

Durham E-Theses

Cognitive and behavioural outcomes of non-organic failure to thrive

Sally Suzanne Corbett

How to cite:

Corbett, Sally Suzanne (1994) Cognitive and behavioural outcomes of non-organic failure to thrive. Masters 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/5835/> 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.

One correction has been made to the original thesis; paragraph 2 on page 5 has been inserted into the text, and the additional reference has been added to page 133.

COGNITIVE AND BEHAVIOURAL OUTCOMES OF
NON-ORGANIC FAILURE TO THRIVE

SALLY SUZANNE CORBETT

Thesis submitted for the degree of Master of Arts

University of Durham
Department of Psychology

1994

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



14 FEB 1995

Statement of Copyright

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

ACKNOWLEDGMENTS

I would like to express thanks to Dr. Robert Drewett for his challenging approach, patience and support during this project.

My thanks are also due to Dr. Charlotte Wright for her untiring efforts to trace subjects and her enormous energy and enthusiasm in co-ordinating our efforts.

I would like to thank Alice Thams who enabled all testing to be carried out 'blind'.

Finally, I would like to thank Andrew Moon and Malcolm Bass for providing training in the use of the WIPPSI-R.

CONTENTS

ABSTRACT.....	4
<u>CHAPTER 1</u>	
1.1. INTRODUCTION.....	5
1.2. CRITERIA FOR FAILURE TO THRIVE.....	7-17
1.2.1. Current criteria for diagnosis.....	7
1.2.2. Diagnosing failure to thrive using normative data....	8-16
1.2.3. Practical difficulties in the collection of data....	16-17
1.2.4. Summary.....	17
1.3 EXPLANATIONS OF THE CAUSE OF FAILURE TO THRIVE.....	18-25
1.3.1. Explanations in terms of emotional deprivation.....	18-19
1.3.2. Explanations in terms of abuse and neglect.....	20
1.3.3. Explanations in terms of nutritional deficiency....	21-22
1.3.4. A synergistic approach.....	23-24
1.3.5. Summary.....	25
1.4. REVIEW OF OUTCOME STUDIES.....	26-44
1.4.1. Studies based on referred cases.....	26-33
1.4.2. Population based or primary care studies.....	33-36
1.4.3. Methodological problems.....	36-43
1.4.4. Summary.....	44
1.5. FACTORS AFFECTING IQ AND BEHAVIOUR.....	45-53
1.5.1. The direct effect of undernutrition on the brain....	45-48
1.5.2. Indirect effects of undernutrition on the brain....	48-49
1.5.3. Effect of the context of undernutrition.....	49-52
1.5.4. Summary.....	53
1.6. OUTCOME EVALUATION.....	54-56
1.6.1. IQ testing.....	54-55
1.6.2. Social competence and behaviour problems.....	55-56
1.6.3. Summary.....	56

CONTENTS CONTINUED.

CHAPTER 2.

2.1. AIMS AND OBJECTIVES OF THE STUDY.....57-59

2.2. METHOD.....60-82

2.2.1. Overview of the study.....60-61

2.2.2. Subjects.....62-68

2.2.3. Comparison of demographic characteristics of those lost
to follow up.....69

2.2.4. Materials.....70-77

2.2.5. Procedure.....77-81

2.2.6. Outcome measures.....82

2.2.7. Growth measures.....82-85

CHAPTER 3.

3.3. ANALYSIS OF DATA AND RESULTS.....86-101

3.3.1. Between groups comparison.....86-95

3.3.2. Further analysis within the case group.....96-99

3.3.3. Summary of Results.....100-101

CHAPTER 4

4.1 DISCUSSION.....102-118

4.1.1 The screening criterion.....104-107

4.1.2. The primary care setting.....107-108

4.1.3. The selection of a control group.....109-110

4.1.4. Exclusions and attrition.....110-113

4.1.5. Tester awareness.....113

4.1.6. Outcome measures.....114-117

4.2. CONCLUSION.....118

APPENDICES.....119-130

REFERENCES.....131-137

LIST OF TABLES.

Table 1. Risk variables shown to be significant predictors of IQ.....51

Table 2. Comparison of unemployment, car and home ownership for the overall area, cases and controls.....65

Table 3. Rate of attrition.....68

Table 4. Comparison of height Z scores at school entry.....88

Table 5. Weschler Pre-School and Primary Scales of Intelligence - Revised: A comparison of case and control groups.....91

Table 6. Teacher's report form: comparison of case and control groups.....94

Table 7. Regression of IQ on the Thrive Index at lowest centile point.....99

Table 8. Thrive Index of those cases lost to follow-up.....112

LIST OF FIGURES

Figure 1. Stem and Leaf of the thrive index at the lowest centile point.....85

Figure 2. Weschler Pre-School and Primary Scales of Intelligence - Revised: A comparison of Full Scale Scores of case and control groups.....92

Figure 3. Teacher's Report Form: A comparison of case and control groups.....95

ABSTRACT

COGNITIVE AND BEHAVIOURAL OUTCOMES OF FAILURE TO THRIVE

In a study of failure to thrive in 1987-8, 54 children (21%) from an annual cohort attending two clinics in a deprived area of Newcastle Upon Tyne were identified as having fallen across two or more major weight centiles for a month or more during the first 18 months of life. They were studied with 52 normally growing controls selected from the same clinics. Eighty nine per cent of these children were traced for a follow up study at age 6-7 years old. This follow up study is reported in this thesis.

IQ was assessed using the Weschler Preschool and Primary Scale of Intelligence. The Teacher's Report Form (Achenbach, 1991) was used to assess behaviour problems in the groups. Testers were unaware of the clinical status of the children. Height was routinely measured at school entry and the original data were analysed to determine age at the lowest centile point and severity of fall in weight gain.

In an independent samples analysis, a small but statistically significant difference in height at school entry age was found, but there was no statistically significant difference between the cases and controls in IQ (mean IQ 83.6 and 87 respectively, $P=0.16$), or ratings of behaviour problems (TRF median problems reported 23 and 14, Mann $U=1.05$, $P=0.297$). Teacher ratings did not reach conventional levels of statistical significance in any subsequent analysis.

A within case group analysis of growth data was carried out to determine if there was a larger effect on a subset of cases sharing characteristics of growth failure. The effects of chronicity, age at the lowest centile point and severity of failure to thrive were analysed. A significant association was found between IQ and severity of failure to thrive ($P=.03$).

Analysis of weight gains showed that while the screening criterion used was sensitive, identifying a group of children with a median rate of weight gain below the 10th centile for expected weight gain, 6 had fallen no lower than their expected weight gain and 17 were only mildly growth retarded. However, the measure for severity of fall used in this study is not only a sensitive criterion, but can also distinguish between a normal fall in rate of weight gain towards the population mean and an abnormal fall away from the mean and it was this measure that was significantly associated with IQ.

CHAPTER 1

1.1. INTRODUCTION

The aim of this project was to determine the long term effects on intellectual and behavioural development of non-organic failure to thrive (poor growth in infancy with no known organic cause).

Many infants in the UK are weighed at intervals after birth, but it is difficult at the moment to provide clear evidence for the functional importance of this screening programme (Hall,1991). One possibility is that weighing provides a means of detecting early malnutrition; this would be important if early malnutrition lead to later cognitive (or other) impairments in children who already suffered from other sources of educational disadvantage. It is, therefore, important to try to provide further evidence on this issue, and especially to investigate the outcomes of failure to thrive defined by different criteria (for example, velocity based criteria rather than criteria based solely on attained weight.

The screening measure used to identify the 'cases' in this study was based on an explicitly described diagnostic criterion (Edwards et al, 1990), which focusses on a single attribute of failure to thrive, a relatively low growth velocity. Our aim was to ascertain if this criterion identified a pattern of growth which was

associated with poor intellectual and behavioural outcomes.

The cohort selected was first studied by me in their homes, in 1991, for my third year undergraduate dissertation (Corbett, 1992). The study compared parental behaviour ratings of cases and controls using the Achenbach Child Behaviour Checklist (1991). Both data sets for the current study, a teacher's behaviour rating and IQ, were to be obtained in the children's schools.

Previous contact with the families had lead me to conclude that some of the children I had visited were 'cases'. Since an important requirement of this study design was that the tester should not know the status of the children tested, a second psychologist was recruited to carry out testing of those whose clinical status was thought to be known. Clearly, if these had been the only children she tested she would have realised that they were 'cases' too. Instead, by supervising her work, I was able to allocate just under half the children tested (46), including those whose status was suspected by me, to her caseload so maintaining our 'blind' status.

1.2. CRITERIA FOR FAILURE TO THRIVE

1.2.1. CURRENT CRITERIA FOR DIAGNOSIS

Although it would seem simple enough to determine whether a child is growing normally or not, no clear agreement about how to define normal growth has as yet emerged. Wilcox et al (1989) reviewed 22 current paediatric textbooks and 13 journals. Of the textbooks 4 made no mention of failure to thrive, 2 gave no definition at all and 6 used subjective definitions which did not include any anthropometry or criteria for abnormality. Failure to thrive was described in quantitative terms in 9/22 textbooks, but there was a lack of consensus in the anthropometric indices used and the criteria for abnormality. The indices used were weight for height, or weight for age, or attained growth or growth velocity. In 8 texts criteria for abnormality were indices below the third centile. However, this cut off point is arbitrary and indices below the tenth and fifth centile have been used as well. The diagnostic criteria are just as confusing for the 13 journals reviewed. Wilcox et al argue that absence of a standard definition has produced an ambiguous body of literature.

1.2.2. DIAGNOSING FAILURE TO THRIVE USING NORMATIVE DATA

In order to clarify the effect different indices for growth have on outcomes these should be examined to determine which populations would be included by each. Firstly, we need to consider two broad categories of growth measure, distance and velocity measures. Distance measures are used to screen for growth problems in the primary care setting. Velocity measures are used more to determine response to treatment.

A distance measure can be defined as an attained height or weight by a certain age, analogous to a child having travelled a certain distance towards adult size (Tanner et al, 1966). Estimates of normality are made by comparing an individual's anthropometry with population norms. Height or weight of a population at any given age is assumed to have a Gaussian distribution. Growth charts have the 97th, 90th, 75th, 50th, 25th, 10th and 3rd centile lines marked to facilitate comparison of an individual's standing with population growth. The height, head circumference and weight norms (and many other anthropometric measures less commonly used) for a population are calculated mainly from cross sectional measures of samples of the target population. This discussion will centre mainly on the norms used for the British population, the British Standards of Tanner and Whitehouse (1976).

There are three main problems with identifying failure to thrive by selecting all individuals below a particular centile on a standardised growth chart. The first problem relates to what constitutes abnormal growth. The second problem is that use of standardised growth charts increases the effect of confounding variables which affect measures of psychological outcomes, and the last problem is that the standards used may not be based on representative populations or adequately updated to allow for changes in childcare practice.

Firstly, the issue of abnormality. In any Gaussian distribution there must be a group of people who occupy the tails of the distribution. This is an indication of normality in a population. It is therefore a contradiction in terms to assume that the smallest 10 or 3 per cent are abnormal and any cut off point must be arbitrary. There is a need to distinguish between small normal and small abnormal individuals. In order to allow for genetically small normal growth a centile line is calculated for mid-parental height. A normal small child's height centile should show a strong correlation with mid-parental height centile. But this definition of normality assumes parental growth has been normal. Also correlations for infants height centile and mid parental height centile before the age of 2 are poor as early length is more strongly influenced by intrauterine experience (Tanner, 1989).

As for weight, in order to ascertain if weight is abnormal, a weight for height ratio can be calculated. Low weight for height is suggestive of malnutrition. It is often thought that a small child with a normal weight for height ratio shows no growth pathology. However, as Tanner (1989) points out, short stature can be an adaptive mechanism to counteract the effects of chronic malnutrition. An average 1 year old infant requires 330 kilojoules per kilogramme of body weight per day for normal body maintenance. The estimated energy required for normal growth is 20 kJ/kg/per day, and the estimated energy required for normal activity is 80 kJ/kg/per day (Waterlow, 1975). In the event that the daily intake falls below this requirement growth and activity are reduced. Since height is restored more slowly to its predicted level than weight and so is less responsive to nutritional change than weight, growth retarded children regulate at a normal weight for height, but neither weight nor height are normal for growth predicted from measures taken before the nutritional deficit.

