported Stress, Temporal Demand and State Anxiety and biomarkers of HR, BP and pupil size) in this study suggest that intense time pressure in executing a translation task was appraised as a factor of threat by the participants, and they had been confronted with the problem of resource shortage to handle the task with a tight deadline. The primary and secondary appraisals led to a stress effect that the tighter the deadline, the sharper the psycho-physiological reactions related to stress. Both groups of participants exhibited such a congruous stress effect when they were working on the tasks so that novice and experienced translators shared similar appraisals to the tasks of different time conditions. The lack of group difference in these results indicates that a tight deadline in a translation task constitutes a salient stressor for translators, no matter whether they are novices or experienced translators.

From the perspective of coping response in the transactional model of stress, the effort invested in the tasks was examined by different measurements (i.e., self-reported Effort, MFD, PWR, GSR and HRV). Generally, the main effect of Group was not evident in these measurements, except that novice translators showed significantly higher SCR Amplitude values than experienced translators. The effect of time pressure was not significant on the *effort* level of experienced translators as revealed by all these measurements. This implies that experienced translators maintained their problem-based coping effort or the active regulatory activity (i.e., changing the cognitive state by recruiting more resources) at a relatively stable level in the translation tasks across different time conditions.

Novice translators, by contrast, showed a more varied level of coping effort across the three tasks as indicated by the self-reported Effort, PWR and GSR. The effect of time condition on MFD and the HRV indices were not statistically significant for novices, but the pattern shown by the HRV indices was fully in line with that of the

self-reported Effort, PWR and GSR. A significant improvement of *effort* from the Short to the Standard condition was captured by the self-reported Effort, PWR and GSR; when time pressure was further alleviated from the Standard condition to the Free condition, their self-reported Effort, PWR and GSR dropped. Thus, time pressure had a more prominent influence on novice translators' coping effort invested in the translation tasks, with the effort level in the moderately time-pressured condition reaching the peak.

This pattern exhibited by novice translators in terms of effort exertion under the influence of different time pressure levels demonstrates an important feature of this group. That is, the moderately time-pressured condition might be an optimal condition for them to devote more effort to the translation task. The drastic surge of effort in this condition signals more task-relevant coping or active regulatory activities. Comparing the two conditions with different deadlines as stressors, novice translators demonstrated a trying-harder reaction by changing the cognitive state in the Standard condition, which is a direct control in regulating time pressure (Maule & Hockey, 1993). This strategy characterises the maintenance of performance goals at the expense of increased effort or recruitment of additional resources. In contrast, under the most time-pressured condition where their effort level was much lowered, novices might adopt the indirect control of changing the target state with a downward adjustment of performance goals. This features a low-effort regulation strategy under highly intense time pressure since the costs of sustained trying-harder reactions may outweigh the benefits gained in such a situation (*ibid.*). Increased use of lower-effort actions such as filtering (Miller, 1960) and a switch from compensatory to non-compensatory rules (Einhorn & Hogarth, 1981) could amount to the adoption of such a regulation strategy in the Short condition for novices.

In addition, the overall cognitive load measures (PWR, GSR Mean and RMSSD) filtered from the pool of *effort* measures were verified through testing the relationships between task performance (translation quality) and extraneous/germane load within the framework of CLT. As a result, the effect of time condition on overall cognitive load was revealed by these measures, which corresponds to the effect of time condition on *effort* indexed by these measures. Apparently, the stringency of time condition did not proportionately exacerbate the overall cognitive load as predicted in H12.

It should also be noted that the conventional indicator of *effort* (or cognitive effort), MFD, did not pronounce a significant effect of either time pressure or experience in the

present study. Nor did it display any significant correlation with other indicators of *effort* under study. This might signal its insensitivity in representing the *overall* intensity of cognitive processing in a translation task of 16-30 minutes. In other words, MFD might be more capable of reflecting the cognitive processing for local or micro-level problem triggers in a task, rather than detecting changes induced by more global or macro-level factors such as time pressure.

7.4 Cognitive activation effect and germane load

The different effect of time condition on stress level and effort exertion observed in this study underlines the nature of stress and effort as different types of activation or energy mobilisation states. Generally, effort-related energy mobilisation is associated with activation states while stress-related mobilisation is associated with disorganised states (Staal, 2004; Gaillard & Wientjes, 1994; Karasek & Theorell, 1990). From this perspective, the most time-pressured condition may represent a typical state of stress-related mobilisation as evidenced by relevant psycho-physiological responses. Such a stress-related mobilisation gradually subsides when time pressure declines. Instead of a uniform effect of time condition for the two groups of translators, the effort-related energy mobilisation exhibited different patterns for novice and experienced translators across different time conditions as evidenced by the effort related measurements. Apparently, the effort-related energy was likely to be maximally mobilised in the Standard condition for novice translators but was kept stabilised across different time conditions for experienced translators.

In another vein, based on the theory of energy mechanisms put forth by Gaillard (2001) that focuses on the way in which our body regulates states of activation needed to perform and process, the energy mobilisations typically occur under one of the following conditions: 1) task-induced stimulation in which the activation results from the task or environment itself, 2) internally guided voluntary mobilisation under a given task load, or 3) emotional arousal mobilisation induced by stress or threatening situations. Correspondingly, the emotional arousal induced by stress is most likely to underpin the activation state in the most time-pressured condition. In this case, although the participants were in an activated state, there was only limited capacity available for processing the task-relevant information. In the moderately time-pressured task, however, their energy was more likely to be mobilised by the task-induced stimulation.

Since time pressure was not intense but still present in this situation, more resources or capacity were available for processing the task-relevant information leading to a higher engagement in the task. In the non-time-pressured condition, the externally imposed time pressure diminished, so that their emotional arousal was lowered, and their energy mobilisation would rest more on their internally guided voluntary effort. Thus, motivation played an essential role under such a condition.

The pattern of effort related energy mobilisation exhibited by novice translators in the Standard condition gave rise to the cognitive activation effect of time pressure on this group of participants. A cognitively activated state in task execution can stimulate the cognitive functioning and processing, and engage the participants in meaningful and higher-order thinking (Teig et al., 2019; Baumert et al., 2010; Klieme et al., 2009; Lipowsky et al., 2009; Depaepe & König, 2018). Drawing upon the framework of CLT, such a cognitive state could be associated with the functioning of germane load which is a type of effective load facilitating task performance and contributing to the overall cognitive load in the meantime.

Since the intrinsic load (i.e., task difficulty) is assumed to be comparable in the three tasks, the effect of time condition on the sum of the intrinsic and germane load (IGL, which is the difference between overall cognitive load and extraneous load) can generally reflect the variation of germane load modulated by time condition. As revealed by the results, the effect of time condition on IGL was different for the two groups of participants. Novices' germane load was at a relatively low level in the Short condition but soared in the Standard condition and remained unchanged in the Free condition, whereas a gradual increase of germane load was captured with the alleviation of time pressure for experienced translators. This pattern conformed with the effect of time condition on translation quality. Thus, H13, which predicts that the effect of time condition on germane load would be similar to its effect on translation quality for the two groups of participants, is corroborated. Further, it is found that certain behavioural metrics such as LP and end revision behaviours might be of particular interest in gauging the germane load levels in translation and thus be indicative of translation quality.

Further analysis revealed that time pressure significantly regulated the functioning of germane load (on translation quality) for novice translators. It was found that the constructive nature of germane load to performance was largely undermined in the Free condition for novices. This effect, however, did not appear in experienced translators –

the functioning of germane load worked effectively on them regardless of the time conditions. Thus, H14 is partially corroborated (i.e., only applicable to novice translators). As mentioned, different mechanisms of energy mobilisation, i.e., task-induced stimulation and internally guided voluntary mobilisation, underline the tasks of Standard and Free conditions respectively. This means that novice translators' germane load in the state of the internally guided voluntary mobilisation may have an inefficiency problem. In other words, the lack of a task-induced stimulation (e.g., a deadline) might account for an undermined effectiveness of germane load in the Free condition for novice translators.

Thus, although the alleviation of time pressure led to a cognitively activated state with a relatively high level of germane load for novices, it impaired the facilitative function or effectiveness of germane load to a great extent. To be specific, although the constructive effect of germane load on performance functioned the best in a highly time-pressured condition, participants could not exert much germane load because the extraneous load was also high in such a condition; on the other hand, when they managed to exert more germane load in the condition without an extraneous load induced by time pressure, the effectiveness of germane load was largely impaired. As a result, it is essential to consider balancing the degree of time pressure to the extent which can optimally stimulate and activate the cognitive processing while maintaining the effectiveness of germane load for novice translators. Such a concern did not apply to experienced translators because time pressure did not modulate the effectiveness of germane load on their performance. In principle, the more the germane load they exert, the better the performance they could achieve. As a result, they performed optimally in the Free condition where their germane load peaked without much extraneous load induced by time pressure taking up the limited capacity.

7.5 Expertise reversal effect

Some conducive instructional elements which are designed for facilitating novices' performance could be destructive to experts. This is called an expertise reversal effect. While the lack of a task-induced stimulation evoked different effects in the Free condition on the effectiveness of germane load of the two groups, their difference in effort exertion in the Standard condition reflected an expertise reversal effect. When relating the translation quality to their coping effort and germane load levels, it was

evident that the moderately time-constrained condition was conducive to novice translators' performance, but this condition was suboptimal for experienced translators.

As shown in the above analysis, the task-induced stimulation or the moderately time-pressured condition functioned as a type of effective instructional design for novices. Thus, it is likely that such a condition can facilitate schema construction and automation through substantially reduced extraneous cognitive load, as well as ample germane load. In other words, the output of the cost of extraneous load (induced by the moderate time pressure) and the benefit of germane load was optimal for novices in the Standard condition. The existence of the task-induced stimulation under such a condition fuelled their motivation and engagement in the task and the functioning of germane load was not undermined. As a consequence, novices could benefit from such translation practices with a moderate deadline which can act as a means of schema construction and automation.

In the Free condition which lacked the task-induced stimulation but was solely driven by the internally guided voluntary mobilisation, novices also managed to invest a considerable amount of germane load which was comparable to that invested to the Standard condition. As mentioned, however, the effectiveness of germane load under such a condition was largely undermined and their translation quality was not improved further in the Free condition. This implies that, although the conventional means of instruction design for effective learning mainly rests on reducing the working memory load or the ineffective extraneous load to facilitate schema construction and automation (Kalyuga et al., 2003), the situation in the current study indicates another possibility. That is, a trade-off effect may exist between the cost of the extraneous load elicited by time pressure and its potential gain which is the facilitative function of a task-induced stimulation on the effectiveness of germane load (for the group of novice translators). In the Standard condition, the gains of the task-induced stimulation outweighed the costs of the extraneous load induced by the time constraint, making it an optimal condition for novice translators in their task performance.

In spite of the plausible constructive effect of such an instructional procedure (i.e., a moderate time constraint) for novices which may facilitate their schema construction and automation, this procedure was not needed by experienced translators whose schemas already provided full guidance for task processing. As mentioned, experienced translators' germane load peaked at the Free condition and time pressure did not modulate the effectiveness of germane load, which made them achieve the best

performance or translation quality in the Free condition. The facilitative function of the task-induced stimulation in the Standard condition displayed on the group of novice translators did not appear in the experienced translators. This characterised an expert reversal effect in the Standard condition. Clearly, experienced translators exhibited a strong commitment to each task, either with or without a deadline, which constitutes an important part of their expertise. Indeed, attitudinal resources such as perseverance, initiative and motivation are included in the psycho-physiological components of the translation competence model developed by the PACTE Group (Albir, 2017). These aspects of translation competence have manifested in the group of experienced translators in the current study. In this regard, an external or task-induced stimulation such as a deadline would only hinder their performance by increasing the ineffective extraneous load.

7.6 Time pressure and translation performance

As discussed by Staal (2004, p. 14), a deadline imposed on a task could carry both direct and indirect effects: the time limit is only a physical boundary that does not require any psychological explanation in understanding its direct effects on performance; but there are often corresponding psycho-physiological consequences that could evoke secondary or indirect effects on performance. The analyses in the previous sections in this chapter have focused on such psycho-physiological and behavioural effects of time pressure and have tapped into further consequences of these effects on the translation product, which represent the indirect effects of time pressure on task performance. For example, it was found that time pressure affected the effectiveness of germane load differently for the two groups of participants, the consequences of which led to the different performance of the two groups. The direct effect of time pressure on translation product quality can be examined disposing of such mediators.

To a certain extent, the results of translation quality regulated by time condition for novice translators revealed a stress-performance relation that corroborates the Yerkes-Dodson Law (Yerkes & Dodson, 1908). Novice translators had optimal performance under the moderately time-pressure condition with a slight downswing in the Free condition and a great fall-off in the Short condition. As such, it is possible that for novice translators, the so-called optimal stress level for performance may exist at the Standard time condition or in between the Standard and the Free condition (i.e., with a

more relaxed deadline than the one in the Standard condition). Experienced translators, however, exhibited a linear negative relationship between the stringency of time condition and translation quality in this study. Assuming that an optimal stress level exists as depicted in the Yerkes-Dodson Law, these results indicate that the threshold of such an optimal stress level was higher for novices than for experienced translators (Figure 7-1). This points to the fact that the “good stress” or “eustress” in need to stimulate novice translators to perform better was not similarly favourable to experienced translators. The expertise reversal effect as discussed in section 7.5 could provide a possible explanation for such a result.