Secondly, how does an objective growth measure introduce confounding variables into a sample of short people? Height and weight are not constant genetically determined measures impervious to environmental effects. Tanner (1989) discusses statistical evidence to show that attained height is lower in children from low social class and with a greater number of

children in a family. At age 7 these two factors can make a difference in mean heights of 5.5 cms. Therefore, the sample of individuals clustered in the bottom 3rd centile can be expected to have a higher representation of poor children from large families. In the same way babies born to parents of low social class are more likely to be low birth weight babies (Rutter and Quine, 1990). These infants are more likely to go on to fail to thrive than normal weight babies (Frank and Zeisel, 1985). These population characteristics also have a confounding effect on psychological outcomes.

Finally, Tanner et al (1966) argue that there is a need to update normative data on growth at least every 10-15 years as there is a secular trend for children to be larger and to mature sooner. The British Standards of Tanner and Whitehouse (1976) are based on a population of 80 boys and 80 girls followed longitudinally from birth to 5.5 years in the 1950s. Further studies carried out in Oxford in the 1950s closely matched the London data and Tanner (1989) also takes the view that recent surveys show that these standards represent the present situation well.

But the standards were devised from a very small localised sample population, at a time (the 1950s) when nutritional practice tended to produce larger infants. For example the original sample of children were all from the London area and were predominantly

bottle fed. There have been important changes in infant feeding practice during the last 40 years. Infant formulae are less rich in comparison with those of the fifties and following cases of obesity caused by incorrect use of formula feeds in the sixties, mothers are exhorted to make up the milk according to instructions without adding extra milk powder. Also an increase in the prevalence of breast feeding in recent years has lead to the separate study of anthropometry in breast fed babies. Dewey et al (1992) found that breast fed babies tend to be lighter after the first 3 months than bottle fed, but with no significant difference in length. They argue that growth of breast fed babies is often thought to be faltering at 3 months, whereas the problem is that growth charts for breast fed babies need to be revised. In addition, Wright et al (1992) have shown in a comparison of poor Newcastle children and better off Oxford children, that there are significant differences in attained height which may be attributed to genetic differences in localised populations, or different nutritional cultures, or economic disadvantage and that the current growth charts fail to take into account the resulting regional variations and should be revised.

It would be unusual to make assessments of an individual's growth status on the basis of a single height, weight or head circumference alone. It would be important to know, if a child were found to be on the 3rd centile for weight, whether the weight gain was constant, increasing or decreasing. It could also be argued that a child whose weight was on the 25th centile was failing to thrive if previous measures were on the 75th centile. Therefore, a second method of assessing growth is to compare rate of weight gain with standardised norms, in other words to use a velocity measure.

There are two ways of assessing growth velocity. The first way involves the use of growth velocity charts. The second way is to monitor changes in centile position on distance charts.

Velocity charts for height and weight are rarely used in primary care clinics, for although they provide data on growth velocity in a population, this is derived cross sectionally. Longitudinal data would produce a quite different picture. After all if a child were to maintain weight gain on the 97th centile it would become pathologically large and vice versa for a child remaining on the 3rd centile (although it would be possible to remain on the 50th centile). As Tanner et al (1966) noted the correlation between height gain from one year to the next is much lower ($r=0.3$) than for height distance ($r=0.9$). Velocity is a more

labile measure than distance. However, Tanner (1989) supports its use in clinical work as 'velocity picks out the pathological cases better than does distance because velocity represents what is happening now, whereas distance represents the sum of all that has happened in the past'. These charts can be used to monitor catch-up growth.

Catch up growth is the result, Waddington (1957) argued, of the process by which any organism will return to its genetically ordained path once the environmental events which have conspired to push it off its original course have been removed. The genetically determined developmental path is known as the Chreod and the tendency to return to the chreod is referred to as canalization. The mechanisms by which canalization occurs are disputed, whether it is a target seeking mechanism or a discrepancy detector.

Casey and Arnold (1985) described catch up growth in 10 infants with severe failure to thrive over a period of 9 months. They found that the degree to which catch up occurs in weight, length and head circumference varies in the same way that the effect of growth retardation on these different anthropometric measures varies. Original weight for age was 54.9% of normal recovering to 95.5%, height for age was 86% of normal recovering to 94%, and head circumference was 92% recovering to 98% of that expected for age. Head circumference appears to be the

least affected and the most likely to show the least deficit after recovery. This is important in the measurement of psychological outcomes as head circumference is an indicator of brain growth.

Studies such as Dowdney et al (1987) used catch up growth to confirm their diagnosis of failure to thrive in the 25 subjects finally selected and Drotar et al (1985) included evidence of catch up growth as one of their criteria for selection of subjects. The main difficulty with using catch-up growth as a criterion for failure to thrive is that it can only be used retrospectively, and so is of little use as a screening measure.

The second measure of velocity is achieved by recording serial measures on a distance chart and noting the growth trajectory along major centile lines. Falls in trajectory often occur in infant growth as a result of acute infection, but Edwards et al (1990) found that if a fall in weight across two or more major centiles persists for a month or more this leads to significant anthropometric differences in the second year of life. Edwards et al (1990) also recommend that the predicted growth trajectory should be calculated from weight attained at 4-8 weeks, partly allowing for the rapid growth of small for dates infants which occurs during the weeks immediately after birth. With this criterion small children following a predicted growth trajectory would not be classified as failing to thrive, and regional,

cultural, class or genetic differences are better controlled as an infant's weight centile position is compared with their own previous weight centile position.

It was this criterion that was used originally by Edwards (1987) to identify the cases in this study.

1.2.3. PRACTICAL DIFFICULTIES IN COLLECTION OF DATA

Assessment of growth velocity requires careful longitudinal recording of anthropometric measures. In infants this generally means weight, as length is more difficult to measure accurately and weight is a more sensitive indicator of the fluctuations which occur in early growth. But many practical problems exist in maintaining records of serial weights in the primary health care setting. Davies and Williams (1983) discuss these problems. Infants are often weighed with clothes on and sometimes the weight of the clothes is not deducted. Poorly calibrated scales are used. Weights are recorded using imperial and decimal indices interchangeably and mistakes are often made when writing weights on records. These problems and inaccuracies often make growth charts difficult to interpret and growth failure difficult to diagnose.

Nevertheless Edwards (1987) found that the mothers of infants who grew poorly in the Scotswood area took their infants to clinics for weighing more often than those who grew normally. Such frequent monitoring increases the possibility that errors would be detected.

1.2.4. SUMMARY

A clearly agreed criterion for failure to thrive based on growth alone would eliminate the problem of diagnosis based on subjective criteria. But growth criteria for failure to thrive are not consistent across studies.

A newer criterion used to detect cases in this study relies on a velocity measure. Identification of a fall in rate of weight gain depends on accurate serial weighing and maintainance of records, however, in practice measurement and recording of weights has been found to be deficient. This is counterbalanced by the tendency of mothers in the Newcastle study to attend clinic more often if their child has growth problems, so enabling errors of recording to be rectified.

1.3 EXPLANATIONS OF THE CAUSE OF FAILURE TO THRIVE

In the absence of agreed anthropometric criteria, causal explanations may bias the diagnosis of cases. It is, therefore, important to appreciate the three strands which have shaped thinking about failure to thrive, emotional deprivation mainly through the inadequacy of institutional or parental care, abuse or neglect, and nutritional insult. These three approaches have resulted in nomenclatures for the condition that reflect the empirical approach, such as 'hospitalism' (Spitz, 1945), 'the maternal deprivation syndrome' (Bowlby, 1953), and 'environmental growth delay' (Skuse, 1987).

1.3.1. EXPLANATIONS IN TERMS OF EMOTIONAL DEPRIVATION

The first approach, mainly based on studies of infants raised in institutional care (Spitz, 1945, Bowlby, 1953, Widdowson, 1951), evokes explanations in terms of emotional or maternal deprivation which induce metabolic and absorption disorders. But there was no attempt to study caloric intake of infants in any of the studies which espoused this view. The quality of nutrition provided in institutions was assumed to be adequate.

More recently, however, a study by Rosenn et al (1980) found support for the emotional deprivation argument. In a carefully controlled 9 stage encounter, infants failing to thrive preferred inanimate objects and distal over proximal relations in comparison with children failing to thrive for organic reasons and normally ill children.

Those failure to thrive infants who started to gain weight did so 24-36 hours before or after the distal/proximal pattern had reversed. Intercorrelations between responses to the 9 stage encounter were highly significant within groups but not between groups, clearly differentiating the failure to thrive group from the others. This strongly suggests an effect of emotional interactions on growth. An alternative interpretation is that increased physical vigour which precedes the growth spurt also reduces apathy. But this does not explain why those children with organic failure to thrive, who may have more than one reason to lack vigour, preferred proximal relations.

The focus has now moved away from institutional care, as failure to thrive is now found in children raised in their own homes. For example, a study by Mitchell et al (1980), reported a prevalence of failure to thrive of 9.6 per cent in a predominantly rural poor population, and ironically, reversal of the condition has only been achieved in some primary care cases when children have been hospitalised. Such cases account for between 1-5% of paediatric hospital admissions (Berwick, 1980. Sills, 1978)

1.3.2 EXPLANATIONS IN TERMS OF ABUSE AND NEGLECT

Failure to thrive is reported to be associated with neglect and abuse (Oates and Hufton, 1977. Taitz and King, 1988). Batchelor and Kerlake (1990) point out that a more careful monitoring of weight gain may have alerted health workers and the social services to the tragic plight of Jasmine Beckford and Kimberley Carlile, both found to be underweight and stunted in growth at post mortem. A problem with this view is that not all children who fail to thrive suffer abuse or intentional neglect (Ayoub and Milner, 1985). Frank and Zeisel (1988) report that 10% of cases of failure to thrive have been hospitalised for non-accidental trauma. Therefore if failure to thrive alone were used to detect abuse it would produce more false positive than positive diagnoses, as it is a minority of infants failing to thrive who are found to be abused. However, its importance as an indicator of abuse should not be ignored.

If the direct cause of failure to thrive was abuse then one would expect to find long term effects associated with early abuse. Oates et al (1984) found that children who had been abused scored significantly lower in general intelligence than a comparison group, whereas the deficit in failure to thrive children was in verbal intelligence only. Abused children were also more socially mature than failure to thrive children. These differences in outcome suggest that the cause of abuse may differ from that of failure to thrive.

1.3.3 EXPLANATIONS IN TERMS OF NUTRITIONAL DEFICIENCY

The preferred explanation in recent work is, on the face of it, the simplest. That is, children who do not eat enough do not grow adequately.

Whitten et al (1969) used a novel experimental paradigm to test this hypothesis. Thirteen infants who presented with characteristic physical and behavioural signs of severe neglect, such as poor growth, developmental lags, apathy, a 'frozen watchfulness' and auto erotic behaviour, were admitted to hospital. The infants were put into a windowless nursery, monitored by close circuit television and handled no more than necessary. They were, in effect, deprived of sensory stimulation and affective relations. During this period the infants were offered diets of at least 140 kcal per kg of ideal weight for actual height. All but two of the infants gained weight. When subsequently stimulated, the infants did not increase weight gain. Following discharge, weight gain fell in 6 out of 10 infants, but this was reversed in 4 out the 6 when the hospital diet was fed in the presence of an observer. Throughout the experiment no attempt was made to improve the mother's handling of the infant.

Inadequate feeding by the mother was attributed to mothers sleeping through meals and sating their infants hunger with a biscuit as a substitute for missed meals. Also, mother's faulty perceptions of the amount their child consumed were compounded by misinterpretation of hunger cues. For example, one mother complained that she could not take her infant to other people's houses because he ate 'like a wolf' and she was embarrassed by such inappropriate behaviour. She perceived him as greedy rather than hungry.

Whitten et al (1969) also noted that the characteristic apathetic behaviour in 5 of the infants diminished as they started to gain weight, but before they were placed in a stimulating environment. Their apathy was attributed to lack of physical vigour as a result of undernutrition, rather than their previous history of neglect.

It is interesting to note that 2 infants remained underweight. Both had been force fed by their parents. Frank and Zeisel (1988) also report mixed success in rates of recovery of growth, with 50% recovering, 40% stabilising at a new lower level and 10% continuing to deteriorate. Clearly, if the deficiency was only nutritional, there would be less variability in recovery rates. Therefore, a more complex causal process must be studied in order to understand not only the variation in growth outcome but also behavioural and intellectual outcomes as well.

1.3.4. A SYNERGISTIC APPROACH

In a review of current issues and controversies in failure to thrive Skuse (1985) proposed that management of a case should be 'guided by a hierarchy of factors that are ranged in their proximity to the individual child'. The first of these factors already discussed is nutritional intake, next the quality of caretaking which should take into account the mismatch of temperament between caretaker and child, and finally, the more distal factors which make childcare difficult, such as poverty, overcrowding and family dysfunction. This type of analysis may result in an overemphasis of socioeconomic factors leading to poor detection of the condition in advantaged families in higher social classes.