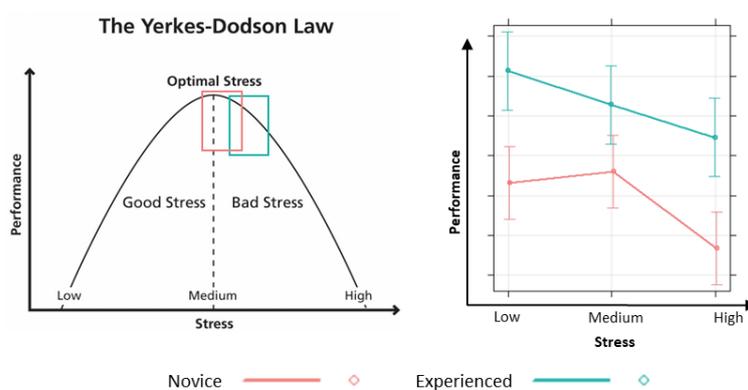


Figure 7-1. Illustration of Yerkes-Dodson Law (left) in relation to the (flip-over presentation of) the stress-performance relation identified in the present study (right)

On the other hand, although novice translators achieved optimal performance (translation quality) in the moderately time-pressured condition, the reduction in the Free condition was not significant. In other words, their performance in the Free condition could be comparable to that in the Standard condition. In this sense, a negative impact of time pressure on both groups was observed in general, although novice translators achieved comparable translation quality in the Standard and Free condition. This featured a negative effect of stress on performance which corroborated H10 and was in line with most previous studies (e.g., Kourouni, 2012; De Rooze, 2003).

Such a negative relationship between time pressure and task performance can be linked to the speed/accuracy trade-off effect whereby participants tend to respond more quickly to the task demands but create more errors (Fitts, 1966). This general phenomenon has been reported in De Rooze (2003) and Ghobadi et al. (2017) on the

relation between time pressure and translation performance, as well as across a wide variety of cognitive domains and under multiple sources of stress (e.g., Driskell et al., 1992; Schellekens et al., 2000; Hygge & Knez, 2001). It represents a “teeter-totter” effect as movement in one end most often results in a corresponding movement in the other (Staal, 2004). The results on both processing speed and pauses in this study showed that time pressure indeed accelerated the processing for both groups, which provided evidence for the use of acceleration (Miller, 1960) as a coping strategy in response to the increased task demands. Overall, acceleration of the translation process and the degraded translation quality under high time pressure characterised a speed/accuracy trade-off effect which can also be seen as a direct effect of time pressure.

As mentioned in Staal (2004), it is usually difficult to distinguish between the direct and indirect effects of time pressure or stress on task performance. In addition, concentrating on the direct effects might give no more than a simplified picture of the one-dimensional relationship between the two variables, i.e., time pressure and product quality. In the present study, some traces of an inverted-U shaped relation were shown for novice translators; but a more evident negative relationship between time pressure and the final product was visible, especially for the experienced translators. Nevertheless, the underlying mechanisms accounting for such relationships could be highly complex which can make the sketchy relationship vary case by case. Thus, looking beyond the direct effect, by understanding and exploiting the underlying mechanisms, is essential for explaining the relationship between time pressure and translation performance.

Overall, the last aim of the present study is addressed. It was found that while experienced translators were affected little by time condition in effort exertion, novice translators tended to have a significantly higher *effort/overall* cognitive load level in the Standard condition than in the other two conditions. In addition, novices’ germane load was at a relatively low level in the Short condition but soared in the Standard condition and remained unchanged in the Free condition, whereas a gradual increase of germane load was captured with the alleviation of time pressure for experienced translators. These tendencies conformed to the effect of time condition on translation quality, which can be generally seen as a negative relation. Time pressure also regulated the facilitative effect of germane load for novice translators (but not for experienced translators), leading to a largely undermined functioning of germane load in the Free condition which lacked a task-induced stimulation. Such an impaired effectiveness of germane

load modulated by time condition for novice translators, in addition to the effects of time condition on overall cognitive load and germane load levels, as well as the expertise reversal effect could supply some insights for the complex underlying mechanisms accounting for the relationship between time pressure and translation quality.

Chapter 8. Conclusion

This thesis explored the effects of time pressure on the cognitive process of translation demonstrated by two groups of translators from psycho-physiological, behavioural-cognitive and performance-based perspectives. By triangulating a wide range of cutting-edge methods, including eye tracking, keystroke logging, psychometric instruments, as well as measurements of several biomarkers, empirical data generated from the translation (English-Chinese) experiment were analysed to address the following three specific aims:

Aim 1: to determine measures of time pressure/stress during translation with translators' psychological and physiological responses being tested as possible indicators, and to recognise possible differences between novice and experienced translators in these responses.

Aim 2: to examine the effects of time pressure on translators' behavioural patterns about automaticity and cognitive rhythm based on their eye-key activities and identify specific features underpinning the experienced translators' translation behaviours.

Aim 3: to explore the effects of time pressure on effort/cognitive load and reveal possible underlying mechanisms accounting for the relationship between time pressure and translation quality by looking into the functioning of cognitive load, especially germane load based on Cognitive Load Theory.

Five research questions (RQs) were formulated with corresponding hypotheses (H) to shape and frame the analyses in response to the research aims. The psycho-physiological responses (RQ1 and H1-H3) were examined in terms of self-reported Stress, State Anxiety, Temporal Demand and all the biomarkers under study (i.e., HR, BP, pupil size, skin temperature, GSR and HRV). Several of these measures were confirmed valid in indicating the time pressure/stress levels during translation, based on which the first aim was achieved. The focus of the second aim, i.e., translators' behavioural patterns, was investigated from the perspectives of automaticity (RQ2 and H4-H5) and cognitive rhythm (RQ3 and H6-H9) during translation. A number of behavioural metrics generated from the eye-key activities designated for each aspect under examination were involved. A series of findings were generated with regards to the effect of time pressure on the translation behavioural patterns of the two groups of

translators, which provided a closer inspection of the translation process as a cognitive activity.

This investigation took Cognitive Load Theory (CLT) as a theoretical foundation in order to discover the interplay between time pressure, cognitive load and translation quality, which was the third aim of this thesis. Before looking into the underlying mechanisms regarding the functioning of cognitive load, the direct effects of time pressure and experience on translation quality was examined (RQ4 and H10-H11). The construct of (overall) cognitive load was established strictly conforming to the relationships assumed within the framework of CLT, such as the negative relationship between extraneous load and performance and the positive relationship between germane load and performance. *Cognitive load* was thus distinguished from the construct *effort* (which refers to the resources recruited in coping with the task demand under certain conditions) in the present study since *effort* was not under the scope of such presumed relationships with performance. The commonly used self-report scales and behavioural and physiological indicators of *effort* were examined with regards to the influence of time pressure and experience; they were then filtered via the analytical method established upon the assumptions of CLT to have the overall cognitive load indicators determined and the germane load indicators generated. These procedures enabled the examination and revelation of the interplay between time pressure, germane cognitive load and translation quality (RQ5 and H12-H14).

This chapter concludes the whole study with a summary of all the key findings relating to the research questions and hypotheses (section 8.1). Considerations on how the findings may be applied to areas of translation practice and translators' training are presented in section 8.2. It is followed by a section (8.3) outlining some strengths and limitations of the whole study and a final section (8.4) providing suggestions and possible directions for future research.

8.1 Summary of the major findings

The results obtained from all the psychological measures (self-reported Stress, State Anxiety and Temporal Demand) indicated that the effect of time pressure on both novice and experienced translators were prominent, and there was little divergence between groups. The same effect was captured by biomarkers of HR, BP and pupil size. Most of, though not all, these psycho-physiological measures were correlated with each

other. These results consolidated their convergence in reflecting the time pressure levels experienced by translators during translation. These measures are thus verified as time pressure measures in translation activities. It was also found that novices did not have stronger psycho-physiological responses than experienced translators when the time condition became increasingly stringent. It is possible that the superior capability of emotion regulation and stress management, which is supposed to have been attained by experienced translators, diminished under extremely stressed conditions. This means that while time pressure exacerbated relevant psycho-physiological responses significantly, translators' experience did not make a difference.

On the other hand, the biomarkers of skin temperature, GSR and HRV did not show the same effect of time pressure as shown by the other psycho-physiological measures mentioned above. Skin temperature was affected little by time pressure for both groups; GSR and HRV were identified as physiological indicators of effort/overall cognitive load, and they displayed different effects of time pressure on the two groups of translators which conformed to that shown by the behavioural and self-reported indicators of effort/overall cognitive load (e.g., PWR and self-reported Effort). An inverted "U-shaped" relationship between time stringency and the effort level was exhibited in novice translators in this pattern, while experienced translators' effort level was almost unaffected by time pressure.

In regard to automaticity in translation, the result confirmed that time pressure generally gave rise to more automated translation processing at the micro-level. This was evidenced from three aspects: faster processing speed, deeper parallel processing and less effort in each processing unit. With time condition getting more stringent, translators generally accelerated the processing with more keystrokes produced per minute and shorter intervals between and within the typing units; their parallel processing degree was improved, especially for experienced translators, with a higher proportion of attention devoted to ST comprehension during the course of producing a TU; the amount of effort (measured by time) devoted to each individual TU was also reduced in the time-constrained tasks.

In particular, the slightly prolonged within-TU intervals exhibited by experienced translators in the most time-pressured condition signalled their intensive mental processing during the typing process in this condition, which implied their enhanced parallel processing as well. In addition, time pressure especially improved experienced translators' rather than novices' level of parallel processing. Since novice translators

were generally of a lower automaticity in translation, they had lower flexibility or buffering capacity in leveraging the capability of parallel processing (as one aspect of automaticity); thus, novice translators' parallel processing degree was less affected by externally imposed time pressure.

It was found that translators' professional experience made a difference on automaticity at the macro-level – experienced translators devoted a higher proportion of attention to TT production than to ST comprehension because of their more automated ST comprehension process. Overall, it turns out that time pressure could only enhance the automaticity level of the translators at the micro- rather than the macro-level. This might be attributed to the fact that automaticity at the macro-level is more of a part of the translators' core competence which can only be improved through repetitive practice but can rarely be affected by externally imposed cues such as momentary time pressure.

Experienced translators took a significantly longer time on orientation than novice translators, and time pressure did not affect the length of their orientation phase. The same effect was also captured in the time of orientation phase as a percentage of the total task time. However, intense time pressure significantly reduced participants' drafting time and their end revision time, with an exception for novice translators' end revision time in the Free condition which was shorter than that in the Standard condition.

Alleviation of time pressure extended the end revision phase and increased the macro-strategy behaviours in a consistent manner for experienced translators. Also, experienced translators' processing style was characterised by a faster pace and a relatively steady rhythm of work influenced little by time pressure, with brief and brisk production activities (segments) flanked by macro-strategy activities (LPs). The Standard condition also prompted novice translators to have substantial end revision time, to shorten the segments and to provoke maximum macro-strategy behaviours, which made their behavioural pattern the most similar to that of the experienced translators. Unlike experienced translators, however, the Free condition did not motivate novice translators to further extend their end revision phase and increase their macro-strategy behaviours; rather, they opted for a shorter end revision phase and a prolonged drafting phase with more online revision and micro-strategy behaviours in this condition.

This means that when a deadline was absent, novices tended to opt for more online revision behaviours while experienced translators tended to opt for more end revision

behaviours; but when the deadline was present, they tended to share a similar (online/end) revision pattern. Their choice of revision pattern in the Standard and Free condition seemed to be associated with their macro-strategy behaviours in these conditions as well: the moderate deadline, which was compulsory but not extremely demanding, gave the impetus for novice translators to have substantial end revision time and provoked maximum macro-strategy behaviours in this phase; however, the time constraint only restricted experienced translators from increasing their macro-strategy behaviours. In addition, the distribution of LPs in the drafting and end revision phases as well as their preference for online/end revision suggest that experienced translators tended to adopt the Knowledge Telling strategy consistently across different time conditions. In contrast, time pressure can facilitate the process of switching from using the Knowledge Transforming strategy, which was the default state for novice translators when there was no time pressure, to using the Knowledge Telling strategy.

Experienced translators maintained their coping effort or the active regulatory activity at a relatively stable level in the translation tasks across different time conditions. However, time pressure had a more prominent influence on novice translators' coping effort invested in the translation tasks, with the effort level in the moderately time-pressured condition reaching a peak. Apparently, the moderately time-pressured condition might be an optimal condition for novice translators to devote more effort to the translation task. The drastic surge of effort in this condition (compared with the other two conditions) might signal more task-relevant coping or task-relevant active regulatory activities in this condition.

In addition, novices' germane load was at a relatively low level in the Short condition but soared in the Standard condition and remained unchanged in the Free condition, whereas a gradual increase of germane load was captured with the alleviation of time pressure for experienced translators. These tendencies conformed to the effect of time condition on translation quality, which can be generally regarded as a negative relation, despite traces of Yerkes-Dodson Law shown on the group of novice translators. Besides, LP and end revision behaviours were of particular interest in gauging the germane load levels in translation and thus were indicative of translation quality as well.

Time pressure also regulated the facilitative effect of germane load for novice translators, leading to a largely undermined functioning of germane load in the Free condition, which lacks a task-induced stimulation. In other words, the lack of a task-induced stimulation (e.g., a deadline) might account for the undermined effectiveness

of germane load in the Free condition for novice translators. Thus, for novice translators, it is essential to consider balancing the degree of time pressure to the extent which can optimally stimulate and activate cognitive processing while maintaining the effectiveness of germane load. Such a concern does not apply to experienced translators because time pressure did not modulate the effectiveness of germane load on their performance. In principle, the more the germane load they exert, the better the performance they could achieve regardless of the time conditions. This is also a reflection of the expert reversal effect. As a result, experienced translators performed optimally in the Free condition where their germane load peaked without much extraneous load induced by time pressure taking up the limited capacity.

8.2 Implications for translation practice and pedagogy

The pronounced relationships found between time stringency and stress, and between time stringency and anxiety, can provide some didactic implications for translation practice and translator training. On the one hand, the more translators are aware of the influence of stressors such as tight deadlines in their daily work, the better they might consciously manage them. On the other hand, stress and anxiety are not completely destructive in task performance. A number of studies have shown that negative emotions such as anxiety can enhance specific aspects of the translation product, such as accuracy and coherence (e.g., Rojo & Ramos, 2016). Although the present study does not involve an examination of such specific aspects of the translation product, this could be of interest to translation training practices. For example, time pressure as a convenient means of triggering emotions such as anxiety could be utilised appropriately in the translation training process.

It is found that intense time pressure could give rise to faster heartbeat, elevated blood pressure and increased anxiety for both groups of translators in the present study. As one of the major stressors at work, exorbitant time pressure can lead to devastating consequences on our mental and physical health. It is important for employers and practitioners to be aware of the presence of stress and its consequences. This could also benefit the practitioners and trainers regarding stress perception and management. In addition, by utilising the validated measures of stress/time pressure in the future, the dynamic stress level of translators at work can be objectively detected and monitored, which could help to keep stress within a healthy limit and consequently boost

performance and productivity.

The comparison of behavioural patterns between novice and experienced translators in this thesis demonstrates some marked features of experienced translators, which could shed some light on translation practice and training. Firstly, experienced translators tend to maintain a faster pace of work than novice translators on account of their superior automaticity. A highly time-pressured written translation task could resemble sight translation or even interpreting, which are inherently time-pressured and fast-paced tasks and are often considered as more automated processes. As indicated by previous research (e.g., Hansen & Hönig, 2000; Hansen, 2002b; De Rooze, 2003), improving the processing speed in translation can in turn, benefit the development of automaticity. Thus, more attention could be paid to improving novice translators or translation students' processing speed under time pressure in translation training which could potentially contribute to the development of automaticity.

Secondly, unlike novice translators, experienced translators' end revision time and macro-strategy behaviours consistently increase when time pressure diminishes. In fact, as mentioned earlier, novice translators achieve the best performance in the Standard condition where their behavioural pattern comes the closest to that of experienced translators, especially regarding end revision and macro-strategy behaviours. Thus, novice translators should consciously attend more to their end revision and macro-strategy behaviours, especially when there is no time pressure.

Another important finding of this thesis is that being free of time pressure largely undermines the effectiveness or constructive nature of germane load for novice translators. This means that a condition without any task-induced stimulation is not optimal for novice translators' performance. Such a finding is particularly beneficial to translation pedagogy since it provides empirical evidence of the role cognitive load, especially germane load, plays in the cognitive process of translation and offers valuable implications for instructional design in settings such as translation examinations and modules. To be specific, balancing the degree of time pressure to the extent which can optimally stimulate and activate cognitive processing while maintaining the effectiveness of germane load, is essential for novices or students to reach their optimal performance. On the other hand, an expertise reversal effect is confirmed in this study, which means that for experienced translators, time pressure only hinders their performance. This finding offers some practical implications to professional translators themselves as well as employers (and even clients) in the

translation industry. For example, it is best to avoid imposing excessively tight time constraints on translation tasks or work since this might largely impair the productivity and product quality of professional translators.

8.3 Strengths and limitations

The present research involved a variety of data types generated by different cutting-edge methods in order to systematically explore the effects of time pressure on translation. Such a multi-methodological approach combining diversified types of data constitutes one of the major strengths of this study. It allows the production of more reliable results and a more comprehensive examination of the topic. Especially, the involvement of several types of physiological data allows a number of significant issues which have not been previously touched upon in related work, to be addressed. These include the psycho-physiological responses of stress during translation and the connections between stress and effort/cognitive load in written translation. The employment of a variety of cutting-edge methods and abundant findings generated in the present study can also make a contribution to the interdisciplinary expansion of Translation Studies. It broadens the interface between Translation Studies and the neighbouring subjects such as learning sciences and cognitive psychology and entails a collision between the psycho-physiological and behavioural-cognitive explorations of the translation process.

In addition, the adoption of CLT as a theoretical framework in the present study constitutes a powerful and effective analytical tool in estimating different types of cognitive load such as overall cognitive load, germane load and extraneous load. Combining the use of diverse psycho-physiological and behavioural measures of stress and effort, this theoretical and methodological application of CLT may lend convenient support to future research on diverse stressors in translation or other cognitively demanding language processing activities. At the same time, in the context of the learning sciences and instructional design, which underlies the nature of CLT, the findings yielded from the present study offer practical suggestions and implications for translation pedagogy which are backed up with solid empirical evidence. In this regard, the value of the applicability of the findings to practice (for translation instructors as well as practitioners in the translation industry) is another marked strength of the present study.

While the sampling population of the present study is of a relatively good size, i.e., 65 participants (35 novice and 30 experienced translators) in the main experiment, the relatively high discard rate of eye tracking data for the group of experienced translators restricts the power of this sample to some extent. This could be regarded as a major limitation of the study. Most discarded eye tracking data files were from the 13 experienced translators whose experiments were conducted with a portable eye tracker and in a relatively unstable environment (i.e., not in the university's eye tracking laboratory). Thus, improvements can be made in this regard for future research. For example, the sampling population could be further enlarged, especially for the group of experienced translators; sample size in each group could be made more balanced and it may be better to conduct the eye tracking experiment in a more controllable environment to ensure higher data quality.

Further, the controlled laboratory settings may inevitably raise some concerns about ecological validity. For example, although the participants could consult any external resources regarding the unfamiliar words provided before the experiment, they were not allowed to access resources during the experiment. This is because the focus of the present study was on the cognitive processes of translation only. This constraint, however, may alter translators' decision-making behaviours compared to what they do during their routine work when resources are at hand. Thus, it would be interesting to involve consultation and documentation in future research.

Besides, the duration of the baseline sessions (3 minutes) for the measurement of continuous biomarkers appears to be short, which, to a great extent, entails a data loss problem for HRV and GSR in some (baseline) sessions. Thus, to reduce the rate of invalid data, slightly longer baseline sessions and more careful guidance to the participants, especially when involving measurements of HRV and GSR with a portable device such as Empatica E4 Wristband are recommended for future research.

8.4 Avenues for future research

One of the major foci of the present study is to examine the change of the overall state in stress levels and the intensity of cognitive processing experienced by a translator in a translation task under different time conditions. This means that for a number of (continuously recorded) physiological and behavioural measures (e.g., HR, pupil size, GSR Mean and MFD), the averaged values are used to represent such a state of a

translation task. It would be of interest in future research to have multiple measures synchronised during the course of a task to inspect the dynamic changes throughout the task session. Although such procedures would be laborious and would come with higher technical requirements such as the registration of timestamps and alignment of various measures, this could allow for analysis of more local changes elicited by particular cues, for example, in the source or target text. In addition, synchronising multiple measures throughout a task might be able to yield more fine-grained findings with regards to the verification of various time pressure/stress and effort/cognitive load indicators.

As mentioned earlier, while the present study concentrated on the role time pressure plays in the genuine or “pure” cognitive process of translation, this might deviate a little from the reality of translation nowadays. With the fast development of artificial intelligence, which has been massively applied in translation technologies in recent years, consultation or seeking help from various computer-assisted translation tools or (neural) machine translation has been a routine for most translators. Thus, the effects of time pressure on the hybrid process of computer-assisted translation or machine translation post-editing process are worth investigating. Such investigations, if conducted successfully in the future, would also allow a comparison of the effects of time pressure on different modes of the translation process, i.e., traditional translation, the hybrid process of computer-aided translation and post-editing. In addition, conducting this kind of research in a more naturalistic rather than the experimental setting in the future, for instance at the translators’ workplace or home, would provide a more comprehensive picture of the translators’ cognitive processes.

Lastly, while the current study adopted a holistic assessment of translation quality, it would be of interest to extend the scope to different aspects of translation quality, such as accuracy and creativity, and to inquire how these aspects are correlated to levels of time pressure/stress and effort/cognitive load. Such explorations could further enrich the findings, especially in relation to the effects of time pressure on the translation product, and thus could provide more pragmatic implications for translation practice and pedagogy.

Bibliography

- Akselrod, S., Gordon, D., Ubel, F. A., Shannon, D. C., Berger, A. C., & Cohen, R. J. (1981). Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control. *Science*, *213*(4504), 220-222. <https://doi.org/10.1126/science.6166045>
- Albir, A. H. (Ed.). (2017). *Researching translation competence by PACTE group*. John Benjamins. <https://doi.org/10.1075/btl.127>.
- Allen, R. D., Hitt, M. A., & Greer, C. R. (1982). Occupational stress and perceived organizational effectiveness in formal groups: An examination of stress level and stress type. *Personnel Psychology*, *35*(2), 359-370. <https://doi.org/10.1111/j.1744-6570.1982.tb02201.x>
- Alves, F. (2003). Triangulation in process oriented research in translation. In F. Alves (Ed.), *Triangulating translation: Perspectives in process oriented research* (pp. vii-x). John Benjamins
- Alves, F., & Liparini Campos, T. (2009). Translation technology in time: Investigating the impact of translation memory systems and time pressure on types of internal and external support. In S. Göpferich, A. L. Jakobsen, & I. M. Mees (Eds.), *Behind the mind: Methods, models and results in translation process research* (pp. 191-218). Samfundslitteratur.
- Alves, F., Pagano, A., & da Silva, I. (2009). A new window on translators' cognitive activity: Methodological issues in the combined use of eye tracking, key logging and retrospective protocols. In I. M. Mees, F. Alves, & S. Göpferich (Eds.), *Methodology, technology and innovation in translation process research* (pp. 267-291). Samfundslitteratur.
- Alves, F., & Vale, D. (2009). Probing the unit of translation in time: Aspects of the design and development of a web application for storing, annotating, and querying translation process data. *Across Languages and Cultures* *10*(2), 251-273. <https://doi.org/10.1556/Acr.10.2009.2.5>
- Amini, M. (2018). How to evaluate the TEFL students' translations: Through analytic, holistic or combined method? *Language Testing in Asia*, *8*(1), 10. <https://doi.org/10.1186/s40468-018-0063-6>
- Anderson, J. R. (2009). *Cognitive psychology and its implications* (7th ed.). Worth Publishers.
- Artino, A. R., Jr. (2008). Cognitive load theory and the role of learner experience: An abbreviated review for educational practitioners. *AACE Review (formerly AACE Journal)*, *16*(4), 425-439.
- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*(1), 403-450. <https://doi.org/10.1146/annurev.neuro.28.061604.135709>
- Ayres, P., & Paas, F. G. W. C. (2012). Cognitive load theory: New directions and challenges. *Applied Cognitive Psychology*, *26*(6), 827-832.

<https://doi.org/10.1002/acp.2882>

- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R* (3rd ed.). Cambridge University Press.
- Baghi, H., & Khoshsaligheh, M. (2019). Stress in written and sight translation in training setting. *Hikma*, 18(2), 237-255. <https://doi.org/10.21071/hikma.v18i2.11539>
- Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, 8, Article 456. <https://doi.org/10.3389/fpsyg.2017.00456>
- Balling, L., Hvelplund, K., & Sjørup, A. (2014). Evidence of parallel processing during translation. *Meta*, 59(2), 234-259. <https://doi.org/10.7202/1027474ar>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using Lmer4. *Journal of Statistical Software* 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133-180. <https://doi.org/10.3102/0002831209345157>
- Bayer-Hohenwarter, G. (2009). Methodological reflections on the experimental design of time-pressure studies. *Across Languages and Cultures* 10(2), 193-206. <https://doi.org/10.1556/Acr.10.2009.2.2>
- Bell, R. T. (1991). *Translation and translating: Theory and practice*. Longman.
- Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods* 190(1), 80-91. <https://doi.org/10.1016/j.jneumeth.2010.04.028>
- Block, R. A., & Reed, M. A. (1978). Remembered duration: Evidence for a contextual-change hypothesis. *Journal of Experimental Psychology*, 4(6), 656-665. <https://doi.org/10.1037/0278-7393.4.6.656>
- Broadbent, D. E. (1971). *Decision and stress*. Academic Press.
- Brown, T. M., & Miller, C. E. (2000). Communication networks in task-performing groups: Effects of task complexity, time pressure, and interpersonal dominance. *Small Group Research*, 31(2), 131-157. <https://doi.org/10.1177/104649640003100201>
- Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 53-61. https://doi.org/10.1207/S15326985EP3801_7
- Buchweitz, A., & Alves, F. (2006). Cognitive adaptation in translation: An interface between language direction, time, and recursiveness in target text production. *Letras de Hoje*, 41(2), 241-272. <https://revistaseletronicas.pucrs.br/index.php/fale/article/view/601>
- Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.),