A different approach is to look for developmental explanations for undernutrition. Chatoor et al (1984) use case histories to illustrate their point that feeding disorders in infants follow a characteristic developmental pattern. From 0-2 months feeding disorders may be associated with constitutional or medical problems. An example of a constitutional problem is that of 'colicky' infants. It is argued that they find difficulty regulating arousal states because they have a labile autonomic nervous system. These infants are vulnerable to overstimulation, cry when fed and have an irregular feeding pattern. A medical problem might be the intubation of infants with respiratory problems. This prevents the development of the sucking reflex into a motivated action

and the coordination of sucking and swallowing. At 2-6 months infant's feeding problems are related to disorders of attachment recognisable by the lack of vocalisation and the avoidance of eye contact. These babies are not cuddly, often stiffening or arching their backs when held. They are more inclined to vomit or ruminate and show little interest in food offered by the mother. The last phase is related to problems of autonomy in the face of a controlling carer. These infants will turn their heads away or spit food out, appear angry and frustrated and are given little opportunity to practice feeding themselves.

There is a complex synergy of medical, temperamental and intellectual characteristics of the child and mother, and the more distal influence of socioeconomic factors which contribute to the cause and expected outcome of early growth retardation. It would be surprising, therefore, if a single outcome characteristic of all types of failure to thrive were found. Instead, it is important to control for risk factors and measure the effects of key variables.

1.3.5.SUMMARY

Early theories about the causes of failure to thrive came from divergent perspectives, based on empirical work with specific populations of children such as orphans in institutions, or abused children. Nutritional intake was not recorded in these studies. Failure to thrive has since been found in children living at home and psychological outcomes of abused children were not consistent with those of children who had failed to thrive.

A convergence of opinion has focussed on nutritional deficiency as the proximal cause of failure to thrive. The origin of nutritional deficiency may depend upon the developmental stage of the infant. Problems relating to poor intake of nutrients in early development may be the result of poor regulatory mechanisms or constitutional weakness, in later development, a lack of attachment and lastly, to questions of autonomy versus dependence.

The factors that contribute to poor growth may be complex. These factors include medical, temperamental, and social characteristics of both the infant and the principal carer and the transactions between them. These factors which contribute to undernutrition may produce heterogenous intellectual and behavioural outcomes, rather than a single outcome characteristic of failure to thrive.

1.4. REVIEW OF OUTCOME STUDIES

The purpose of this review of outcome studies is to evaluate the evidence that cognitive deficits and behavioural problems are associated with failure to thrive in infancy. For this purpose I have divided a sample of studies into those based on cases referred to hospital and those based on an unpreselected population.

1.4.1 STUDIES BASED ON REFERRED CASES.

Overwhelmingly, studies based on cases of failure to thrive referred to hospital have found developmental delay and cognitive deficits. This apparent consistency disappears upon closer examination of the findings, both in terms of when the delay occurs, the degree to which nutritional insult causes permanent damage, the reversibility of developmental delay in response to environmental change, and how the characteristics of failure to thrive, such as age at onset, duration and severity, affect outcome.

The first issue to examine is whether developmental delay is a presenting feature of failure to thrive. Glaser et al (1968) found that 'failure to develop motor skills' was one of the presenting problems of 50 growth retarded

children. They also found that 6 cases were mentally retarded at follow up and 7 of the 19 children over the age of 5 were having problems at school.

Drotar et al (1985) argued that the inclusion of developmental delay in the diagnostic criteria is likely to have a confounding effect on outcome. This is questionable since McCall (1979) argued that scores for mental performance assessed during the first 18 months of life do not predict later IQ to any practical extent. However, Largo et al (1990) found that there is greater stability in test scores at the lower levels of mental performance since cerebral impairment limits the variability in development over time. How true this is of failure to thrive depends on the extent to which poor growth has caused permanent neurological damage, and whether the cause of poor intellectual development can be remedied. Therefore, the effect of family environment and intervention strategies on a child's development need to be considered when measuring outcome of failure to thrive.

Family environment contributed to the pessimistic prognosis for mental and emotional development in the Glaser et al (1968) study, as developmental delays occurred within the context of family dysfunction. It was reported that a small number of children from more favourable home environments experienced spontaneous recovery.

Field (1984) also found that infants failing to thrive were delayed upon admission; but they subsequently improved to the normal range. She followed up 17 infants aged between 3 and 19 months, 1 month, 3 months and 6-13 months after discharge from hospital. When admitted, mean scores for mental quotient were 83.3, Bayley Mental Development Index (MDI) 81.5, and motor quotient, 80.5. These scores improved to the normal range for mental quotient (95) and Bayley MDI (93.4) at follow up, although there was continuing evidence of delayed motor development (motor quotient, 85).

Delayed motor development was explained by Field either as a result of maternal practices, for example rarely holding the baby, or damage to the cerebellum during the post natal growth spurt as a result of nutritional deficiency. No uniform intervention was planned following discharge, but 4 of the children were placed in foster care and various forms of contact were maintained with the remaining cases which, Field argues, spurred the mothers to maintain weight gain.

These studies together point to a pattern of developmental delay and subsequent recovery depending on the caregiving environment, with the exception of Field et al (1984) who raises the possibility of permanent neurological damage affecting motor development.

By contrast Drotar and Sturm, (1988) found the opposite pattern when they tested 59 infants weighing less than the 5th centile of the National Center For Health Statistic (NCHS) norms. The infants were not delayed at intake but were at follow up. Bayley scores at intake had a mean of 99.6, whereas Stanford-Binet scores administered when the children were aged 3 years old had a mean of 85.4. The effect of three types of intervention, family centered, parent centered and advocacy, were also studied. There was no significant effect of intervention. This suggests that damage during a critical period of development had not been remedied by environmental factors.

In order to discover if the deficit was a result of neurological damage during a critical period or the result of a poor caretaking environment Singer and Fagan (1984) carried out a controlled study of infants in their first year. They reported no significant difference in the performance of a visual memory test of infants with failure to thrive of organic or non-organic aetiology compared to normally growing controls. However, Bayley MDI was significantly lower in both the organic and non-organic failure to thrive groups compared with controls both at the time they completed the visual memory test, at 20 months old and 3 years old. From this Singer and Fagan concluded that the impairment was not therefore neurological but environmental. But in a later study of 29 children followed up at 3 years old, although Singer

(1986) again found depressed scores using the McCarthy Scales of Children's Ability, she found no significant difference in IQ attributable to different placement and intervention strategies, suggesting a more limited effect of environment.

Although these studies have all found an association between developmental delay and failure to thrive, there is little agreement about when the delay occurs, or the degree to which recovery can be achieved through intervention. The variations in these findings may be explained by the particular characteristics of failure to thrive in an infant, such as age at onset, severity, and duration of the episode of growth retardation.

In a controlled retrospective study of 19 infants diagnosed as failing to thrive, significant associations were found between failure to thrive after 4 months of age and developmental deficits at follow up (Chase and Martin, 1970). No associations were found in this study between infants whose growth faltered before 4 months and developmental delay. It is not possible to analyse the other characteristics of failure to thrive in this study as the criteria for diagnosis were not specified.

Using a follow up design instead of a retrospective design, Field (1984) found that age at onset did not predict outcome, but again the criterion for inclusion in the study was vague, using only the 'physician's diagnosis'.

By contrast, Drotar et al (1985) in their follow up study of 80 children, using no less than 11 criteria for inclusion, found that early age at onset significantly predicted Bayley MDI at 24 months so that the earlier the onset the lower the predicted Bayley MDI. This finding supports the view of Chatoor et al (1984) that early onset failure to thrive identifies a distinct subtype of infants who present with disorders of homeostatic mechanisms or constitutional difficulties, such as poor suck. This type of disorder suggests that neurological damage may have preceded the condition, though these infants were not found to be delayed at intake.

Degree of wasting (severity) and chronicity (duration) defined as the discrepancy between the estimated age at which the child reached the 5th centile and the age at intake, was not found to significantly predict Bayley MDI in this study. But, as 87% of cases showed a mild degree of wasting (only 7% were severely malnourished) and the age at intake was between 1-9 months, the range of these measures may have been too small to predict developmental delay. From these studies it could be concluded that

certain characteristics of failure to thrive such as age of onset, contribute to depressed intellectual and developmental quotients in some infants. But, despite this apparent consensus there are unexplained differences, for example whether developmental delay is detected at intake or follow up and whether it is late or early onset that is associated with IQ deficits. The variation in findings could possibly be accounted for by individual differences and protective factors, but this seems unlikely as the same kinds of differences would emerge in each study. The most probable explanation for the inconsistencies of results lies in the multiplicity or lack of specific criteria used to identify study populations. In other words, since there are no common diagnostic criteria, we may not be comparing like populations of children failing to thrive.

The issue of diagnostic criteria has been discussed already, but in order to discover if there is an intellectual deficit directly attributable to undernutrition, using case studies with a clear diagnostic basis, it is possible to use another strategy. Several studies exist which have examined the intellectual outcomes of early malnutrition with an organic or known aetiology. Berglund and Rabo (1973) found a small but not significant IQ difference in 180 individuals tested at the time of national service who had untreated pyloric stenosis in infancy. Lloyd-Still et al (1974) followed up 41 individuals with various diagnoses that affected

nutritional intake and weight gain. They found significant developmental differences in the younger children when compared with siblings, but in the older children, these differences were small (less than 3 IQ points) and not statistically significant .

Age at follow up was much longer in these latter studies. It could be argued that recovery from intellectual deficits may be a function of time.

1.4.2 POPULATION BASED OR PRIMARY CARE OUTCOME STUDIES

There is less consensus about the inevitability of developmental delay in studies screening whole populations: indeed, it is contentious whether IQ deficits have been found at all.

Mitchell, Gorrell and Greenburg (1980) identified 30 cases of failure to thrive from the medical records of a cohort of 312 children followed in three rural primary care centres. These cases were defined as under 80% of normal weight in the first 24 months if the previous weight was above 80%. The cases were compared with a matched control group between 3-6 years of age at follow up using the McCarthy Scale of Children's Abilities. No statistically significant difference was found in cognitive ability between the cases and controls. Unfortunately, only 19 of the 30 cases were sought for follow up, and only 12 completed a full examination. Consent was refused for the

further participation of the 7 cases traced but not followed up. The high rate of attrition in this study leaves these findings open to question.

On the other hand, a recent, well conducted population study was carried out which found a worryingly large level of intellectual impairment in their failure to thrive subjects. Dowdney et al (1987) gathered information from child health records of 1868 infants registered with an inner city community paediatric service. Subsequently, 25 children were identified as failing to thrive. The last recorded weight for these children was below the 10th centile, they were still below the 10th centile for height and weight when traced at 4 years old and remained below the 10th centile when the height of both parents was taken into account. Cognitive assessment of 23 'cases' was carried out using the McCarthy Scales. Results were compared with those of a normally growing control group selected from the remainder of the cohort. A 20 point difference (Cases 77.1 SD 17.6, Controls 97.7, SD 15.2) was found in general cognitive Index (GCI) scores. This represents a massive fall in IQ points especially as the control group were themselves 10 points below the mean GCI scores found in a sample of British children (Lynch et al, 1982).

This study followed many of the design protocols required to produce convincing evidence, in that it was a controlled population based study, with a low rate of attrition, using selection criteria based on growth alone and testing was carried out blind. It is interesting to note that only 4 of the cases had been referred for investigation of failure to thrive, therefore the 'cases' had not benefitted from (or were uncontaminated by) any intervention.

The main problem with this study is that the exclusion criteria were so rigorous. After premature babies, those with congenital defects and those whose current weight was above the 10th centile were excluded, 61 of the original 138 'cases' remained. Those whose mid parental height was below the 10th centile were further weeded out. The final selection through paediatric examination and interview with the mother reduced the study population to only 23 'cases', so limiting the generalizability of findings.

It should be borne in mind that the cases in this study represented a chronic form of growth failure as growth retardation had been detectable for three years; but it is not clear how severely growth must be retarded before intellectual impairment occurs. Also the selection criteria used have limited clinical use as they can only be applied in retrospect after the damage to cognitive development has apparently already been done.

Dowdney et al (1987) argued that subsequent analysis of their growth data revealed that all the cases could have been detected from a fall in weight gain in the first year of life. Weight gain, expressed in standard deviations from the mean of National Growth Standards (Tanner and Whitehouse, 1976), fell from $-.81$ at birth to -1.95 SDs at 12 months, compared with controls who were -1.09 at birth and $-.49$ SDs at 12 months. This shows that the growth failure was not only chronic, but also severe and could have been detected earlier by a fall in weight gain.