- Language production: Vol. 1. speech and talk* (pp. 155–176). Academic Press.
- Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of Management Review*, 13(1), 40-52. <https://doi.org/10.5465/amr.1988.4306775>
- Campbell, S., & Wakim, B. (2007). Methodological questions about translation research: A model to underpin research into the mental processes of translation. *Target*, 19(1), 1-19. <https://doi.org/10.1075/target.19.1.02cam>
- Carl, M. (2012). Translog-II: A program for recording user activity data for empirical reading and writing research. The Eighth International Conference on Language Resources and Evaluation (LREC-2012) Istanbul, Turkey.
- Carl, M., & Dragsted, B. (2012). Inside the monitor model: Processes of default and challenged translation production. In O. Czulo & S. Hansen-Schirra (Eds.), *Crossroads between contrastive linguistics, translation studies and machine translation: TC3 II* (pp. 5-30). Language Science Press. <https://doi.org/10.5281/zenodo.1019685>
- Carl, M., Dragsted, B., & Jakobsen, A. L. (2011). A taxonomy of human translation styles. *Translation Journal*, 16(2). <https://translationjournal.net/journal/56taxonomy.htm>
- Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S. Z., Khawaji, A., & Conway, D. (2016). *Robust multimodal cognitive load measurement*. Springer.
- Chesterman, A. (2009). The name and nature of translator studies. *Hermes* 22(42), 13-22. <https://doi.org/10.7146/hjlc.v22i42.96844>
- Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21-30). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816796.002>
- Chiang, Y.-N. (2009). Foreign language anxiety in Taiwanese student interpreters. *Meta*, 54(3), 605-621. <https://doi.org/10.7202/038318ar>
- Chiang, Y.-N. (2010). Foreign language anxiety and student interpreters' learning outcomes: Implications for the theory and measurement of interpretation learning anxiety. *Meta*, 55(3), 589-601. <https://doi.org/10.7202/045079ar>
- Cho, D., Ham, J., Oh, J., Park, J., Kim, S., Lee, N.-K., & Lee, B. (2017). Detection of stress levels from biosignals measured in virtual reality environments using a kernel-based extreme learning machine. *Sensors*, 17(10), 2435. <https://doi.org/10.3390/s17102435>
- Choi, H.-H., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (2014). Effects of the physical environment on cognitive load and learning: Towards a new model of cognitive load. *Educational Psychology Review*, 26(2), 225-244. <https://doi.org/10.1007/s10648-014-9262-6>
- Cifuentes-Férez, P., & Fenollar-Cortés, J. (2017). On the impact of self-esteem, emotion regulation and emotional expressivity on student translators' performance. *Vigo International Journal of Applied Linguistics*, 14(1), 71-97.
- Cifuentes-Férez, P., & Meseguer Cutillas, P. (2018). Can self-esteem and creative

- intelligence foster accuracy and creativity in professional translators? *Translation, Cognition & Behavior*, 1(2), 341-360. <https://doi.org/10.1075/tcb.00015.cif>
- Cohen, J. D., Aston-Jones, G., & Gilzenrat, M. S. (2004). A systems-level perspective on attention and cognitive control: Guided activation, adaptive gating, conflict monitoring, and exploitation versus exploration. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 71-90). The Guilford Press.
- Colina, S. (2008). Translation quality evaluation: Empirical evidence for a functionalist approach. *The Translator*, 14(1), 97-134. <https://doi.org/10.1080/13556509.2008.10799251>
- Colina, S. (2009). Further evidence for a functionalist approach to translation quality evaluation. *Target*, 21(2), 235-264. <https://doi.org/10.1075/target.21.2.02col>
- Cooper, C. L., Davies, R., & Tung, R. L. (1982). Interpreting stress: Sources of job stress among conference interpreters. *Multilingua*, 1(2), 97-107. <https://doi.org/10.1515/mult.1982.1.2.97>
- Courtney, J., & Phelan, M. (2019). Translators' experiences of occupational stress and job satisfaction. *Translation & Interpreting*, 11(1), 100-113. <https://doi.org/10.12807/ti.111201.2019.a06>
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87-114. <https://doi.org/10.1017/S0140525X01003922>
- Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, 26(2), 197-223. <https://doi.org/10.1007/s10648-013-9246-y>
- Cox, T. (1978). *Stress*. Macmillan.
- Cox, T. (1985). The nature and measurement of stress. *Ergonomics*, 28(8), 1155-1163. <https://doi.org/10.1080/00140138508963238>
- Cox, T. (1987). Stress, coping and problem solving. *Work & Stress*, 1(1), 5-14. <https://doi.org/10.1080/02678378708258476>
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications.
- Critchley, H. D. (2002). Review: Electrodermal responses: What happens in the brain. *The Neuroscientist*, 8(2), 132-142. <https://doi.org/10.1177/107385840200800209>
- De Dreu, C. K. (2003). Time pressure and closing of the mind in negotiation. *Organizational Behavior and Human Decision Processes*, 91(2), 280-295. [https://doi.org/10.1016/S0749-5978\(03\)00022-0](https://doi.org/10.1016/S0749-5978(03)00022-0)
- De Rooze, B. (2003). *La traducción, contra reloj* [PhD Thesis, University of Granada]. Granada.
- De Rooze, B. (2008). La traducción contra reloj. Consecuencias de la falta de tiempo en el proceso de traducción. In M. Fernández & R. Muñoz (Eds.),

- Aproximaciones cognitivas al estudio de la traducción e interpretación* (pp. 1-37). Comares.
- Deckert, M. (2017). Asymmetry and automaticity in translation. *Translation and Interpreting Studies*, 12(3), 469-488. <https://doi.org/10.1075/tis.12.3.06dec>
- Depaepe, F., & König, J. (2018). General pedagogical knowledge, self-efficacy and instructional practice: Disentangling their relationship in pre-service teacher education. *Teaching and Teacher Education*, 69, 177-190. <https://doi.org/10.1016/j.tate.2017.10.003>
- DiMaggio, P. (1997). Culture and cognition. *Annual Review of Sociology*, 23(1), 263-287. <https://doi.org/10.1146/annurev.soc.23.1.263>
- Dobson, C. B. (1982). *Stress: The hidden adversary*. MTP Press Limited.
- Dragsted, B. (2004). *Segmentation in translation and translation memory systems: An empirical investigation of cognitive segmentation and effects of integrating a TM system into the translation process* [PhD Thesis, Copenhagen Business School]. Copenhagen.
- Dragsted, B. (2005). Segmentation in translation: Differences across levels of expertise and difficulty. *Target*, 17(1), 49-70. <https://doi.org/10.1075/target.17.1.04dra>
- Dragsted, B. (2010). Coordination of reading and writing processes in translation. In G. M. Shreve & E. Angelone (Eds.), *Translation and cognition* (pp. 41–62). John Benjamins.
- Driskell, J. E., Willis, R. P., & Copper, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology*, 77(5), 615-622. <https://doi.org/10.1037/0021-9010.77.5.615>
- Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review* 66(3), 183-201. <https://doi.org/10.1037/h0047707>
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgement and choice. *Annual Review of Psychology* 32(1), 53-88. <https://doi.org/10.2307/2490959>
- Eysenck, M. (1982). *Attention and arousal*. Springer. <https://doi.org/10.1007/978-3-642-68390-9>
- Fitts, P. M. (1966). Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of Experimental Psychology*, 71(6), 849-857. <https://doi.org/10.1037/h0023232>
- Fraisse, P. (1963). *The psychology of time* (J. Leith, Trans.). Harper & Row.
- Frankenhaeuser, M. (1959). *Estimation of time: An experimental study*. Almqvist & Wiksell.
- Frankenhaeuser, M. (1986). A psychobiological framework for research on human stress and coping. In M. H. Appley & R. A. Trumbull (Eds.), *Dynamics of stress* (pp. 101-116). Plenum Press.
- Friend, K. E. (1982). Stress and performance: effects of subjective work load and time

- urgency. *Personnel Psychology*, 35(3), 623-633.
<https://doi.org/10.1111/j.1744-6570.1982.tb02214.x>
- Furlan, S., Agnoli, F., & Reyna, V. F. (2016). Intuition and analytic processes in probabilistic reasoning: The role of time pressure. *Learning and Individual Differences*, 45, 1-10. <https://doi.org/10.1016/j.lindif.2015.11.006>
- Gaillard, A. W. K. (2001). Stress, workload, and fatigue as three biobehavioral states: A general overview. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 623-639). Lawrence Erlbaum Associates Publishers.
- Gaillard, A. W. K., & Wientjes, C. J. E. (1994). Mental load and work stress as two types of energy mobilization. *Work & Stress*, 8(2), 141-152. <https://doi.org/10.1080/02678379408259986>
- Gass, S. M., Behney, J., & Plonsky, L. (2013). *Second language acquisition: An introductory course* (4 ed.). Routledge. <https://doi.org/10.4324/9780203137093>
- Ghobadi, M., Madadi, G., & Najafian, B. (2017). A study of the effects of time pressure on translation quantity and quality. *International Journal of Comparative Literature and Translation Studies*, 5(2), 7-13. <https://doi.org/10.7575/aiac.ijclts.v.5n.2p.7>
- Gile, D. (1995). *Basic concepts and models for interpreter and translator training*. John Benjamins
- Gile, D. (1997). Conference interpreting as a cognitive management problem. In J. H. Danks, G. M. Shreve, S. B. Fountain, & M. McBeath (Eds.), *Cognitive processes in translation and interpreting* (pp. 196-214). Sage Publications.
- Greer, C. R., & Castro, M. A. D. (1986). The relationship between perceived unit effectiveness and occupational stress: The case of purchasing agents. *The Journal of Applied Behavioral Science*, 22(2), 159-175. <https://doi.org/10.1177/002188638602200208>
- Hamilton, P., Hockey, G. R. J., & Rejman, M. (1977). The place of the concept of activation in human information processing theory: An integrative approach. In S. Dornic (Ed.), *Attention and performance VI* (pp. 463-486). Lawrence Erlbaum.
- Han, C. (2020). Translation quality assessment: a critical methodological review. *The Translator*, 26(3), 257-273. <https://doi.org/10.1080/13556509.2020.1834751>
- Hansen, G. (2002a). Selbstaufmerksamkeit im Übersetzungsprozess. In G. Hansen (Ed.), *Empirical translation studies* (pp. 9-28). Samfundslitteratur.
- Hansen, G. (2002b). Zeit und Qualität im Übersetzungsprozess. In G. Hansen (Ed.), *Empirical translation studies* (pp. 29-54). Samfundslitteratur.
- Hansen, G. (2005). *Störquellen in Übersetzungsprozessen: eine empirische Untersuchung von Zusammenhängen zwischen Profilen, Prozessen und Produkten* [PhD Thesis, Copenhagen Business School]. Frederiksberg.
- Hansen, G. (2006a). Retrospection methods in translator training and translation research. *Journal of Specialised Translation*, 5, 2-41.

- Hansen, G. (2006b). Time pressure in translation teaching and translation studies. *Interdisciplinarité en Traduction*, 2, 71-80.
- Hansen, G., & Hömig, H. G. (2000). Kabine oder Bibliothek? Überlegungen zur Entwicklung eines interinstitutionell anwendbaren Versuchsdesigns zur Erforschung der mentalen Prozesse beim Übersetzen. In M. Kadric, K. Kaindl, & F. Pöchhacker (Eds.), *Translationswissenschaft: Festschrift für Mary Snell-Hornby zum 60. Geburtstag* (pp. 319-338). Stauffenburg.
- Harris, J. H., & Berger, P. K. (1983). Antecedents of psychological stress. *Journal of Human Stress*, 9(2), 24-31. <https://doi.org/10.1080/0097840X.1983.9936121>
- Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9), 904-908. <https://doi.org/10.1177/154193120605000909>
- Henckens, M. J., Hermans, E. J., Pu, Z., Joëls, M., & Fernández, G. (2009). Stressed memories: How acute stress affects memory formation in humans. *Journal of Neuroscience*, 29(32), 10111-10119. <https://doi.org/10.1523/JNEUROSCI.1184-09.2009>
- Herborn, K. A., Graves, J. L., Jerem, P., Evans, N. P., Nager, R., McCafferty, D. J., & McKeegan, D. E. F. (2015). Skin temperature reveals the intensity of acute stress. *Physiology and Behavior*, 152, 225-230. <https://doi.org/10.1016/j.physbeh.2015.09.032>
- Hild, A. (2014). The role of self-regulatory processes in the development of interpreting expertise. *Translation and Interpreting Studies*, 9(1), 128-149. <https://doi.org/10.1075/tis.9.1.07hil>
- Hockey, G. R. J. (1979). Stress and the cognitive components of skilled performance. In V. Hamilton & D. M. Warburton (Eds.), *Human stress and cognition* (pp. 141-177). Wiley.
- Hockey, G. R. J. (1986). A state control theory of adaptation and individual differences in stress management. In G. R. J. Hockey, A. W. K. Gaillard, & M. G. H. Coles (Eds.), *Energetics and human information processing* (pp. 285-298). Dordrecht.
- Hockey, G. R. J. (1993). Cognitive-energetical control mechanisms in the management of work demands and psychological health. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent* (pp. 328-345). Clarendon Press.
- Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, 45(1), 73-93. [https://doi.org/10.1016/S0301-0511\(96\)05223-4](https://doi.org/10.1016/S0301-0511(96)05223-4)
- Hockey, G. R. J. (2011). A motivational control theory of cognitive fatigue. In P. L. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 167-187). American Psychological Association. <https://doi.org/10.1037/12343-008>
- Hockey, G. R. J. (2013). *The psychology of fatigue: Work, effort and control*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139015394>
- Horsten, M., Erigson, M., Perski, A., Wamala, S. P., Schenck-Gustafsson, K., & Orth-

- Gomer, K. (1999). Psychosocial factors and heart rate variability in healthy women. *Psychosomatic Medicine*, 61(1), 49-57. <https://doi.org/10.1097/00006842-199901000-00009>
- Hubscher-Davidson, S. (2016). Trait emotional intelligence and translation: A study of professional translators. *Target*, 28(1), 132-157. <https://doi.org/10.1075/target.28.1.06hub>
- Hui, T. K. L., & Sherratt, R. S. (2018). Coverage of emotion recognition for common wearable biosensors. *Biosensors*, 8(2), 30. <https://doi.org/10.3390/bios8020030>
- Hvelplund, K. T. (2009). Indicators of text complexity. In S. Göpferich, A. L. Jakobsen, & I. M. Mees (Eds.), *Behind the mind: Methods, models and results in translation process research* (pp. 61-80). Samfundslitteratur.
- Hvelplund, K. T. (2011). *Allocation of cognitive resources in translation: An eye-tracking and key-logging study* [PhD Thesis, Copenhagen Business School]. Copenhagen.
- Hvelplund, K. T. (2014). Eye tracking and the translation process: Reflections on the analysis and interpretation of eye-tracking data. In R. Muñoz Martín (Ed.), *Minding translation* (pp. 201-223). Universitet d'Alacant.
- Hvelplund, K. T. (2016). Cognitive efficiency in translation. In R. Muñoz Martín (Ed.), *Reembedding translation process research* (pp. 149-170). John Benjamins.
- Hvelplund, K. T. (2017). Translators' use of digital resources during translation. *Hermes* 56, 71-87. <https://doi.org/10.7146/hjlc.v0i56.97205>
- Hvelplund, K. T. (2019). Digital resources in the translation process: Attention, cognitive effort and processing flow. *Perspectives*, 27(4), 510-524. <https://doi.org/10.1080/0907676X.2019.1575883>
- Hygge, S., & Knez, I. (2001). Effects of noise, heat, and indoor lighting on cognitive performance and self-reported affect. *Journal of Environmental Psychology*, 21(3), 291-299. <https://doi.org/10.1006/jevp.2001.0222>
- Immonen, S. (2011). Unravelling the processing units of translation. *Across Languages and Cultures* 12(2), 235-257. <https://doi.org/10.1556/Acr.12.2011.2.6>
- Jääskeläinen, R., Kujamäki, P., & Mäkisalo, J. (2011). Towards professionalism — or against it? Dealing with the changing world in translation research and translator education. *Across Languages and Cultures* 12(2), 143-156. <https://doi.org/10.1556/Acr.12.2011.2.1>
- Jääskeläinen, R., & Tirkkonen-Condit, S. (1991). Automatised processes in professional vs. non-professional translation: A think-aloud protocol study. In S. Tirkkonen-Condit (Ed.), *Empirical research in translation and intercultural studies: Selected papers of the TRANSIF Seminar, Savonlinna 1988* (pp. 89-109). Narr.
- Jakobsen, A. L. (2002). Orientation, segmentation, and revision in translation. In G. Hansen (Ed.), *Empirical translation studies: Process and product* (pp. 191-204). Samfundslitteratur.
- Jakobsen, A. L. (2003). Effects of think aloud on translation speed, revision, and

- segmentation. In F. Alves (Ed.), *Triangulating translation: Perspectives in process oriented research* (pp. 69-95). John Benjamins.
- Jamal, M. (1984). Job stress and job performance controversy: An empirical assessment. *Organizational Behavior and Human Performance*, 33(1), 1-21. [https://doi.org/10.1016/0030-5073\(84\)90009-6](https://doi.org/10.1016/0030-5073(84)90009-6)
- Jamal, M. (1985). Relationship of job stress to job performance: A study of managers and blue-collar workers. *Human Relations*, 38(5), 409-424. <https://doi.org/10.1177/001872678503800502>
- Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. Free Press.
- Jankowiak, K., & Korpala, P. (2018). On modality effects in bilingual emotional language processing: Evidence from galvanic skin response. *Journal of Psycholinguistic Research*, 47(3), 663-677. <https://doi.org/10.1007/s10936-017-9552-5>
- Jennings, J. R. (2007). Heart rate. In G. Fink (Ed.), *Encyclopedia of stress* (2nd ed., pp. 274-277). Elsevier.
- Jensen, A. (1999). Time pressure in translation. In G. Hansen (Ed.), *Probing the process in translation: Methods and results* (pp. 103-119). Samfundslitteratur.
- Jensen, A. (2000). *The effects of time on cognitive processes and strategies in translation* [PhD Thesis, Copenhagen Business School]. Copenhagen.
- Jensen, A., & Jakobsen, A. L. (2000). Translating under time pressure: An empirical investigation of problem-solving activity and translation strategies by non-professional and professional translators. In A. Chesterman, N. G. S. Salvador, & Y. Gambier (Eds.), *Translation in context: Selected contributions from the EST Congress, Granada 1998* (Vol. 39, pp. 105-116). John Benjamins.
- Jiménez-Crespo, M. A. n. (2012). Translation under pressure and the Web: a parallel corpus-study of Obama's Inaugural speech in the online media. *Translation & Interpreting*, 4(1), 56-76.
- Jiménez Ivars, A., & Pinazo Calatayud, D. (2001). I failed because I got very nervous. Anxiety and performance in interpreting trainees: An empirical study. *The Interpreters' Newsletter*, 9, 21-39.
- Julian, L. J. (2011). Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). *Arthritis Care & Research*, 63(S11), S467-S472. <https://doi.org/doi.org/10.1002/acr.20561>
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8(4), 441-480. [https://doi.org/10.1016/0010-0285\(76\)90015-3](https://doi.org/10.1016/0010-0285(76)90015-3)
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review* 87(4), 329-354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23-31. https://doi.org/10.1207/S15326985EP3801_4
- Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, 21(5), 348-354. <https://doi.org/10.1177/0963721412454875>
- Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74(4), 657-690. <https://doi.org/10.1037/0021-9010.74.4.657>
- Kantowitz, B. H. (1987). Mental workload. In P. A. Hancock (Ed.), *Advances in psychology* (Vol. 47, pp. 81-121). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)62307-9](https://doi.org/10.1016/S0166-4115(08)62307-9)
- Kao, P.-C., & Craigie, P. (2013). Evaluating student interpreters' stress and coping strategies. *Social Behavior and Personality*, 41(6), 1035-1043. <https://doi.org/10.2224/sbp.2013.41.6.1035>
- Karasek, R., & Theorell, T. (1990). *Healthy work: Stress, productivity, and the reconstruction of working life*. Basic Books.
- Kataoka, H., Kano, H., Yoshida, H., Saijo, A., Yasuda, M., & Osumi, M. (1998). Development of a skin temperature measuring system for non-contact stress evaluation. *Vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond The 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Hong Kong.
- Klepsch, M., & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science*, 48(1), 45-77. <https://doi.org/10.1007/s11251-020-09502-9>
- Klieme, E., Pauli, C., & Reusser, K. (2009). The pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137-160). Waxmann.
- Klonowicz, T. (1994). Putting one's heart into simultaneous interpretation. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 213-224). John Benjamins.
- Korpal, P. (2016). Interpreting as a stressful activity: Physiological measures of stress in simultaneous interpreting. *Poznan Studies in Contemporary Linguistics*, 52(2), 297-316. <https://doi.org/10.1515/psicl-2016-0011>
- Korpal, P. (2021). Stress experienced by Polish sworn translators and interpreters. *Perspectives*, 29(4), 554-571. <https://doi.org/10.1080/0907676X.2021.1889004>
- Korpal, P., & Jasielska, A. (2019). Investigating interpreters' empathy: Are emotions in simultaneous interpreting contagious? *Target*, 31(1), 2-24. <https://doi.org/10.1075/target.17123.kor>
- Kourouni, K. (2012). *Translating under time constraints in an undergraduate context:*

A study of students' products, processes and learning styles [PhD Thesis, Universitat Rovira i Virgili]. Tarragona.

- Krings, H. P. (2001). *Texte reparieren. Empirische Untersuchungen zum Prozeß der Nachredaktion von Maschinenübersetzungen* [Repairing texts: empirical investigations of machine translation post-editing processes] (G. S. Koby, G. M. Shreve, K. Mischerikow, & S. Litzer, Trans.). Kent State University Press. (1994)
- Künzli, A. (2007). Translation revision: A study of the performance of ten professional translators revising a legal text. In Y. Gambier, M. Shlesinger, & R. Stolze (Eds.), *Doubts and directions in translation studies* (pp. 115-126). John Benjamins.
- Kurz, I. (2002). Physiological stress responses during media and conference interpreting. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st century: Challenges and opportunities* (pp. 195-202). John Benjamins.
- Kurz, I. (2003). Physiological stress during simultaneous interpreting: A comparison of experts and novices. *The Interpreters' Newsletter*, 12, 51-67.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1-26. <https://doi.org/10.18637/jss.v082.i13>
- Lacruz, I. (2017). Cognitive effort in translation, editing, and post-editing. In J. W. Schwieter & A. Ferreira (Eds.), *The handbook of translation and cognition* (pp. 386-401). John Wiley & Sons Inc.
- Lacruz, I., & Shreve, G. M. (2014). Pauses and cognitive effort in post-editing. In S. O'Brien, L. W. Balling, M. Carl, M. Simard, & L. Specia (Eds.), *Post-editing of machine translation: Processes and applications* (pp. 246-272). Cambridge Scholars Publishing.
- Lagace, R. R. (1988). Role-stress differences between salesmen and saleswomen: Effect on job satisfaction and performance. *Psychological Reports*, 62(3), 815-825. <https://doi.org/10.2466/pr0.1988.62.3.815>
- Lazarus, R. S. (1966). *Psychological stress and the coping process*. McGraw-Hill.
- Lazarus, R. S. (1976). *Patterns of adjustment*. McGraw-Hill.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer.
- Levy, J. (1967). Translation as a decision making process. In *To honor Roman Jakobson: Essays on the occasion of his 70. birthday, 11. October 1966* (Vol. 2, pp. 1171-1182). De Gruyter Mouton.
- Lim, Y. M., Ayesha, A., & Stacey, M. (2014, Aug. 27-29). Detecting cognitive stress from keyboard and mouse dynamics during mental arithmetic. 2014 Science and Information Conference, London, UK.
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, 19(6), 527-537. <https://doi.org/10.1016/j.learninstruc.2008.11.001>

- Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading & Writing Quarterly*, *13*(2), 123-146. <https://doi.org/10.1080/1057356970130203>
- Lommel, A. R., Burchardt, A., & Uszkoreit, H. (2013). Multidimensional quality metrics: A flexible system for assessing translation quality. Proceedings of ASLIB: Translating and the Computer, London, UK.
- Lundberg, U. (1982). Psychophysiological aspects of performance and adjustment to stress. In H. W. Krohne & L. Laux (Eds.), *Series in clinical & community psychology: Achievement, stress, & anxiety* (pp. 75-91). Hemisphere.
- Luque-Casado, A., Perales, J. C., Cárdenas, D., & Sanabria, D. (2016). Heart rate variability and cognitive processing: The autonomic response to task demands. *Biological Psychology*, *113*, 83-90. <https://doi.org/10.1016/j.biopsycho.2015.11.013>
- MacGregor, D. (1993). Time pressure and task adaptation: Alternative perspectives on laboratory studies. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 73-82). Plenum Press.
- Mackintosh, J. (2003). The AICC workload study. *Forum*, *1*(2), 189-214. <https://doi.org/10.1075/forum.1.2.09mac>
- Magne, C., Aramaki, M., Astesano, C., Gordon, R. L., Ystad, S., Farner, S., Kronland-Martinet, R., & Besson, M. (2005, 2005). Comparison of rhythmic processing in language and music: An interdisciplinary approach. *Journal of Music and Meaning*, *3*. <https://hal.archives-ouvertes.fr/hal-00105862>
- Maule, A. J., & Hockey, G. R. J. (1993). State, stress, and time pressure. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 83-101). Plenum Press. https://doi.org/10.1007/978-1-4757-6846-6_6
- Maule, A. J., Hockey, G. R. J., & Bdzola, L. (2000). Effects of time-pressure on decision-making under uncertainty: Changes in affective state and information processing strategy. *Acta Psychologica* *104*(3), 283-301. [https://doi.org/10.1016/s0001-6918\(00\)00033-0](https://doi.org/10.1016/s0001-6918(00)00033-0)
- Maule, A. J., & Svenson, O. (1993). Theoretical and empirical approaches to behavioral decision making and their relation to time constraints. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 3-25). Plenum Press.
- McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, *136*(2), 188-197. <https://doi.org/10.1037/a0018298>
- Meglino, B. M. (1977). Stress and performance: Are they always incompatible? *Supervisory Management*, *22*(3), 2-12.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* *63*(2), 81-97. <https://doi.org/10.1037/h0043158>

- Miller, J. G. (1960). Information input overload and psychopathology. *American Journal of Psychiatry*, 116(8), 695-704. <https://doi.org/10.1176/ajp.116.8.695>
- Miller, K. S., & Sullivan, K. P. H. (2006). Keystroke logging: An introduction. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging and writing: Methods and applications* (pp. 1-9). Elsevier. <http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-13155>
- Minkel, J., & Phillips, S. (2015). Methodological considerations when integrating experimental manipulations of sleep and emotion. In K. A. Babson & M. T. Feldner (Eds.), *Sleep and affect* (pp. 119-138). Elsevier. <https://doi.org/10.1016/B978-0-12-417188-6.00006-2>
- Mitchell, D. C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & J. Charles Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 15-32). Psychology Press.
- Moray, N. (Ed.). (1979). *Mental workload*. Springer.
- Moreno, R., & Park, B. (2010). Cognitive load theory: Historical development and relation to other theories. In J. L. Plass, R. Brünken, & R. Moreno (Eds.), *Cognitive load theory* (pp. 9-28). Cambridge University Press. <https://doi.org/10.1017/CBO9780511844744.003>
- Moser-Mercer, B. (2003). *Remote interpreting: Assessment of human factors and performance parameters*. Retrieved 16 Aug 2021 from https://aiic.org/document/516/AIICWebzine_Summer2003_3_MOSER-MERCER_Remote_interpreting_Assessment_of_human_factors_and_performance_parameters_Original.pdf
- Moser-Mercer, B. (2005). Remote interpreting: The crucial role of presence. *Bulletin VALS-ASLA (Swiss Association of Applied Linguistics)* 81, 73-97.
- Moser-Mercer, B., Künzli, A., & Korac, M. (1998). Prolonged turns in interpreting: Effects on quality, physiological and psychological stress (pilot study). *Interpreting*, 3(1), 47-64. <https://doi.org/10.1075/intp.3.1.03mos>
- Muñoz Martín, R. (2009). Expertise and environment in translation. *Mutatis Mutandis: Revista Latinoamericana de Traducción*, 2(1), 24-37.
- Muñoz Martín, R. (2014). Situating translation expertise: A review with a sketch of a construct. In J. W. Schwieter & A. Ferreira (Eds.), *The development of translation competence* (pp. 2-56). Cambridge Scholars Publishing.
- Muñoz Martín, R., & Cardona Guerra, J. M. (2019). Translating in fits and starts: Pause thresholds and roles in the research of translation processes. *Perspectives*, 27(4), 525-551. <https://doi.org/10.1080/0907676X.2018.1531897>
- Muñoz Martín, R., & Martín de Leon, C. (2018). Fascinating rhythm – and pauses in translators' cognitive processes. *Hermes* 57, 29-47. <https://doi.org/10.7146/hjlc.v0i57.106192>
- Muse, L. A., Harris, S. G., & Feild, H. S. (2003). Has the inverted-U theory of stress and job performance had a fair test? *Human Performance* 16(4), 349-364. https://doi.org/10.1207/S15327043HUP1604_2