1.4.3 METHODOLOGICAL PROBLEMS

The contrary findings of the outcome studies reviewed lead one to question whether developmental delay and behaviour problems are caused by failure to thrive, or merely an artifact of the sociology of the referral process and defective study design. Evidence in support of this argument comes from children who recover from failure to thrive with an organic or clearly identified aetiology. These children show little evidence of long term cognitive or behavioural effects (Berglund and Rabo, 1973, Lloyd-Still et al, 1974).

In a discussion of sampling problems in research with non-organic failure to thrive children, Drotar (1990) indicates five sources of variation as a result of samples used: the criteria used to define failure to thrive, the populations from which samples are drawn, the rate of attrition, the extent and type of intervention locally

available, and the inclusion or exclusion of environmental and biologic risk factors, such as family dysfunction or low birth weight. To these sampling problems could be added further problems in some studies of small sample sizes, lack of comparison groups and tester awareness of the clinical status of subjects.

The first issue of diagnostic criteria has already been discussed. The remaining design issues will now be examined. Firstly, sampling issues such as dependence on referred cases for subjects, small sample sizes and too many exclusionary criteria will be discussed, then, lack of comparison groups, high rates of attrition at follow up and tester awareness of the status of subjects.

The consistent finding of developmental delay at some stage in referred cases may be an artifact of the referral process. Children referred to a specialist clinic or hospital as a result of growth failure may not represent the whole population of children who fail to thrive. Clinic samples depend upon detection and referral decisions. The absence of clearly defined diagnostic criteria leads to the detection of different populations of infants failing to thrive at the level of the community. As a result of this lack of clarity, Batchelor and Kerslake (1990) found that a diagnosis of 'small', as opposed to failure to thrive, seemed to 'rest on the absence of factors associated with deprivation, rather than on positive evidence of a child being

constitutionally small'. Decisions about referral of identified cases depend on the severity of the condition, how difficult it is to diagnose or how responsive it has been to outpatient management (Drotar, 1990). The latter may depend on a physician's judgment about the ability of the caregivers to effect the child's recovery. Therefore infants from the most dysfunctional and deprived families could be expected to be found in specialist clinic populations. This may have increased the likelihood of spurious relationships being found between failure to thrive and intellectual and behavioural outcomes.

An example of the effect on outcomes of these implicit referral criteria can be found in a study by Hufton and Oates (1977) which found a high level of problems amongst the families of their referred 'cases'. As dysfunctional families affect behavioural outcomes, so it should not have been unexpected when they also found that, at follow up, emotional status, as assessed by teacher's questionnaire, rated 10 out of the 21 children abnormal, 5 antisocial and 5 neurotic. Verbal IQ was also found to be 20 points lower than performance scores.

Hufton and Oates (1977) report;

'The families of these children have high incidence of marital instability and economic difficulties.'

This clearly implies that family problems are associated with failure to thrive, not merely a reason for referral and behavioural problems and IQ deficits are an outcome of early failure to thrive, not of the referral criteria applied at the time of diagnosis.

Pollitt and Eichler (1976) tried to circumvent the problem of referral by choosing both cases and controls from an outpatient clinic. In both groups, with a few exceptions, the reason for attending the clinic was gastro-intestinal, ear or upper respiratory tract infection. Some of these children were found to be growth retarded. They found an overall non-specific increase in behaviour problems reported by the mothers of children who fail to thrive and a significant increase in reported feeding problems with more skipped or skimpy meals. However, when matching of socioeconomic variables was checked, there were significant differences between per capita income and mother's education. These differences may have accounted for the differences in behavioural rating.

Reasons for inconsistency of findings between studies, such as the characteristics of growth which affect outcome may be found in sample size. Small samples do not have the statistical power to detect quite large differences (Cohen, 1992). Field's (1984) sample of 17 subjects, only 13 of whom completed developmental testing, was not a large enough sample to detect the effect of age of onset of failure to thrive. The sample of 80 infants (69 followed up) in a study by Drotar et al (1985) provides much stronger evidence of an effect of early onset failure to thrive on subsequent IQ tests.

Exclusion of subjects based on the differentiation of infants according to organic and non-organic aetiology of growth retardation and perinatal factors has also limited the usefulness of research findings. The differentiation between organic and non-organic failure to thrive was questioned by Homer and Ludwig (1981) when they proposed that a third category should be created to accommodate those children whose failure to thrive had an organic aetiology, but in whom the effects of the organic disease were compounded by psychosocial factors. Furthermore, infants who fail to thrive have been found to have a greater susceptibility to otitis media and other low grade persistent illnesses (Sherrod et al, 1984).

Many studies also exclude children with low birth weight from their study populations. Yet, as Frank and Zeisel (1988) point out 10-40% of failure to thrive cases have birth weights below 2500gm compared to 7% of the general population and even when low birth weight babies are excluded, the remaining infants have significantly lower weights than controls. Therefore, by excluding these categories, an important body of the failure to thrive population is not studied and conclusions cannot be extrapolated to apply to them.

The use of clinic populations as a sampling strategy, according to Drotar (1990), 'introduces bias, limits generalizability of findings, leads to an erroneous assumption of sample homogeneity and results in contradictory findings in studies conducted in different settings'.

A safeguard against the effects of sample bias would be comparison with a control group matched for variables known to be associated with outcomes. However, in some instances, the studies of referred cases have not had the benefit of a comparison group (Hufton and Oates, 1977, Glaser et al, 1968, Elmer et al, 1969, Field, 1984), or a comparison group has been recruited retrospectively which may only share demographic variables at the time of study, but not at the time of diagnosis (Oates et al, 1984, 1985).

Another critical problem with any follow up design is the rate of sample attrition or parental refusal. Singer and Fagan (1984) lost 14 out of 39 subjects in a three year period and Oates et al (1984, 1985) had a rate of attrition of 16 out of 30 after 12 years. Mitchell, Gorrel and Greenburg (1980) in one of the few studies in the primary care setting only completed testing at follow up of 12 of the original 30 cases. Drotar (1990) expresses concern that the effects of this form of sample self selection is not adequately evaluated as demographic variables of those lost are not often compared to ensure that they do not differ from the follow up sample. It has been found that families of infants that cannot be traced have more psychological problems than those who can (Aylward et al, 1985) so introducing a possible source of bias into the sample available for follow up. In the same way parents who refuse to participate in a follow up study have also been found to have more psychological problems (Beck et al, 1984).

A further methodological problem which has been recognised in the conduct of case and intervention studies is the difficulty of maintaining the 'blind' status of testers. Many studies do not specify whether psychological testing was carried out by workers who were aware of the clinical status of subjects or not.

However, it must be inferred, from the method described, that the researchers did indeed know the status of the child. It was shown by Rosenthal (1966) that knowledge of the subject's status accounts for error which creates bias in favour of the hypothesis being tested at all stages of the investigation. Another problem is that mothers are usually aware of the growth status of their child, which may influence their responses when rating their child. Pollitt and Eichler (1976) are amongst the few who specify that they were unable to use a double blind design because of the nature of the study.

In order to prevent bias it is necessary to incorporate certain design features into a study of this nature. Selection should be based on growth criteria alone which can be applied to a whole unpreselected population. A sufficiently large number of cases should be matched with a group which controlled for demographic and family variables. The rate of attrition would need to be minimal and follow up testing carried out blind.

1.4.4. SUMMARY.

Studies using cases already identified by primary care physicians and referred to hospital or out patient clinic consistently support the view that cognitive impairment and behavioural problems are related to failure to thrive in infancy. Within this consistency there are disagreements over when developmental delay becomes manifest, the efficacy of intervention and the characteristics of failure to thrive which are associated with delay. There is also evidence to show that there is less effect of early developmental delay the older the age at follow up.

Unlike studies of referred cases, population based studies of children who failed to thrive in infancy do not uniformly find an effect on cognitive function.

This diversity of findings can be explained by design problems in many studies of failure to thrive. The comparability of results is reduced by varying criteria for identifying cases. Samples using 'cases' have been biased by the referral criteria applied and have often been too small to detect statistically significant differences. The generalizability of results is limited by the use of many exclusionary criteria. Many studies have not benefitted from comparison with a matched control group, have suffered from high rates of attrition and may have been biased by tester awareness of the subjects clinical status.

1.5. FACTORS AFFECTING IQ AND BEHAVIOUR

Any discussion of tests of possible cognitive and behavioural outcomes of failure to thrive must take into account the direct effect of undernutrition on the brain, indirect effects such as increased susceptibility to teratogens, and the context in which children are found to be growing poorly.

1.5.1. THE DIRECT EFFECT OF UNDERNUTRITION ON THE BRAIN.

Study of the effects of undernutrition on the human brain have relied heavily on inferences based on the effects of undernutrition on animals, principally rats. Rat pups undernourished between 13 and 28 days old have smaller brains, a disproportionately small cerebellum, fewer glial cells and fewer lipids characteristic of myelin (Dobbing and Smart, 1974). Also altered enzyme activity lead to an increase in acetylcholinesterase which is related to a disproportionate increase in cholinergic nerve endings.

The behavioural outcome of these structural and functional changes was found to be poorer co-ordination. Undernourished rats could not walk backwards on a revolving drum as well as controls and missed their footing more often on a bridge of parallel rods. They were also found to react more to aversive stimuli and were more excitable and aggressive.

The use of the rat as a model for studying the effects of undernutrition on cellular growth was questioned by Winick et al (1970) when they compared wet weight, dry weight, total protein, RNA and DNA in the brains of normal infants post mortem and those who died of marasmus. In normal brains they found a linear increase in the first two years of life in all measures except DNA. DNA (cell number) increased in the cerebrum between 6-8 months and the cerebellum between 8-10 months. There were reductions in all measures in all regions of the brains of undernourished infants; but the reductions in DNA were disproportionately greater in the cerebrum and cerebellum than the brain stem. It was argued that the growth spurt period was not comparable with that of the rat and that conclusions from animal experimentation should be treated with caution.

Dobbing and Sands (1973) argued that the disparity of findings was due to the different maturational timetables followed by different species. The brain growth spurt was completed in utero in guinea pigs, post nately in rats and perinatally in pigs and humans. Comparisons between species should take the timing of the brain growth spurt in relation to birth into account.

In order to ascertain the timing of the human brain growth spurt in relation to birth to facilitate comparison with other species Dobbing and Sands (1973) dissected 139 human brains between the ages from 10 weeks gestation to 7 years old. In addition they dissected 9 adult brains. They found that total cell numbers rise reaching a plateau at about 2 years. The cerebellum starts to grow later than the forebrain or stem but finishes earlier. Dobbing (1984) also reports a post natal increase in cell population in the hippocampus. The difference in the timing of the growth spurt from those found by Winick et al (1970) can be accounted for by the fall in cellularity (ie. the proportion of DNA to tissue) despite an increase in cell number. Dobbing and Sands (1973) conclude that the brain growth spurt is much more post natal than previously supposed and in this way the timing of human brain development resembles that of rats.

I know of no studies, as yet, that have established a clear relationship between a known abnormality in the brains of undernourished human infants and behavioural or cognitive outcomes.

Newer techniques, such as Magnetic Resonance Imaging (MRI), have allowed the study of the effects of undernutrition on the brains of subjects in vivo. Gunston et al (1992) used MRI to monitor the effect on the brain of recovery from Kwashiorkor in 12 subjects aged 6-37 months over a period of 90 days. All subjects showed signs of brain shrinkage in white and grey matter on admission. After 90 days shrinkage had reversed and in 9 of the subjects had recovered completely. They suggest that shrinkage is not due to loss of myelin lipids as myelin was assessed as normal, but to fluid shifts as a result of changes in osmolality. They conclude that 'if environmental factors and social stimulation at the time of undernutrition are favourable the chances of retarded development ... are diminished'.

1.5.2. INDIRECT EFFECTS OF UNDERNUTRITION ON THE BRAIN

One effect of undernutrition, mediated by increased absorption of lead, may be a reduced rate of production of serotonin, a CNS neurotransmitter. Bithoney (1986) reported that lead and other heavy metals are absorbed more readily by infants who are undernourished and that the behavioural and intellectual problems reported in children who failed to thrive were of a similar nature to those reported of children with elevated lead levels.

In a study of 45 cases (weights and heights for age less than the 5% of NCHS norms) referred from primary care he found a significant difference between cases and controls in blood lead levels and that the cases were significantly more anaemic. When tested using the Denver Developmental Screening Test high failure rates were reported in both language and motor skills.