- Myrtek, M., Weber, D., Brügger, G., & Müller, W. (1996). Occupational stress and strain of female students: Results of physiological, behavioral, and psychological monitoring. *Biological Psychology*, 42(3), 379-391. [https://doi.org/10.1016/0301-0511\(95\)05168-6](https://doi.org/10.1016/0301-0511(95)05168-6)
- Neumann, E., & Blanton, R. (1970). The early history of electrodermal research. *Psychophysiology*, 6(4), 453-475. <https://doi.org/10.1111/j.1469-8986.1970.tb01755.x>
- Norman, R., Mendolicchio, L., Mordeniz, C., Pierangeli, E., Pannarale, P., Orsucci, F., & Conte, E. (2016). Galvanic skin response and its neurological correlates. *Journal of Consciousness Exploration and Research*, 7(7), 553-572.
- Nourbakhsh, N., Chen, F., Wang, Y., & Calvo, R. A. (2017). Detecting users' cognitive load by galvanic skin response with affective interference. *ACM Transactions on Interactive Intelligent Systems*, 7(3), 1-20. <https://doi.org/10.1145/2960413>
- Ornstein, R. E. (1969). *On the experience of time*. Penguin.
- Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429-434. <https://doi.org/10.1037/0022-0663.84.4.429>
- Paas, F. G. W. C., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1-4. https://doi.org/10.1207/S15326985EP3801_1
- Paas, F. G. W. C., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, 32(1), 1-8. <https://doi.org/10.1023/B:TRUC.0000021806.17516.d0>
- Paas, F. G. W. C., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, 24(1), 27-45. <https://doi.org/10.1007/s10648-011-9179-2>
- Paas, F. G. W. C., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63-71. https://doi.org/10.1207/S15326985EP3801_8
- Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6(4), 351-371. <https://doi.org/10.1007/BF02213420>
- Paas, F. G. W. C., Van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79(1), 419-430. <https://doi.org/10.2466/pms.1994.79.1.419>
- PACTE. (2003). Building a translation competence model. In F. Alves (Ed.), *Triangulating translation: Perspectives in process oriented research* (pp. 43-66). John Benjamins.
- Pahlavani, P., & Asroush, M. H. (2013). The relationship between emotional intelligence (EQ) and the EFL learners' oral translation performance. *The*

Iranian EFL Journal, 9(5), 39-50.

- Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, 59(1-2), 185-198. [https://doi.org/10.1016/S1071-5819\(03\)00017-X](https://doi.org/10.1016/S1071-5819(03)00017-X)
- Pedrotti, M., Mirzaei, M. A., Tedescho, A., Chardonnet, J.-R., Merienne, F., Benedetto, S., & Baccino, T. (2014). Automatic stress classification with pupil diameter analysis. *International Journal of Human-Computer Interaction*, 30(3), 220-236. <https://doi.org/10.1080/10447318.2013.848320>
- Pijera-Díaz, H. J., Drachsler, H., Kirschner, P. A., & Järvelä, S. (2018). Profiling sympathetic arousal in a physics course: How active are students? *Journal of Computer Assisted Learning*, 34(4), 397-408. <https://doi.org/10.1111/jcal.12271>
- Poynter, W. D. (1983). Duration judgment and the segmentation of experience. *Memory & Cognition*, 11(1), 77-82. <https://doi.org/10.3758/BF03197664>
- Poynter, W. D. (1989). Inferring time's passage. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life span perspective* (pp. 305-332). North Holland.
- Poynter, W. D., & Homa, D. (1983). Duration judgment and the experience of change. *Perception & Psychophysics*, 33(6), 548-560. <https://doi.org/10.3758/BF03202936>
- Priestly, J. B. (1968). *Man and time*. Dell.
- Pym, A. (2009). Using process studies in translator training: Self-discovery through lousy experiments. In I. M. Mees, F. Alves, & S. Göpferich (Eds.), *Methodology, technology and innovation in translation process research* (pp. 135-156). Samfundslitteratur.
- Qin, S., Hermans, E. J., van Marle, H. J., Luo, J., & Fernández, G. (2009). Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biological Psychiatry* 66(1), 25-32. <https://doi.org/10.1016/j.biopsych.2009.03.006>
- Rasinger, S. M. (2008). *Quantitative research in linguistics: An introduction*. Continuum.
- Rastegary, H., & Landy, F. J. (1993). The interactions among time urgency, uncertainty, and time pressure. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 217-239). Plenum Press.
- Rattat, A.-C., Matha, P., & Cegarra, J. (2018). Time flies faster under time pressure. *Acta Psychologica* 185, 81-86. <https://doi.org/10.1016/j.actpsy.2018.01.014>
- Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, 5(4), 443-448. <https://doi.org/10.3758/BF03197383>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422. <https://doi.org/10.1037/0033-2909.124.3.372>

- Regan, J. E. (1981). Automaticity and learning: Effects of familiarity on naming letters. *Journal of Experimental Psychology*, 7(1), 180-195. <https://doi.org/10.1037/0096-1523.7.1.180>
- Ren, P., Barreto, A., Huang, J., Gao, Y., Ortega, F. R., & Adjouadi, M. (2014). Off-line and on-line stress detection through processing of the pupil diameter signal. *Annals of Biomedical Engineering*, 42(1), 162-176. <https://doi.org/10.1007/s10439-013-0880-9>
- Rendón-Vélez, E., Horváth, I., & Vegte, W. F. V. d. (2016). Motivating subjects to drive in haste using time pressure in a simulated environment. *International Journal of Computer Aided Engineering and Technology*, 8(1-2), 99-124. <https://doi.org/10.1504/IJCAET.2016.073264>
- Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, 32(4), 371-388. <https://doi.org/10.1177/0305735604046096>
- Rojo, A. (2017). The role of emotions. In J. W. Schwieter & A. Ferreira (Eds.), *The handbook of translation and cognition* (pp. 369-385). John Wiley & Sons Inc.
- Rojo, A., Foulquié-Rubio, A.-I., Espín López, L., & Martínez Sánchez, F. (2021). Analysis of speech rhythm and heart rate as indicators of stress on student interpreters. *Perspectives*, 29(4), 591-607. <https://doi.org/10.1080/0907676X.2021.1900305>
- Rojo, A., & Ramos, M. (2014). The impact of translators' ideology on the translation process: A reaction time experiment. In R. Muñoz Martín (Ed.), *Minding translation* (pp. 247-271). Universitet d'Alacant.
- Rojo, A., & Ramos, M. (2016). Can emotion stir translation skill? Defining the impact of positive and negative emotions on translation performance. In R. M. Martín (Ed.), *Reembedding translation process research* (pp. 107-130). John Benjamins.
- Rojo, A., & Ramos, M. (2018). The role of expertise in emotion regulation: Exploring the effect of expertise on translation performance under emotional stir. In I. Lacruz & R. Jääskeläinen (Eds.), *Innovation and expansion in translation process research* (pp. 105-129). John Benjamins. <https://doi.org/10.1075/ata.18.06roj>
- Rojo, A., Ramos, M., & Valenzuela, J. (2014). The emotional impact of translation: A heart rate study. *Journal of Pragmatics*, 71, 31-44. <https://doi.org/10.1016/j.pragma.2014.07.006>
- Rosenqvist, S. (2015). *Developing pause thresholds for keystroke logging analysis* [Bachelor's Thesis, University of Umeå]. Sweden. <http://www.diva-portal.org/smash/get/diva2:834468/FULLTEXT01>
- Rosiers, A., Eyckmans, J., & Bauwens, D. (2011). A story of attitudes and aptitudes? Investigating individual difference variables within the context of interpreting. *Interpreting*, 13(1), 53-69. <https://doi.org/10.1075/intp.13.1.04ros>
- Roziner, I., & Shlesinger, M. (2010). Much ado about something remote: Stress and performance in remote interpreting. *Interpreting*, 12(2), 214-247.

<https://doi.org/10.1075/intp.12.2.05roz>

- Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology* 76(5), 805-819. <https://doi.org/10.1037/0022-3514.76.5.805>
- Russell, L. (2020). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.4.4. In <https://CRAN.R-project.org/package=emmeans>
- Saldanha, G., & O'Brien, S. (2014). *Research methodologies in translation studies*. Routledge.
- Samuels, S. J. (2002). Reading fluency: Its development and assessment. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 166-183). International Reading Association.
- Samuels, S. J. (2006). Toward a model of reading fluency. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 24-46). International Reading Association.
- Samuels, S. J. (2013). Toward a theory of automatic information processing in reading, revisited. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 698-718). International Reading Association.
- Samuels, S. J., & Flor, R. F. (1997). The importance of automaticity for developing expertise in reading. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 13(2), 107-121. <https://doi.org/10.1080/1057356970130202>
- Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, 53(1), 61-97. [https://doi.org/10.1016/0001-6918\(83\)90016-1](https://doi.org/10.1016/0001-6918(83)90016-1)
- Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Advances in Applied Psycholinguistics; Vol 2. Reading, Writing, and Language Learning* (pp. 142-175). Cambridge University Press.
- Schellekens, J. M. H., Sijtsma, G. J., Vegter, E., & Meijman, T. F. (2000). Immediate and delayed after-effects of long lasting mentally demanding work. *Biological Psychology*, 53(1), 37-56. [https://doi.org/10.1016/S0301-0511\(00\)00039-9](https://doi.org/10.1016/S0301-0511(00)00039-9)
- Schilperoord, J. (1996). *It's about time: Temporal aspects of cognitive processes in text production*. Rodopi.
- Schmaltz, M., da Silva, I. A. L., Pagano, A., Alves, F., Leal, A. L. V., Wong, D. F., Chao, L. S., & Quaresma, P. (2016). Cohesive relations in text comprehension and production: An exploratory study comparing translation and post-editing. In M. Carl, S. Bangalore, & M. Schaeffer (Eds.), *New directions in empirical translation process research: Exploring the CRITT TPR-DB* (pp. 239-263). Springer. https://doi.org/10.1007/978-3-319-20358-4_11
- Schönpflug, W. (1983). Coping efficiency and situational demands. In G. R. J. Hockey (Ed.), *Stress and fatigue in human performance* (pp. 299-330). Wiley.
- Seeber, K. G. (2011). Cognitive load in simultaneous interpreting: Existing theories — new models. *Interpreting*, 13(2), 176-204.

<https://doi.org/10.1075/intp.13.2.02see>

- Seeber, K. G. (2013). Cognitive load in simultaneous interpreting: Measures and methods. *Target*, 25(1), 18-32. <https://doi.org/10.1075/target.25.1.03see>
- Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). Oxford University Press.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review* 96(4), 523-568. <https://doi.org/10.1037/0033-295X.96.4.523>
- Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. Freeman.
- Selye, H. (1976). *The stress of life* (Revised ed.). McGraw-Hill.
- Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms [Review]. *Frontiers in Public Health*, 5(258). <https://doi.org/10.3389/fpubh.2017.00258>
- Shallice, T., Stuss, D. T., Alexander, M. P., Picton, T. W., & Derkzen, D. (2008). The multiple dimensions of sustained attention. *Cortex*, 44(7), 794-805. <https://doi.org/10.1016/j.cortex.2007.04.002>
- Shangarffam, N., & Abolsaba, A. (2009). The relationship between emotional intelligence and the quality of translation. *Journal of Teaching English as a Foreign Language and Literature*, 1(2), 103-114.
- Sharmin, S., Špakov, O., Rähkä, K.-J., & Jakobsen, A. L. (2008). Where and for how long do translators look at the screen while translating? In S. Göpferich, A. L. Jakobsen, & I. M. Mees (Eds.), *Looking at eyes: Eye-tracking studies of reading and translation processing* (pp. 31-51). Samfundslitteratur.
- Sherwood, A., & Carels, R. A. (2007). Blood pressure. In G. Fink (Ed.), *Encyclopedia of stress* (2nd ed., pp. 335-342). Elsevier.
- Shreve, G. M. (2006). The deliberate practice: Translation and expertise. *Journal of Translation Studies* 9(1), 27-42.
- Sjørup, A. C. (2013). *Cognitive effort in metaphor translation: An eyetracking and key-logging study* (Publication Number 18.2013) [PhD Thesis, Copenhagen Business School]. Copenhagen. <http://hdl.handle.net/10419/208853>
- Skulmowski, A., & Rey, G. D. (2017). Measuring cognitive load in embodied learning settings. *Frontiers in Psychology*, 8(1191). <https://doi.org/10.3389/fpsyg.2017.01191>
- Slavich, G. M., Taylor, S., & Picard, R. W. (2019). Stress measurement using speech: Recent advancements, validation issues, and ethical and privacy considerations. *Stress*, 22(4), 408-413. <https://doi.org/10.1080/10253890.2019.1584180>
- Sloan, R. P., Shapiro, P. A., Bagiella, E., Boni, S. M., Paik, M., Bigger, J. T., Steinman, R. C., & Gorman, J. M. (1994). Effect of mental stress throughout the day on