The study illustrates the increased susceptibility of the brain to teratogens as a result of undernutrition, which may produce a more generalised deficit which would be detected by IQ tests and behavioural assessment. It should, nevertheless, be noted, that no significant correlation was found between lead levels, anaemia and the Denver Developmental Test, therefore, the effect found may be an effect of undernutrition alone.

1.5.3. EFFECT OF THE CONTEXT OF UNDERNUTRITION

Poor performance on IQ tests may be explained more easily by factors which contribute to undernutrition, rather than the undernutrition itself.

Feeding in infancy is a social interaction with the caregiver, since infants cannot select or prepare food themselves, and learning to feed to themselves successfully occurs throughout infancy. Feeding depends on the ability of the infant to elicit food from the caregiver, the caregiver's responsiveness to signals from the infant, and, as Stern's (1977) work showed, the ability of the caregiver and child to time their interactions and to moderate stimulation to a level acceptable to the child. Temperamental differences and behavioural problems in either the mother or child may result in a failure of these social interactions to result in adequate care of the child. It has been noted that these problems may be exacerbated by social factors which make caring for an infant difficult (Skuse, 1985).

A recent study which attempts to disentangle the separate environmental components that contribute to social class differences in IQ was conducted by Sameroff et al (1987). In this study the effects on verbal IQ of 10 variables, maternal mental health, maternal anxiety, parental perspectives, interaction, maternal education, occupation of the head of the household, membership of a minority group, social support, life events and family size were separately assessed. Then the cumulative effects on verbal IQ of these variables were calculated.

All risk variables were found to have a significant effect on verbal IQ when compared with the group of subjects with the risk factor absent. They are shown in descending order of statistical significance in Table 1.

Table 1.

RISK VARIABLES SHOWN TO BE SIGNIFICANT PREDICTORS OF IQ.
(After Sameroff, Seifer, Barocas, Zax and Greenspan, 1987)

	WIPPSI VERBAL IQ		
	low risk	high risk	t value
Minority group membership	109.6	91.3	8.06
Maternal education	108.6	92.5	6.73
Parental perspectives	107.3	90.4	6.17
Occupation of head of house	105.9	87.7	6.02
Family support	107.3	95.3	4.87
Family size	102.5	92.0	4.08
Interaction	105.6	95.7	2.90
Maternal anxiety	105.1	97.6	2.51
Life events	102.5	92.0	2.39
Maternal mental health	104.5	99.3	1.98

All except the last three variables were significant at the $P < .01$ level. The last three were significant at the $P < .05$ level.

Children in lower social classes were found to have a larger number of risk factors present and the greater the number of risk factors, the greater the IQ deficit. The difference between the mean IQ of children with 0 and 8 risk factors present was in excess of 30 IQ points.

It is of particular interest to this investigation that the same risk variables related to verbal IQ in the Sameroff et al study (1987), such as occupation, family size and interaction, have also been linked to poor growth and incidence of failure to thrive (Tanner, 1989. Chatoor et al, 1984).

A further consideration is that intelligence and emotional development may be affected by different aspects of the environment. Schaffer (1977) argues that language and emotional responsiveness is stimulated by the way mothers interact with their infants. However, language and emotional development are also partly independent of each other, as it has been found that affectionate mothers of large families may produce emotionally healthy children who have a lower IQ (Tizard, 1977).

It appears that different aspects of the caring environment affect physical, cognitive, emotional and behavioural development independently. This point was argued by Rutter (1972) in his reassessment of 'the maternal deprivation syndrome'. What is not clear is the extent to which poor growth underlies poor outcomes in any of these domains, consequently more than one test of developmental outcomes is required.

1.5.4 SUMMARY

Differences in structure and function of the brain have been found in both undernourished humans and animals. Animal studies have shown that there are long term effects of undernourishment on motor control and levels of excitability and aggression. However, it has been found that shrinkage of the brain in infants with kwashiorkor is reversible.

Undernourished infants are more susceptible to environmental teratogens such as increased absorption of lead which is associated with lower IQ and behavioural problems.

Performance on IQ tests is associated with environmental variables, such as social class and family size. Some of these variables are also associated with growth. At the same time growth, intelligence and behavioural and emotional development can be independently affected. It is important therefore, to assess these separately.

1.6. OUTCOME EVALUATION

1.6.1. IQ TESTING

An IQ test has the practical value of identifying a group of children who may be educationally disadvantaged and require special education. In addition, if it can be shown that IQ is associated with a particular pattern of growth in infancy, early intervention may be instituted to prevent later learning difficulties.

Many studies of failure to thrive outcomes have followed up subjects within months of original diagnosis, but IQ or developmental tests carried out in infancy, such as the Bayley scales, have low predictive validity. They are better used to ascertain the child's current developmental status (Bayley, 1970) but not to predict long term outcomes.

Since the predictive validity of an IQ test is one of the main justifications for its use, then it is clearly preferable to delay follow up assessment until scores stabilise.

After infancy, an IQ test is a stable indicator of the skills acquired for formal schooling and it is a moderate predictor of later academic achievement. Typically the correlation between test performance and school performance is .50. (Snow and Yalow, 1981). This has considerable practical value when evaluating the importance of the implications for failure to thrive.

1.6.2. SOCIAL COMPETENCE AND BEHAVIOURAL PROBLEMS

Weschler (1974) discussed the role of non-intellectual factors, such as motivation and temperament, in determining test scores. Scarr (1981) argues that these factors are just as important in functioning in later life as intelligence and cognition; socially competent school-retarded children can adjust normally to the demands of life after leaving school.

The behavioural and motivational qualities which constitute social competence are nurtured in the interactive environment. Scarr (1981) found that mothers who reasoned with their child and encouraged them in an experimental task and reported that they typically used these techniques had children who were more co-operative, attentive, less overactive, had better communication skills and higher IQ. These behaviours predicted later IQ and IQ predicted these behaviours in a later test. Scarr found that these measures of the mother's control and disciplinary techniques were a better predictor of the child's competence than mother's education or IQ.

Assessment of behavioural problems is justified on the basis that they may affect the ability to 'blend into the general population' (Scarr, 1981) as much as IQ test performance, that behaviour measures may predict IQ and behavioural problems have been reported alongside school failure in previous studies of failure to thrive (Elmer et al, 1969. Glaser et al, 1968. Hufton and Oates, 1977). If it was found that children who failed to thrive had lower IQ and more behavioural problems than a matched group, then this would be a useful pointer to future research and possible interventions.

1.6.3. SUMMARY

IQ test scores become more reliable, with better predictive validity after infancy and the constructs of intelligence tests are better understood when they become related to the skills necessary for school performance. IQ tests predict academic achievement which has strong implications for interventions in disadvantaged groups.

An individual's ability to integrate into the community can be predicted by social competence. Socially competent school retarded people are able to adapt to the requirements of life after school. Previous studies have shown that failure to thrive is associated with poor social competence and behaviour problems and these constructs should be separately assessed.

CHAPTER 2

2.1. AIMS AND OBJECTIVES OF THE STUDY.

It is clear from the preceding discussion that there are several requirements for an effective test of the hypothesis that failure to thrive in infancy results in IQ deficits and behavioural problems.

The first requirement is that the growth criteria used for selection should be made explicit as different criteria identify different populations and it will only be by comparing the outcomes of these populations that relevant criteria will emerge.

Growth measures which select children below an arbitrarily selected centile must include normal small children, therefore, a growth velocity measure is preferred. Velocity charts are not used in primary care, but rate of weight gain can easily be ascertained by plotting serial weights on a distance chart. Serial weights plotted from 4-8 weeks have been found to be a more stable predictor of growth trajectory than birthweight.

Secondly, failure to thrive should be studied in the primary care setting and inclusion as a case should be decided according to the records of growth kept in child care clinics or by GPs. In this way, all children whose pattern of growth meets the criterion used will be classified as 'cases', not merely those referred to specialised clinics because they are more resistant to treatment, or because they come from poor home environments. This also eliminates the problem of using too many exclusionary criteria, so limiting the conclusions that can be drawn from such a study. A large enough sample size should be provided to detect an effect on IQ and behaviour. The outcome for every child who shares the specified criteria can then be measured irrespective of background or subjective diagnostic criteria.

A control group must be used. The advantage of selecting a control group at the time when cases were first identified is that we know, from records studied contemporaneously, that controls were growing normally at the time and differed in their growth patterns from cases. We also know that they shared matching criteria at the time of diagnosis, if not subsequently.

A disadvantage of selecting a control group at the time cases were identified in infancy is that many subjects may be lost to follow up or may refuse to participate. Every effort should be made to prevent loss of subjects. In the event of a high rate of attrition, the characteristics of subjects lost must be compared with those remaining in the study to determine how their loss may bias results.

Testing and analysis of results should as far as possible be carried out blind. Unfortunately, it is not possible to use a double blind method on tests completed by mothers, as they may be aware of the growth problems experienced by their own infants. But neither the tester nor other participants, such as teachers need know the original status of the child. In this way results cannot be influenced in favour of the hypothesis being tested.

IQ testing should be carried out after infancy when results are more stable and have greater predictive value.

The design characteristics outlined here should provide an effective test of the hypothesis that children who failed to thrive in infancy have lower IQ and more behaviour problems than a comparison group.

2.2. METHOD

2.2.1 OVERVIEW OF THE STUDY

This is a prospective investigation of the cognitive and behavioural outcomes of a group of children who were classified as failing to thrive in infancy. They were originally studied by Edwards (1987). Edwards identified 54 children as failing to thrive from a cohort of children, born during one year, in two deprived wards of Newcastle upon Tyne. Failure to thrive was determined by a single growth criterion alone. Cases were matched with controls at the time of identification at 18 months of age. Approval to study these children further was gained from the Newcastle Health Authority Medical Ethics Committee.

In 1991 I traced 91% of the cases and 88% of the controls and sought consent from the families for a follow up study of these children in which mothers were asked to rate their behavioural problems using the Achenbach Child Behaviour Checklist (1991). The results were reported in my third year undergraduate dissertation (Corbett, 1992).

In 1992 I retraced the families to inform them of the results of the first study and to seek consent for a further follow up study to rate the children's behaviour and test IQ in school. Consent was also sought from the Local Education Authority and Head Teachers. All those families traced the year before consented to the participation of their children in this study.

All follow up investigations of behaviour and cognitive development were carried out by a researcher 'blind' to the status of the subjects. This was achieved by engaging the services of a second psychologist, Alice Thams, a visiting student from Aarhus, Denmark. Testing was carried out under my supervision and just under half the children were allocated by me for her to test, including those whose status had become known to me.

The independent variable was the growth criterion used. The outcome variables were IQ, as measured by the WIPPSI-R, and behaviour problems reported by teachers using the Teacher's Report Form.

The aim of the follow up study was to test the hypothesis that failure to thrive in infancy as identified by the screening criterion used by Edwards (1987) was associated with a subsequent lower IQ and an increase in reported behaviour problems.

2.2.2 SUBJECTS

For the original study (Edwards, 1987) a cohort of children born between November 1985 and November 1986 attending the Armstrong Road Clinic, Scotswood and Cruddas Park Clinic, West City were serially weighed a minimum of six times. Follow up weights were acquired by GPs, child health clinics, and health visitors for all the children on the register. The children were then screened using a criterion which identified those with poor growth as failure to thrive cases. A case was defined as follows:

' A child whose weight deviates downwards across two or more major centiles from the maximum centile achieved at 4-8 weeks for a period of a month or more.' (Edwards et al, 1990).

When applied to the Newcastle Cohort, the criterion identified 63 children as having failed to thrive at one time or another during the first 18 months of life. A further 219 children were found to be growing normally. Using this screening criterion, a prevalence of failure to thrive of 20.9% was found at the level of the community.

Children excluded from the study were those whose characteristics were associated with different growth trajectories; two pairs of twins, 12 Asian infants, and of the 63 cases, 5 children who had growth problems with a known organic aetiology. A further 24 children could not be included as they did not have the minimum requirement of six serial weights on record. Six cases were lost as 3 had moved, 1 was untraceable, 1 parent refused consent and 1 GP expressed concern that further professional intervention may prove detrimental to the child.

The remaining fifty two cases were matched with controls selected from the 219 children in the cohort found to be growing normally. The next child on the same clinic register that matched on sex and age to within one month was selected as a control. In one instance a control from West City was matched with a child on the Armstrong Road register to achieve a better age matched pairing. Where consent from the parent was not obtained the control was replaced by the next child on the register to fulfill the criteria.

Groupwise matching of socio-economic factors was believed to have been achieved by matching the control child from the clinic register as those attending were from a highly localised deprived population with few people choosing to come into the area to attend these clinics, and homeowners tend to register elsewhere.

An analysis of deprivation and health in the North of England (Townsend, Phillimore and Beattie, 1988) ranked West City as the 3rd most deprived ward out of 678 wards in the North. Scotswood was ranked 12th most deprived. This was characterised by high unemployment and overcrowding, low levels of car and home ownership and greater numbers of people classified as class 4 and 5. West City was also ranked the second ward with the worst health out of 678. Scotswood was ranked 12th. Poor health was determined by the standardised mortality ratio, number of deaths, number and percentage of those permanently sick, and number and percentage of low birth weights. The data on low birth weight is of particular interest because of its association with subsequent failure to thrive (Frank and Zeisel, 1985). The percentage of low birth weights were 18.6% for West City and 19.6% for Scotswood. These compare poorly with middle ranked Crawcrook and Greenside (12%) and some of the higher ranked wards had no incidence of low birth weight at all for that year.

Edwards (1987) found that socio-economic factors of the cases and controls did not significantly differ, although cases tended to be of a slightly lower class and slightly more deprived, as measured by levels of car and home ownership, unemployment and presence of the natural father in the home.

It is possible to compare the Townsend et al (1988) figures for employment, car and home ownership for the area with the same figures gathered contemporaneously by Edwards (1987) for the cases and controls followed up in this study (Table 2).

Table 2.

COMPARISON OF UNEMPLOYMENT, CAR AND HOME OWNERSHIP FOR THE OVERALL AREA, CASES AND CONTROLS.

	WEST CITY (From Townsend et al, 1988)		SCOTSWOOD CASES N=50 CONTROLS N=52 (From Edwards, 1987)	
% UNEMPLOYED	29.8%	26.5%	78%	69%
% NO CAR	84.3%	73.7%	92%	76.9%
% NOT OWNER OCC.	97.1%	68.2%	94%	88.5%

Both the cases and controls have higher rates of unemployment than those found in the wards, cases have lower levels of car and home ownership and car and home ownership amongst controls are similar to the low levels found in the area. It can be seen that not only was this study carried out in two of the poorest wards in the North, but also amongst some of the poorest people in those wards. The groups were closely matched on these variables.

I visited each of the subject's mothers at home twice. For the first follow up study it was to gain consent for their further participation and to ask them to complete a behaviour report (Corbett, 1992). Five subjects refused consent. Newcastle Upon Tyne Social services refused consent for a child recently taken into care. A further 3 families had moved from the district. The 2 cases in the Edwards study where consent was refused were also followed up. One continued to refuse consent, but the carers of the other child agreed to participate. Two other subjects were unobtainable. Therefore the first follow up study consisted of 49 cases and 46 controls, of whom 30 cases and 31 controls were males and 19 cases and 15 controls were female.

For this second study I visited each family to report on our previous findings and to obtain permission for us to visit the children at school to test their IQ and to ask their teachers for a report on behaviour problems. All those contacted for my first study agreed to the participation of their children in this second follow up, with the exception of a female 'case' who completed the first part of the study but subsequently moved away from the area.

Consent to visit the children at school was also sought from the Heads of the schools to be visited. All but one head agreed. For children in this school IQ testing was carried out in the home, but we were unable to obtain teacher's reports for 4 cases and 3 controls.

A summary of those lost to follow up and the numbers of subjects completing each stage of the project is given below (Table 3). It can be seen that over 80% of the subjects were scored on all three measures and that the same number of cases and controls (6) did not complete the IQ test. For the Teacher's Report ten cases and nine controls were not rated. Therefore the number of cases lost to follow up did not differ significantly from the number of controls lost.

	CASES	%	CONTROLS	%
<u>ORIGINAL NUMBER</u>	<u>54</u>	<u>100%</u>	<u>52</u>	<u>100%</u>
<u>FIRST FOLLOW UP (1992)</u>				
NON CONSENT	-3		-3	
MOVED OUT OF AREA	-2		-1	
UNOBTAINABLE			-2	
<u>TOTAL PARENTS REPORT (CBCL)</u>	<u>49</u>	<u>91%</u>	<u>46</u>	<u>88%</u>
<u>THE PRESENT STUDY</u>				
MOVED OUT OF AREA	-1			
<u>TOTAL IQ TESTS</u>	<u>48</u>	<u>89%</u>	<u>46</u>	<u>88%</u>
HEAD TEACHER REFUSED CONSENT				
NO TEACHER REPORT	-4		-3	
<u>TOTAL TEACHER REPORT</u>	<u>44</u>	<u>81%</u>	<u>43</u>	<u>83%</u>

Table 3. RATE OF ATTRITION

In the present study out of a possible 106 subjects to follow up, 94 (89%) were tested for IQ and 87 (82%) were rated by teachers. The same number of cases and controls (6) did not complete the IQ test, and a further 4 cases and 3 controls were not rated by their teachers.

2.2.3. COMPARISON OF DEMOGRAPHIC FACTORS OF THOSE LOST TO FOLLOW UP.

The data collected in the earlier study by Edwards was used to provide an indication of the extent of deprivation in the study group. The measures used were whether the natural father was living in the family home, whether the principal wage earner was unemployed, car and home ownership. These data and sex of the child were reanalysed to determine whether those lost differed in any of the deprivation measures from those studied. The calculations were mainly based on the numbers of children who did not have a Teacher's Report Form since this measure showed the highest rate of attrition and numbers of those who did not complete other measures were too small to carry out a useful statistical analysis. The data and results can be seen in appendix 1. The cases and controls lost did not differ significantly in the deprivation measures recorded from those followed up.

No interventions were planned or carried out as a result of the findings of the Edwards (1987) study and interventions such as contact with social services were not recorded. But it should be noted that four of the cases had been removed from the home and two of these cases had been permanently adopted. None of the controls had been removed from home.

2.2.4 MATERIALS

The intelligence test was carried out using the Weschler Preschool and Primary Scale of Intelligence - Revised (WIPPSI-R, 1990). The full test was administered omitting the optional animal pegs and sentences subtests as these additional tests would have added at least another 20 minutes on top of the one to one and a half hours taken to administer the full test. In order to fit most unobtrusively into the school schedule which is broken up into one and a half hour units of time, it was important to ensure that testing time did not exceed this.

The age range for which the WIPPSI-R is applicable has been extended from 4-6 years of age to 3-7 years 3 months. Test-retest reliability is reported as .92 on the performance quotient (PIQ), .95 on the verbal quotient (VIQ) and .96 on the fullscale score (FSIQ). These scores remain stable over time (tested after a mean of 4 weeks (PIQ = .88, VIQ = .90, FSIQ = .91) (WIPPSI-R, 1990).

However, WIPPSI-R scores show less reliability in the older range of children tested (PIQ = .85, VIQ = .86, FSIQ = .90). This is explained in terms of a ceiling effect. It could be argued that since our subjects were in the older range (6-7 years), the test for older children (WISC-R) would be a preferable instrument to use. It is clear that, although all the children in the study were in the top age ranges, they were living in very deprived circumstances and could be expected to perform at levels

below a normal population sample. The advantages of using the WIPPSI-R are that it is the most recently updated Weschler test and can be expected to be more accurate in terms of age standardisation, especially as IQ changes over generations, but it also contains attractive and colourful items which make it easier to maintain the interest of young children.

The revised version has the added virtue that stronger scoring guidelines have been given on those items where some subjective judgement is required, so improving interscorer reliability. Interscorer reliability is reported as .96 on the comprehension test, .94 on the vocabulary test, .96 on similarities, .94 on mazes, and .88 on geometric design (WIPPSI-R Manual, 1990). To improve interscorer reliability for this study, Alice Thams and I were trained to use the WIPPSI-R together under the supervision of two clinical psychologists, Andrew Moon and Malcolm Bass. Both observed the other on trial tests and discussed a standardised approach to scoring trial tests particularly those where some judgements are made.

Construct validity of the WIPPSI was assessed by factor analysing subtest scores (WIPPSI-R, 1990). Two clear factors, performance and verbal, emerged. It has also been found that although the median for subtests correlate with performance scores (.55) and verbal scores (.65), nevertheless, the subtest measures a unique factor as part of the variance is not accounted for. This supports the view that interpretation of IQ scores should be at subtest and composite score level. However, the WIPPSI-R manual (1990) reports conflicting results for use of subtest analysis with no clear patterns emerging for either gifted or mentally deficient individuals. Therefore, analysis will rely on the PIQ, VIQ and FSIQ composite scores for which better reliability and validity data are available.

The work on predictive validity reported in the WIPPSI-R manual is mainly related to the former WIPPSI. Much of the content of the WIPPSI-R is similar, but easier and more difficult items have been added to ensure that the expanded age range is adequately tested. However, concurrent validity of the WIPPSI-R and WISC-R is reported as PIQ .75, VIQ .76, FSIQ .85 (WIPPSI-R Manual, 1990). A similar result was reported with the WIPPSI (PIQ .68, VIQ .8, FSIQ .84). As the correlations between WIPPSI-R and WISC-R scores are as strong as those between the WIPPSI and WISC-R, work on the predictive validity of the WIPPSI can also apply to the WIPPSI-R.

The WIPPSI was found to be a good predictor of WISC-R scores one year later (PIQ = .8, VIQ = .81, FSIQ = .91) (Rasbury, McCoy, and Perry, 1977). It is also a moderate predictor of reading scores (Maxwell, 1972. Feshback, Adelman and Fuller, 1977. White and Jacobs, 1979) and later school achievement as measured by the Metropolitan Achievement Test, Teacher assigned grades and a personal-social behaviour rating scale (Segerstrom, 1976).

It has been argued that in addition to IQ social competence and behavioural problems should be assessed (Scarr, 1981) especially as previous research on children who failed to thrive found lower levels of social competence and more behaviour problems (Hufton & Oates, 1977. Oates et al, 1984, 1985. Pollitt & Eichler, 1976. For this reason another assessment instrument was used, The Teacher's Report Form or TRF (Achenbach, 1991). The questionnaire contains two sections, competence items which report on the child's ability to integrate into peer group, family and school activities, and problem items which consists of the 118 problems most commonly reported to mental health professionals in the United States.

The problem items only were used as some of the competence items were not applicable to English children (i.e. 'Has your child repeated a grade?'). It was also found that competence scores discriminated cases and non-referred children less well in the lower age ranges than problem items (Achenbach, 1991, TRF Manual). Work in Holland using problem items only found a higher correlation between parents ratings using the Child Behaviour Checklist (Achenbach, 1991) and teacher's ratings using The Teacher's Report Form ($r=.57$) in learning disabled children than normal children ($r=.3$) suggesting that the problem items alone are an indicator of learning competency (Verhulst and Aakerhuis, 1989). In addition, the aim of the study is clearly to discover if children who failed to thrive had more behaviour problems, so no measure of positive achievement was required.

To test the validity of the behaviour rating scales a trial of 4,220 children was conducted, half of whom had been referred for mental health services, the referred children scored significantly higher on the Child Behaviour Checklist (CBCL) problem items than the non-referred children (Achenbach, 1991). Twenty-five of the problem items in the CBCL were replaced by items that were more easily measured or appropriate for teachers. The association between teachers ratings and referral for mental health services was tested on a sample of 2550 children half of whom had been referred. The

regression of raw scores for problem items on referral status accounted for 24-32% of the variance depending on age range. The problem items discriminated between referred and non-referred children at the $P < .01$ level.

Problem items identified in the United States have been found to have cross cultural validity. In a study carried out in Holland few differences were found in the prevalence rates of the 118 behavioural items in comparison with the United States (Achenbach, Verhulst, Baron and Aakerhuis, 1987).

The manual for the Teacher's Report Form (TRF) explains that 118 problem items were subjected to principal components analysis and were further divided into syndromes comprising problems which occur in association with one another. The syndromes are 'withdrawn', 'somatic complaints', 'anxious/depressed', 'attention', 'social problems', 'thought problems', 'delinquent behaviour' and 'aggression'. No problem occurs in more than one syndrome, although different problems may be grouped into the same syndrome at different ages. However, the syndromes do not account for all the scores at any age, as there are a number of items which do not load onto one specific factor in each age group and there are a number of low prevalence items. These would not be included in an analysis of syndrome scores.