- cardiac autonomic control. *Biological Psychology*, 37(2), 89-99. [https://doi.org/10.1016/0301-0511\(94\)90024-8](https://doi.org/10.1016/0301-0511(94)90024-8)
- Spielberger, C. D. (1972). *Anxiety: Current trends in theory and research: I*. Academic Press.
- Staal, M. A. (2004). *Stress, cognition, and human performance: A literature review and conceptual framework*. National Aeronautics & Space Administration.
- Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., Borowiec, A., & Katz, D. I. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia*, 43(3), 396-417. <https://doi.org/10.1016/j.neuropsychologia.2004.06.010>
- Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A multidisciplinary approach to anterior attentional functions. *Annals of the New York Academy of Sciences*, 769(1), 191-212. <https://doi.org/10.1111/j.1749-6632.1995.tb38140.x>
- Sun, S. (2015). Measuring translation difficulty: Theoretical and methodological considerations. *Across Languages and Cultures* 16(1), 29-54. <https://doi.org/10.1556/084.2015.16.1.2>
- Sun, S., & Shreve, G. M. (2014). Measuring translation difficulty: An empirical study. *Target*, 26(1), 98-127. <https://doi.org/10.1075/target.26.1.04sun>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
- Sweller, J. (2003). Evolution of human cognitive architecture. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215-266). Academic Press.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296. <https://doi.org/10.1023/A:1022193728205>
- Taelman, J., Vandeput, S., Spaepen, A., & Van Huffel, S. (2009). Influence of mental stress on heart rate and heart rate variability. In J. Vander Sloten, P. Verdonck, M. Nyssen, & J. Haueisen, 4th European Conference of the International Federation for Medical and Biological Engineering. IFMBE Proceedings, vol 22, Berlin, Heidelberg.
- Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. A., Ranta-aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV – Heart rate variability analysis software. *Computer Methods and Programs in Biomedicine*, 113(1), 210-220. <https://doi.org/10.1016/j.cmpb.2013.07.024>
- Teig, N., Scherer, R., & Nilsen, T. (2019). I know I can, but do I have the time? The role of teachers' self-efficacy and perceived time constraints in implementing cognitive-activation strategies in science. *Frontiers in Psychology*, 10(1697). <https://doi.org/10.3389/fpsyg.2019.01697>
- Tullis, T., & Albert, B. (2013). *Measuring the user experience: Collecting, analyzing, and presenting usability metrics* (2nd ed.). Morgan Kaufmann. <https://doi.org/10.1016/C2011-0-00016-9>

- Turner-Cobb, J. M., & Hawken, T. (2019). Stress and coping assessment. In C. D. Llewellyn, C. McManus, J. Weinman, K. J. Petrie, S. Newman, S. Ayers, & T. A. Revenson (Eds.), *Cambridge handbook of psychology, health and medicine* (3rd ed., pp. 229-236). Cambridge University Press.
- Unsworth, N., & Robison, M. K. (2020). Working memory capacity and sustained attention: A cognitive-energetic perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(1), 77-103. <https://doi.org/10.1037/xlm0000712>
- Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17(2), 147-177. <https://doi.org/10.1007/s10648-005-3951-0>
- Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies. *Medical Education*, 44(1), 85-93. <https://doi.org/10.1111/j.1365-2923.2009.03498.x>
- Varzande, M., & Jadidi, E. (2015). The impact of translators' academic experience on their translation quality. *English Language Teaching*, 8(9), 11-16. <https://doi.org/10.5539/elt.v8n9p11>
- Vieira, L. N. (2015). *Cognitive effort in post-editing of machine translation: Evidence from eye-movements, subjective ratings, and think-aloud protocols* [PhD Thesis, Newcastle University]. Newcastle.
- Vieira, L. N. (2016). How do measures of cognitive effort relate to each other? A multivariate analysis of post-editing process data. *Machine Translation*, 30(1), 41-62. <https://doi.org/10.1007/s10590-016-9188-5>
- Vroom, V. H. (1964). *Work and motivation*. Wiley.
- Waddington, C. (2001). Should translations be assessed holistically or through error analysis? *Hermes* 14(26), 15-37. <https://doi.org/10.7146/hjlc.v14i26.25637>
- Wang, J. T.-Y., Spezio, M., & Camerer, C. F. (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth-telling and deception in sender-receiver games. *American Economic Review*, 100(3), 984-1007. <https://doi.org/10.1257/aer.100.3.984>
- Welford, A. T. (1973). Stress and performance. *Ergonomics*, 16(5), 567-580. <https://doi.org/10.1080/00140137308924547>
- Weng, Y., & Zheng, B. (2020). A multi-methodological approach to studying time-pressure in written translation: Manipulation and measurement. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 19. <https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/view/548>
- Weng, Y., Zheng, B., & Dong, Y. (in press). Time pressure in translation: Psychological and physiological measures. *Target*.
- Wengelin, Å. (2006). Examining pauses in writing: Theory, methods and empirical data. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging and writing: Methods and applications* (pp. 107-130). Elsevier.
- Westbrook, A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach.

- Cognitive, Affective, & Behavioral Neuroscience*, 15(2), 395-415.
<https://doi.org/10.3758/s13415-015-0334-y>
- Westman, M., & Eden, D. (1991). Implicit stress theory: The spurious effects of stress on performance ratings. *Journal of Social Behavior & Personality*, 6(7), 127-140.
- Westman, M., & Eden, D. (1996). The inverted-U relationship between stress and performance: A field study. *Work & Stress*, 10(2), 165-173.
<https://doi.org/10.1080/02678379608256795>
- Whyatt, B., Stachowiak, K., & Kajzer-Wietrzny, M. (2016). Similar and different: Cognitive rhythm and effort in translation and paraphrasing. *Poznan Studies in Contemporary Linguistics*, 52(2), 175-208. <https://doi.org/10.1515/psicl-2016-0007>
- Wilkinson, R. T. (1962). Muscle tension during mental work under sleep deprivation. *Journal of Experimental Psychology*, 64(6), 565-571.
<https://doi.org/10.1037/h0043570>
- Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes*, 37(1), 60-82.
[https://doi.org/10.1016/0749-5978\(86\)90044-0](https://doi.org/10.1016/0749-5978(86)90044-0)
- Yang, X., & Kim, J. H. (2018). Measuring workload in a multitasking environment using fractal dimension of pupil dilation. *International Journal of Human-Computer Interaction*, 35(15), 1352-1361.
<https://doi.org/10.1080/10447318.2018.1525022>
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18(5), 459-482.
- Young, D. L., Goodie, A. S., Hall, D. B., & Wu, E. (2012). Decision making under time pressure, modeled in a prospect theory framework. *Organizational Behavior and Human Decision Processes* 118(2), 179-188.
<https://doi.org/10.1016/j.obhdp.2012.03.005>
- Zakay, D. (1993). The impact of time perception processes on decision making under time stress. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 59-72). Plenum Press.

Appendices

Appendix 1: Source texts

Text A:

(From: *The Economists*. 23/06/2018. “Free exchange: Trading peace for war”)

China might yet hew more closely to rich-country trade rules, and intervene less in its economy and foreign-exchange markets. Trade between America and China could continue to grow, even as the technological gap between them, and their bilateral imbalance, shrink. Expensive goods, investment and services could flow both ways, as between America and Europe. But this sunny future looks increasingly remote. America already limits some Sino-American trade on national-security grounds. Past spats over dumping and other unfair trade practices led to punitive duties on some goods, as allowed under WTO rules. An all-out trade war would blow the two economies apart. The higher tariffs being mooted on half, or nearly all, of America’s imports of Chinese goods would cause serious economic pain in both countries. In America the prices of many goods would jump and those of others, like the soybeans exported to China by the shipload, would plummet. A sudden drop in China’s trade surplus with America, now over 3% of Chinese GDP, would be a heavy blow. Even though a weaker currency would make it easier to export more to other countries, China would probably need both monetary and fiscal stimulus to avoid a socially disruptive rise in unemployment.

Text B:

(From: *The Economists*. 05/05/2018. “Casting illusions aside”)

Tech-nationalists on both sides argue that China and America, their economies intertwined for so long, must now cleave and go their own ways. In China the propaganda doesn’t favour common sense. “Amazing China”, currently smashing box-office records for a documentary, extols Chinese technological prowess. And the press likes to talk of high-speed rail, e-commerce, mobile payments and bike sharing as China’s “new four great inventions” (to rival the past accomplishments of papermaking, printing, gunpowder and the compass). They are not China’s at all. American tech-nationalists also harbour delusions. The Trump administration has flirted with the idea of huge government support for the development of a 5G network. That would never fly politically. Mr Trump insists that America and China will “always be friends, no matter what happens with our dispute on trade”. There is an echo in that of old-think—of a time when American and Chinese officials believed that no matter how much they disagreed, they would always find a way of getting on because the consequences of falling out would be so devastating for both. The two sides’ techno-sparring is evidence of how hard it is becoming to separate their economic and strategic rivalries.

Text C:

(From: *The Economists*. 26/05/2018. “Trade with America: Assessing the pain”)

Leading the charge against China on economic matters has been Robert Lighthizer, the United States Trade Representative. In March, after an investigation into China’s trade practices, he alleged that China had, time and again, stolen American technology or forced firms to hand it over. He called on China to stop subsidising industries that it deems strategic, from renewable energy to electric vehicles. From China’s standpoint, this is a non-starter. Its plan known as “Made in China 2025” identifies ten high-tech industries and sets out global market-share goals. For policymakers in Beijing, it is their blueprint for reaching the next level of development—a reasonable desire for a middle-income country, as 19th-century Americans would have agreed. But foreign governments and businesses see it as a declaration of intent to seek global dominance. The more the rest of the world complains, the more irascible China sounds. Mei Xinyu, a researcher in the commerce ministry, likened America’s demands to what are known in China as the country’s “unequal treaties” with foreign powers in pre-communist days. The most notorious of these accords was forced on China in 1842 by Britain after a war over British opium sales. It required China to open its doors to foreign trade and cede Hong Kong.

Appendix 2: Glossary list

Text A

1	hew	vi. 砍; 劈; 坚持
2	spat	n. 口角; 蚝卵; 蚝仔; 掌击
3	dumping	n. 倾销; 倾泻
4	tariff	n. 关税表; 收费表
5	moot	vt. 提出...供讨论
6	plummet	vi. 垂直落下; (价格、水平等) 骤然下跌
7	shrink	vi. 收缩; vt. 使缩小, 使收缩

Text B

1	cleave	vt. 砍开; 使分开; 打通
2	propaganda	n. 宣传; 传道总会
3	smash	vt. 粉碎; 使破产; 溃裂
4	extol	vt. 颂扬; 赞美; 赞颂
5	prowess	n. 英勇; 超凡技术; 勇猛
6	delusion	n. 迷惑, 欺骗; 错觉; 幻想
7	spar	vi. 争论; 拳击

Text C

1	allege	vt. 宣称, 断言; 提出...作为理由
2	subsidise	vt. 资助 (等于 subsidize); 给...补助金
3	non-starter	无希望成功的人或事; 不值得考虑的人或方案
4	irascible	adj. 易怒的
5	liken	vt. 比拟; 把...比作
6	accord	n. 符合; 一致; 协议; 自愿
7	cede	vt. 放弃; 割让 (领土)

Appendix 3: Translation guidelines/brief**Task:**

You are asked to translate the following texts from English into Chinese within the given deadlines.

Your commissioner:

A Chinese newspaper.

Your status:

You are working as a translator for them.

Target readership:

The newspaper is addressed to the educated general public. The text you deliver will be included in the future edition.

Source:

The text is taken from *The Economist*.

Situation:

The task is an urgent one. You should complete the translation task within the given deadline. Please give priority to the completeness since your performance will be evaluated based on the completeness of the translation product and an incomplete product will be marked down significantly. You should make full use of the time available and do not quit before the deadline.

Deadline:

16:15/ 20:25/ no time limit

Appendix 4: Instructions before task**Time condition: Short**

Please note that the time given for the following task is 16 min 15 sec and this timeframe is shorter than the average time generally needed. Please try your best to complete the translation task and if possible, prioritise completeness. Your translation will be sent to an expert for marking and you will receive feedback for your work later. You are not allowed to use any online or offline tools for consultation during the task, neither can you ask any questions related to the text during the translation task. But you can check out the word list now using any tools. If you complete the task before the deadline, please take the time to review or revise your work. When the time is up, you will hear a beeping sound reminder and the researcher will end the task for you.

Time condition: Standard

Please note that the time given for the following task is 20 min 25 sec and this timeframe is the average time generally needed. Please try your best to complete the translation task and prioritise completeness. Your translation will be sent to an expert for marking and you will receive feedback for your work later. You are not allowed to use any online or offline tools for consultation during the task, neither can you ask any questions related to the text during the translation task. But you can check out the word list now using any tools. If you complete the task before the deadline, please take the time to review or revise your work. When the time is up, you will hear a beeping sound reminder and the researcher will end the task for you.

Time condition: Free

Please note that there is no time limit for the following task. Please do your best to complete the translation task. Your translation will be sent to an expert for marking and you will receive feedback for your work later. You are not allowed to use any online or offline tools for consultation during the task, neither can you ask any questions related to the text during the translation task. But you can check out the word list now using any tools. When you complete the task, please inform the researcher to end the task for you.