Syndrome scores have also been factor analysed to produce two categories which distinguish internalising and an externalising syndromes. Internalising factors were 'withdrawn', 'somatic complaints' and 'anxious/depressed'. Externalising factors were 'aggression' and 'delinquent behaviour'. Three syndromes 'attention', 'thought problems' and 'social problems' do not load into either category and would be omitted in an analysis of externalising and internalising factors.

Because of these omissions, some of which may be particularly relevant to the group we are studying (ie. 'Eats things that aren't food' is a low prevalence item in girls aged 6-11), the most appropriate analysis would be carried out on total problem scores. It is also the case that positive correlations have been found between the externalising and internalising scores ($r=.52$), showing that children who tend to score high on one factor also score high on the other, even though they may be primarily an externaliser or internaliser (Achenbach, 1991). As the aim of the study is to discover if failure to thrive children score high, all high scores should initially be included.

The correlation coefficient for test-retest reliability of the TRF was found to be .92 when 44 8-9 years olds were retested after 15 days. Mean correlations for inter-teacher agreement range from .5 to .62 for total problem items, reaching the greatest level of agreement on scores for aggressive behaviour ($r = .68$).

Ratings from the TRF were scored using the Achenbach computer scoring programme. The programme involves a double data entry procedure. Any differences in entry of the same data are signalled and errors can be checked, so ensuring accurate entry of scores.

2.2.5. PROCEDURE

Initially each mother was visited at home in 1991 and given a letter personally to describe the first follow up study (Appendix 2). They consented, with the exceptions already detailed, to complete a CBCL according to the instructions at the top of the questionnaire.

The interviewer was blind to the status of the subjects as all subjects had been assigned a dummy identifier by Dr C. M. Wright. These identifiers were unscrambled by Dr R.F. Drewett after all data had been entered using a code which allocated case or control identifiers. In this way the researcher did not acquire information about the case or control status of subjects except by accident during contact with mothers, when revealing information came to light about some of the children. The interviewer noted which of these children were believed to be cases and controls to ensure that there was no direct contact with them in subsequent studies. The actual status (cases or controls) was never revealed to the interviewer.

The data on the CBCL were collected before the start of this project, over a three month period from September 1991 to November 1991. A report was completed by March 1992 as part of my third year dissertation (Corbett, 1992).

For this study, the parents were visited during the summer of 1992 to gain consent for their child to be visited at school in order to complete an IQ test and to acquire a Teacher's Report Form (Appendix 3 - letter personally delivered. Appendix 4 - consent form). None of those parents previously visited refused consent for follow up at school.

Visiting the parents at home may seem to be a time consuming exercise when we could have gained consent by post. Home visits were carried out for two reasons. The first reason was that personal contact gave the parents an opportunity to ask questions about the study to allay any concerns they had (i.e. confidentiality of reports and test results). Secondly, it was essential to follow up as many subjects as possible and the rate of return of postal communications is often low. There was genuine concern that parents would not receive a letter, be unable to read it, or be unwilling to reply.

The difficulty in receiving a letter is explained by the mobility of the population. There is a considerable amount of rebuilding and renovation of old housing stock in the area. This alongside many other reasons for moving, which were not recorded, resulted in a highly mobile population. At follow up in 1991, only 17% of the families were at the same address they had occupied in 1987.

When visited again a year later, in 1992, only 30% were at the address they had occupied a year earlier. Of the 70% who had moved, 40% had moved more than once, the most mobile moving as many as 6 times. Fortunately, all but one of the families moved within the locality and were retraceable through child health records or the school nurse.

Head Teachers were contacted by letter (Appendix 5) to ask permission for a psychologist to visit the children at school to carry out an IQ test and to ask the child's teacher to complete a report form. The letter was then followed by a telephone call to the head to find out if permission had been given and to arrange suitable facilities for testing (a quiet room) and a time to visit the school. In most instances a quiet room was available, but 3 tests were carried out in the only space available, an often noisy corridor. Some heads asked to meet the researcher and to see the signed consents from parents. One head refused permission despite a visit to the school, so IQ testing was carried out in the home and no teachers reports were obtained.

All contact with the teachers and children were made by a tester who was unaware of the status of the child. Where I suspected the status of a child after the home visit (a total of 12 children, some thought to be controls), that child was tested by the tester who had no previous contact with the family. This was easy to do as schools attended by the known children were allocated by me to the second tester.

Children were tested according to the instructions in the WIPPSI-R (1990) manual. At the end of the test they were given a letter for their parents to say that we had seen their child and to thank them for their help (Appendix 6) and the child was given a notelet to thank them (Appendix 7).

Teachers were given the Teacher's Report Form (TRF) with personal details for the child already completed and the competence section deleted. They were asked to complete only the problem items section according to the instructions at the top of the questionnaire and to note how long they had known the child on the front of the form. Where a teacher had not known the child for more than three months, they were asked to wait until a later date to complete the report so that no child received a report from a teacher who had not known them for three months or more. These late reports were mostly posted back to Durham University, although some were collected personally.

The TRF was scored using the Achenbach computer software. The WIPPSI-R was hand scored and results entered into an ascii file. All outcome data were analysed using SPSSX on the University of Durham mainframe computer.

2.2.6. OUTCOME MEASURES.

The outcome measures derived from the assessment instruments reviewed are:

WIPPSI-R	PERFORMANCE IQ (PIQ)
	VERBAL IQ (VIQ)
	FULL SCALE IQ (FSIQ)
TEACHER'S REPORT FORM	TOTAL PROBLEM SCORE. (TRF)

2.2.7 GROWTH MEASURES

In order to determine if there was an effect of chronicity or severity of growth retardation, or an effect of the age at which growth retardation was most severe, three measures derived from growth data were also prepared for analysis by Dr. Charlotte Wright. These were :

1. The first growth variable is height Z score at school entry (FUHTZ). This provides a measure of an individual's relative standing for height in comparison with the population mean for height of children at the same age. Z scores are standard deviations from the population mean ($Z=0$), and may therefore be positive or negative values. The height Z score has the advantage of being standardised for age so producing a value which is easily comparable between subjects and groups. Those children who showed long term growth delay could be expected to be short relative to their peers at follow up, which would reduce the overall mean for the group.

2. The second growth variable was a measure of failure to gain weight adequately during the first two years of life. This was calculated by using weight standard deviations from the population mean where $Z=0$. Predicted weight is estimated by converting the weight at 6 weeks to a standard deviation score with a positive or negative value and multiplying it by .65. This value allows for regression to the mean: the tendency over time for a large infant's standardised weight to fall towards the population mean and a small infant's standardised weight to rise towards the population mean. The predicted weight standard deviation when subtracted from the actual weight standard deviation score shows when there is an abnormally low weight gain, given the infant 6 week weight. Those with values less than 0 have failed to reach their predicted weight. Those with values less than -1.5 standard deviations below predicted weight are in the slowest growing 3%; those with values less than -1.26 are in the slowest 5%; and those with values less than -.89 are in the slowest 10% (Wright et al, 1994). This value (The Thrive Index) was calculated at the age at which the infants relative weight was at its lowest.

This Index measures a child against its own predicted weight. The values for regression to the mean, which are based on newly acquired population statistics, allow for the level of severity to be defined. The measure also produces a continuous variable which is a considerable improvement on the categorical measure of growth failure (e.g. <3rd centile).

Values for the Thrive Index at lowest centile point for the case group are shown in figure 1.

3. Age at lowest centile point for weight (ALOW) was the age at which a child's weight reached the greatest number of standard deviations below the weight predicted from weight at 6 weeks.

Figure 1.

LEAF UNIT = 0.10 N=53

CUMULATIVE CASES TO MEDIAN	THRIVE INDEX VALUES
2	-2 55
3	-2 3
4	-2 0
7	-1 998
8	-1 6
13	-1 54444
20	-1 3332222
(7)	-1 1100000
26	-0 9998
22	-0 7777766
15	-0 5555
11	-0 3
10	-0 1100
6	0 01
4	0 222
1	0
1	0 6

Figure 1. STEM AND LEAF OF THE THRIVE INDEX AT THE LOWEST CENTILE POINT

It can be seen that whilst the weight gain of 16 cases was below the slowest growing 5%, 17 cases show only mild growth failure, and weights of 6 cases are above the predicted value.

CHAPTER 3

3.1. ANALYSIS OF DATA AND RESULTS

3.3.1. BETWEEN GROUPS COMPARISON

The matching was groupwise since cases and controls were chosen from two clinics in two deprived areas of Newcastle Upon Tyne, Scotswood and West City. The groups were found to be closely matched on socio-economic variables. There was, however an element of casewise matching, as the next child to the case on the clinic list was chosen as the control. The matching criteria (age, sex and list position) were selected without a follow up study in mind and there was no reason to suppose that these matching criteria would predict the outcomes measured here.

Before analysing the data the pairings were, however, checked by correlating outcome variables to determine if there was an association between matching criteria and outcome, and so whether a matched pairs analysis or an independent samples analysis would be most appropriate. Using Pearson's product moment correlation coefficient the correlation for IQ was $r=.21$. As the distribution of total scores for the Teacher's Report Form was skewed, Spearman's Rho was used to calculate correlations; for these $Rho = 0.00$.

Neither of these correlations reached conventional levels of statistical significance ($P > 0.1$ for both correlations), so there was no advantage in using a paired analysis as the matching criteria were not significantly related to the outcome.

HEIGHT

These data were prepared for this study by Dr. Charlotte Wright and are shown with her kind permission.

The cases were smaller than controls during their second year of life (Edwards, 1987), and there was a small but significant difference in height at school entry. It can be seen in Table 4. that the mean height of cases at school entry remains below the population mean, whereas the mean heights of controls are above the population mean. The standard deviations of the distributions differ slightly, indicating a greater spread of height in controls. A comparison of heights between the two groups shows a statistically significant difference (using an unpaired t-test $t = 2.48$, $P = .015$).

	<u>CASES</u>	<u>CONTROLS</u>
	N = 51	N = 44
Z SCORE	-.19	.37
STDEV	.92	1.24

t = 2.48

p = .015

Table 4. COMPARISON OF HEIGHT Z SCORES AT SCHOOL ENTRY.

The mean height of cases is -.19 standard deviations from the population mean (Z=0), whereas the mean height for controls is .37 standard deviations above the population mean.

IQ

The Full Scale, performance and verbal IQ scores for cases and controls were compared using a t test for independent samples. The means and standard deviations and the results of statistical analysis are shown in Table 5.

Mean values for IQ in this study were approximately 1 standard deviation (15 points) below the population mean of 100 IQ points, placing the mean scores for both cases and controls in the low average range.

The cases showed a small (3.4 point) deficit in Full Scale IQ scores in comparison with controls (mean for cases 83.6, mean for controls 87.0). The Full Scale IQ scores were normally distributed in each group. The height and spread of the distributions were also similar (STDEV = 11.5 and 11.9 for cases and controls respectively). Although there was a trend towards lower Full Scale IQ scores in the case group, the difference between the case and control group was not statistically significant ($t = 1.42, p = .16$).

These results are shown graphically on p92 in Figure 2.

Mean values for Performance IQ were higher than Full Scale scores in both groups, but cases showed a 3 point deficit in comparison with controls (mean for cases 88, mean for controls 91). The variances were similar (STDEV = 11.3 for cases, STDEV = 12.4 for controls). The difference between groups was not significant ($t = 1.17$, $p = .243$).

Verbal IQ scores were lower than Full Scale IQ for both groups and a 3 point verbal IQ deficit was found in the case group in comparison with controls (mean cases 82, mean for controls 85). Variances were very similar (STDEV for cases = 13.0, STDEV for controls = 13.2). But again the difference was not statistically significant ($t = 1.08$, $p = .282$).

FULL SCALE IQ

	CONTROLS	CASES
N	46	48
MEAN	87	83.6
STDEV	11.5	11.9

$t = 1.42$ $p = .160$

PERFORMANCE IQ

	CONTROLS	CASES
N	46	48
MEAN	91	88
STDEV	11.3	12.4

$t = 1.17$ $p = .243$

VERBAL IQ

	CONTROLS	CASES
N	46	48
MEAN	85	82
STDEV	13.0	13.2

$t = 1.08$ $p = .282$

Table 5. WESCHLER PRE-SCHOOL AND PRIMARY SCALES OF INTELLIGENCE -REVISED; COMPARISON OF CASE AND CONTROL GROUPS

Mean and standard deviation scores are shown for Full Scale IQ, Performance IQ and Verbal IQ. Although cases score less, the difference between groups fails to reach conventional levels of statistical significance (all $p > .05$).