Appendix 5: Consent form**CONSENT TO PARTICIPATE IN RESEARCH****Identification of Investigator & Purpose of Study**

You are invited to participate in a research study conducted by Yu Weng, a research postgraduate from School of Modern Languages and Cultures at Durham University. This study will contribute to the researcher's completion of her Ph.D project, which is about the influence of time pressure on translation process and product.

Research Procedures

You are asked to translate 3 short texts from English into Chinese and your data of eye movement, keystroke logging, pupil size, blood pressure, heart rate, skin temperature, galvanic skin response and heart rate variability will be recorded or measured. You will also be asked to complete a set of questionnaires (State-Trait Anxiety Inventory, NASA Task Load Index and a basic background information sheet). You will be given a £10 Tesco Voucher after completion of the whole experiment.

Privacy and Confidentiality

The result of this research will be coded in a way in which respondents' identity will not be attached to the final presentation of the study. The researcher retains the right to use and publish non-identifiable data. While individual responses are confidential, the overall result and data will be presented representing averages or generalisations of each group of participants as a whole. All the data will be stored in a secure place and only accessible to the researcher.

Participation and Withdrawal

Your participation is completely voluntary. If you choose to participate, you can withdraw at any time without consequences of any kind.

Right as Research Subjects

You are not waiving any legal claims, rights or remedies because of your participation in this study.

Questions about the study

If you have questions or concerns about the study, please contact:

Researcher's name: Yu Weng

Department: School of Modern Languages and Cultures, Durham University

Email: yu.weng@durham.ac.uk

Giving of Consent

I have read this consent form and I understand what is being requested of me as a participant of the study. I freely consent to participate. I also give the researcher my consent for the use of my data for any anticipated future research.

Name of participant: _____ (Signed) Date:

Name of researcher: _____ (Signed) Date:

Appendix 6: Background information questionnaire**Background Information**

Contact Information:

Name: _____ Email: _____ Date: _____

Please answer the following questions to the best of your knowledge.

1. Age (in years):
2. Sex (circle one): Male / Female
3. Education (highest degree obtained and your major):
4. When did you start to learn English as a second language?
5. For how many years have you been working as a professional freelance/inhouse translator (with stable income from the translation work)? Please specify the kind of company you work for and list your specialisation of translation (e.g., legal, medical, business & finance etc.).
6. Are you experiencing hypertension (high blood pressure) condition? If yes, please indicate your average blood pressure (systolic/ diastolic) value here.
7. If you have taken a standardized test of proficiency for English (e.g., TOEFL, IELTS, TEM8 etc.), please indicate the scores you received for each.

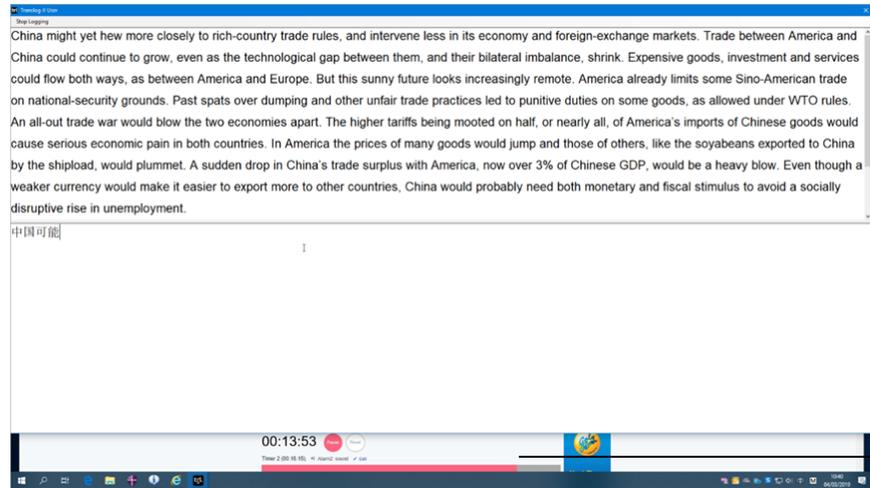
Test	Scores				
	Total	Listening	Speaking	Reading	Writing
IELTS					
TOEFL					
TEM8					

Appendix 7: Task rotation table

Order	1		2		3	
P01	Text A	Short	Text B	Standard	Text C	Free
P02	Text A	Standard	Text B	Free	Text C	Short
P03	Text A	Free	Text B	Short	Text C	Standard
P04	Text B	Short	Text C	Standard	Text A	Free
P05	Text B	Standard	Text C	Free	Text A	Short
P06	Text B	Free	Text C	Short	Text A	Standard
P07	Text C	Short	Text A	Standard	Text B	Free
P08	Text C	Standard	Text A	Free	Text B	Short
P09	Text C	Free	Text A	Short	Text B	Standard

Appendix 8: Task interface

Short and Standard task:



China might yet hew more closely to rich-country trade rules, and intervene less in its economy and foreign-exchange markets. Trade between America and China could continue to grow, even as the technological gap between them, and their bilateral imbalance, shrink. Expensive goods, investment and services could flow both ways, as between America and Europe. But this sunny future looks increasingly remote. America already limits some Sino-American trade on national-security grounds. Past spats over dumping and other unfair trade practices led to punitive duties on some goods, as allowed under WTO rules. An all-out trade war would blow the two economies apart. The higher tariffs being mooted on half, or nearly all, of America's imports of Chinese goods would cause serious economic pain in both countries. In America the prices of many goods would jump and those of others, like the soybeans exported to China by the shipload, would plummet. A sudden drop in China's trade surplus with America, now over 3% of Chinese GDP, would be a heavy blow. Even though a weaker currency would make it easier to export more to other countries, China would probably need both monetary and fiscal stimulus to avoid a socially disruptive rise in unemployment.

中国可能

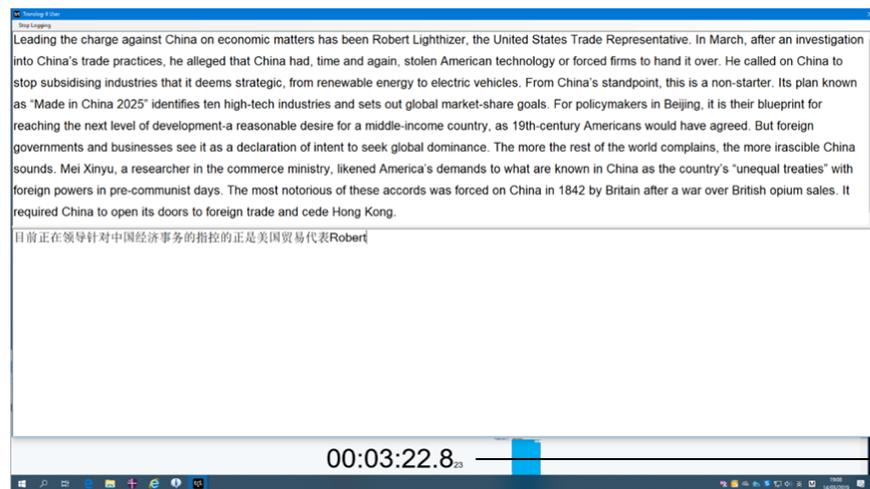
00:13:53

ST

TT

Countdown timer

Free task:



Leading the charge against China on economic matters has been Robert Lighthizer, the United States Trade Representative. In March, after an investigation into China's trade practices, he alleged that China had, time and again, stolen American technology or forced firms to hand it over. He called on China to stop subsidising industries that it deems strategic, from renewable energy to electric vehicles. From China's standpoint, this is a non-starter. Its plan known as "Made in China 2025" identifies ten high-tech industries and sets out global market-share goals. For policymakers in Beijing, it is their blueprint for reaching the next level of development—a reasonable desire for a middle-income country, as 19th-century Americans would have agreed. But foreign governments and businesses see it as a declaration of intent to seek global dominance. The more the rest of the world complains, the more irascible China sounds. Mei Xinyu, a researcher in the commerce ministry, likened America's demands to what are known in China as the country's "unequal treaties" with foreign powers in pre-communist days. The most notorious of these accords was forced on China in 1842 by Britain after a war over British opium sales. It required China to open its doors to foreign trade and cede Hong Kong.

目前正在领导针对中国经济事务的指控的正是美国贸易代表Robert

00:03:22.8

ST

TT

Stopwatch

Appendix 10: GSR and HRV data completeness in Baseline 1 and Baseline 2

Group=Novice ID	GSR			HRV		
	Baseline1	Baseline2	Status	Baseline1	Baseline2	Status
NP01	√	√	remain	√	√	remain
NP02	√	√	remain	√	√	remain
NP03	√	√	remain	√	√	remain
NP04	√	√	remain	√	√	remain
NP05	√	√	remain	√	√	remain
NP06	√	√	remain	√	√	remain
NP07	√	√	remain	√	√	remain
NP08	√	√	remain	√	√	remain
NP09	√	√	remain	√	√	remain
NP10	√	√	remain	√	√	remain
NP11	√	√	remain	√	×	remove
NP12	√	√	remain	√	√	remain
NP13	√	√	remain	√	√	remain
NP14	√	√	remain	√	×	remove
NP15	√	√	remain	√	×	remove
NP16	√	√	remain	√	√	remain
NP17	√	√	remain	√	×	remove
NP18	√	√	remain	√	√	remain
NP19	√	√	remain	√	√	remain
NP20	√	√	remain	√	√	remain
NP21	√	√	remain	√	√	remain
NP22	√	√	remain	×	√	remove
NP23	√	√	remain	√	√	remain
NP24	√	√	remain	√	√	remain
NP25	√	√	remain	√	√	remain
NP26	√	√	remain	√	√	remain
NP27	√	√	remain	√	√	remain
NP28	√	√	remain	√	√	remain
NP29	√	√	remain	√	√	remain
NP30	√	√	remain	√	√	remain
NP31	√	√	remain	√	√	remain
NP32	√	√	remain	√	×	remove
NP33	√	√	remain	√	√	remain
NP34	√	√	remain	√	√	remain
NP35	√	√	remain	√	√	remain

Group=Experienced ID	GSR			HRV		
	Baseline1	Baseline2	Status	Baseline1	Baseline2	Status
EP01	√	√	remain	√	√	remain
EP02	√	√	remain	√	√	remain
EP03	√	√	remain	√	√	remain
EP04	√	√	remain	√	√	remain
EP05	√	√	remain	√	√	remain
EP06	√	√	remain	√	√	remain
EP07	√	√	remain	√	√	remain
EP08	√	√	remain	√	√	remain
EP09	√	√	remain	√	√	remain
EP10	√	√	remain	√	√	remain
EP11	×	√	remove	√	√	remain
EP12	√	√	remain	√	√	remain
EP13	√	√	remain	√	√	remain
EP14	√	√	remain	√	√	remain
EP15	√	√	remain	√	×	remove
EP16	√	√	remain	√	×	remove
EP17	√	√	remain	√	×	remove
EP18	√	√	remain	×	√	remove
EP19	√	√	remain	×	×	remove
EP20	√	√	remain	√	√	remain
EP21	√	√	remain	√	×	remove
EP22	√	√	remain	√	×	remove
EP23	√	√	remain	√	√	remain
EP24	√	√	remain	×	×	remove
EP25	√	√	remain	√	√	remain
EP26	√	√	remain	√	√	remain
EP27	√	√	remain	×	×	remove
EP28	√	√	remain	√	√	remain
EP29	√	√	remain	√	×	remove
EP30	√	√	remain	√	√	remain

Appendix 11: Adapted NASA TLX questionnaire

Mental Demand: How much mental and perceptual activity did you spend for this task?

_____ |
Low High

Definition: How much mental and perceptual activity was required (for example, thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, forgiving or exacting?

Temporal Demand: How much time pressure did you feel in order to complete this task?

_____ |
Low High

Definition: How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Effort: How hard did you have to work to accomplish your level of performance?

_____ |
Low High

Definition: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Stress: How stressed were you during this task?

_____ |
Low High

Definition: How stressed versus relaxed did you feel during the task?

Appendix 12: Biodata of participants

Group=Novice				Group=Experienced			
ID	Age	Gender	IELTS	ID	Age	Gender	Experience (years)
NP01	29	F	7	EP01	25	F	4
NP02	29	M	8	EP02	28	F	4
NP03	25	F	8	EP03	37	M	15
NP04	34	F	7	EP04	27	F	7
NP05	23	F	7.5	EP05	29	M	5
NP06	26	F	7	EP06	28	M	5
NP07	28	F	7	EP07	50	M	23
NP08	22	F	7.5	EP08	24	F	4
NP09	24	F	7.5	EP09	24	F	3
NP10	24	F	7.5	EP10	24	F	3
NP11	25	M	7.5	EP11	24	F	4
NP12	23	F	7.5	EP12	25	F	4
NP13	23	F	7.5	EP13	24	F	4
NP14	24	F	7.5	EP14	28	F	5
NP15	23	F	7.5	EP15	27	F	5
NP16	23	F	7	EP16	28	F	4
NP17	22	F	8	EP17	43	M	18
NP18	23	M	7	EP18	27	M	3
NP19	22	F	7.5	EP19	40	M	17
NP20	22	F	7.5	EP20	29	F	5
NP21	22	F	7.5	EP21	24	F	3
NP22	23	F	7.5	EP22	24	F	3
NP23	23	F	7	EP23	24	F	4
NP24	21	F	7	EP24	30	M	9
NP25	22	F	7	EP25	47	F	15
NP26	23	F	7	EP26	46	F	16
NP27	24	F	7.5	EP27	30	M	6
NP28	22	F	7	EP28	36	F	8
NP29	23	F	7.5	EP29	30	F	5
NP30	24	M	8	EP30	40	F	6
NP31	29	F	7				
NP32	24	M	7				
NP33	24	F	7.5				
NP34	23	F	7.5				
NP35	24	F	7.5				