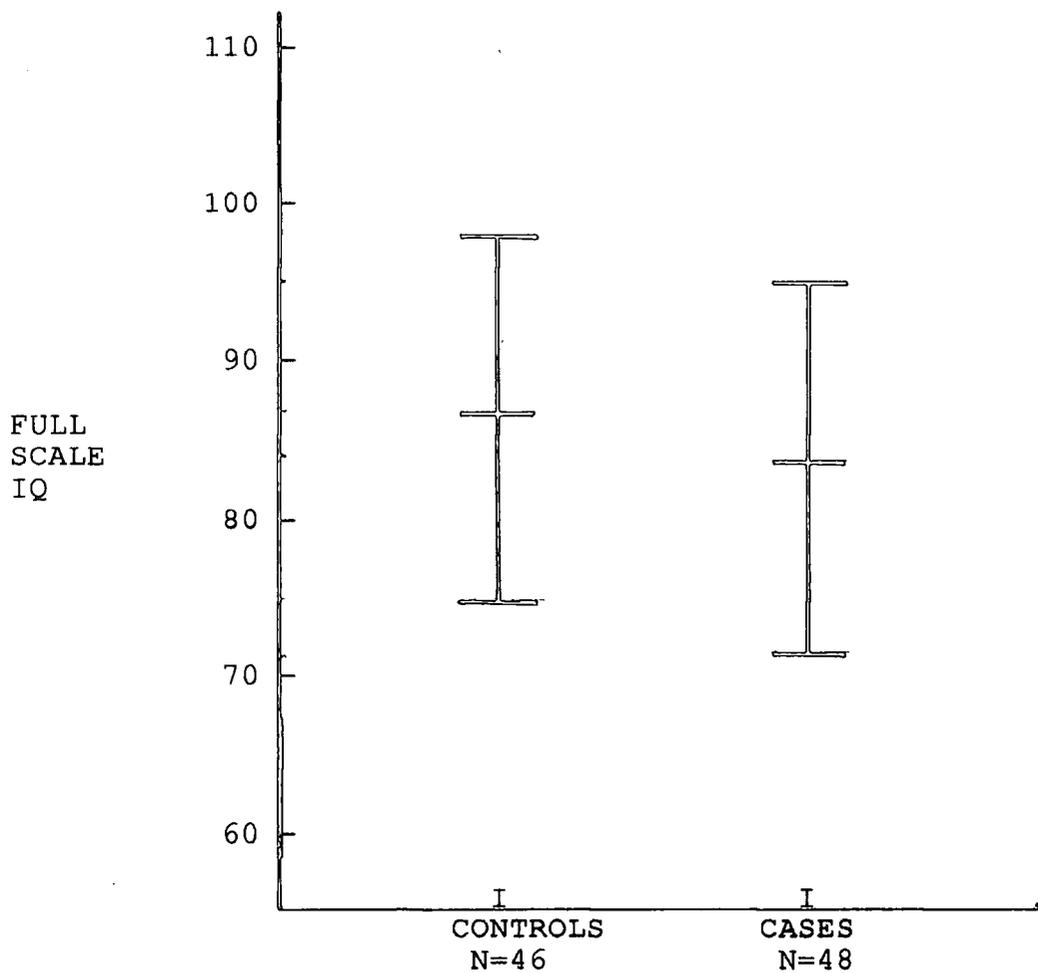


Figure 2. WESCHLER PRE-SCHOOL AND PRIMARY SCALES OF INTELLIGENCE-REVISED: A COMPARISON OF FULL SCALE SCORES OF CASE AND CONTROL GROUPS.

The case group Full Scale IQ scores show a 3.6 IQ point deficit in comparison with controls. The variance is similar (STDEV 11.5 and 11.9 for cases and controls respectively).

TEACHER'S REPORT FORMS

As the distribution for the Teacher's Report Form was skewed, a Mann Whitney U test was used to carry out a statistical comparison between the groups. These results are shown in Table 6.

The median value for problems was 23 problems for the cases, the standard interquartile range was 5 to 47. The median value for the controls was 14 problems with a standard interquartile range of 2 to 38 problems. However this trend did not reach statistical significance (Mann Whitney $U = 1.05$, $p = .296$). These data are shown graphically in figure 3. (p94).

TOTAL PROBLEMS

	CONTROLS	CASES
N	43	44
MEDIAN	14	23
MANN WHITNEY U = 1.05		p = .296

Table 6. TEACHER'S REPORT FORM: A COMPARISON OF CASE AND CONTROL GROUPS

The median value for total problems reported by teachers was 9 points higher in the case group. This difference was not statistically significant using Mann Whitney (U=1.05, P=.296).

Median + - SIR

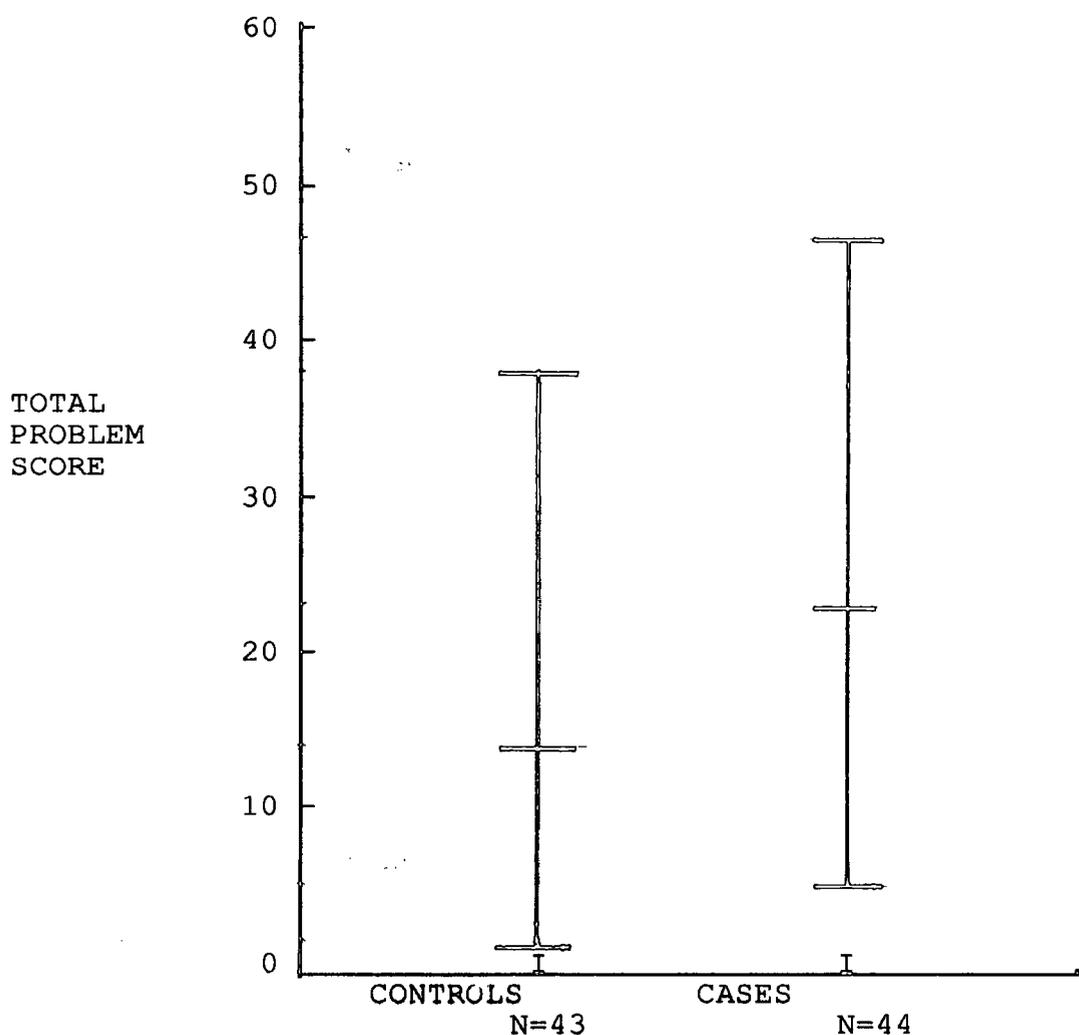


Figure 2. TEACHER'S REPORT FORM: COMPARISON OF CASE AND CONTROLS

The median value for total problems reported by teachers was 9 points higher in the case group and there were more high scores. The distribution of total problem scores was more skewed in the control group. This shows that cases not only scored more in the higher ranges, but also had more problems reported in the lower range of scores. However, the difference did not reach conventional levels of statistical significance ($p=.296$).

3.3.2. FURTHER ANALYSIS WITHIN THE CASE GROUP.

As weak trends towards lower IQ and more behaviour problems can be seen in the between groups comparison it was important to establish if these could be explained by the presence of a small group of severely affected children within the case group who shared certain growth characteristics. Therefore an analysis of the growth characteristics of the 'cases' is needed.

Three continuous variables which reflect the infant's salient growth characteristics were used to analyse the data within the case group using a regression analysis. The three variables, described earlier, were follow up height Z scores (FUHTZ) calculated from height at school entry, which provided an indication of chronicity of failure to thrive; age at the lowest centile position (ALOW); and the THRIVE INDEX (Wright et al, 1994), which provides a measure of severity of failure to thrive.

WIPPSI-R Full Scale IQ, Child Behaviour Checklist (using data from the first follow up study) and Teacher's Report Form were separately regressed on to the three independent variables FUHTZ, ALOW, and the THRIVE INDEX using the regression program in SPSSX.

FOLLOW UP HEIGHT Z SCORES (FUHTZ)

In the regression of Full Scale IQ scores on FUHTZ a very weak positive association was found between the two ($r=.12$). This accounted for 5% of the variance ($Rsq=.055$). Although a weak trend was found, IQ was not shown to be dependent upon FUHTZ; $F=2.496$ which was significant only at the $p=.0884$ level.

In the regression of Child Behaviour Checklist scores on FUHTZ, no association was found between the two ($r=.007$, $F=.131$, $p=.8775$).

Similarly Teacher's Report Form scores were not found to be associated with FUHTZ ($r=.044$, $F=1.14$, $p=.324$).

AGE AT LOWEST CENTILE POINT (ALOW)

In the regression of IQ on ALOW a weak negative association was found ($r=-.214$). Although this result is in the expected direction, the analysis of variance showed that the association was not statistically significant ($F=1.46$, $p=.237$).

CBCL scores also showed a negative association with ALOW ($r=-.276$), but again this failed to reach statistical significance ($F=.80$, $p=.45$).

Similarly, TRF scores were negatively correlated with ALOW ($r=-.18$) but the association was not statistically significant ($F=.847$, $p=.432$).

FUHTZ and ALOW when used as covariates showed a very weak association with outcome variables which failed to reach conventional levels of significance ($P>.1$ in all cases). Neither were the CBCL and TRF outcome variables significantly associated with any of the growth covariates ($P>.1$ for both outcome variables with all covariates).

THE THRIVE INDEX

The correlation coefficient for CBCL scores and the Thrive Index was expected to have a negative value so that the poorer the relative growth the more reported behaviour problems, but a positive value was found ($r=.12$) and this was not found to be statistically significant ($F=.39$, $p=.6779$). TRF scores and The Thrive Index were negatively associated ($r=-.11$), but the association was not statistically significant ($F=.85$, $p=.431$).

However, there was a significant association between WIPPSI-R FSIQ and Thrive Index at the lowest centile point. The results of this analysis are shown below (Table 7).

SOURCE	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARE	F	P
REGRESSION	1	637	637	4.95	.03
RESIDUAL	43	5528	128		
TOTAL	44	6165			

For the cognitive outcome the regression equation is:

$$\text{Predicted IQ} = 88.6 + 5.8(Z)$$

Table 6. REGRESSION OF IQ ON THRIVE INDEX AT THE LOWEST CENTILE POINT.

IQ correlated with The Thrive Index at the lowest centile point ($r=.32$). The regression equation predicts that for every 1 standard deviation weight fell below its predicted value there was a 5.8 fall in IQ points, where Z is the value of the lowest deviation from expected weight in standard deviation units, as measured by the Thrive Index at lowest centile point. A statistical significance was found between the covariates ($F=4.95$, $p=.03$). The residual standard deviation is 11.4.

3.3.3. SUMMARY OF RESULTS

The screening criterion used has selected a group of children who were significantly shorter 3 years later at school entry. The case group differed anthropometrically in the second year of life, and the mean height Z score of cases at school entry was below a standardised population mean ($Z = -.19$), whereas the mean height Z score of controls was above the mean ($Z = .37$); this difference is also statistically significant at the $p=.015$ level.

However, the case group did not differ significantly when compared with the control group in any of the psychological outcome measures. There was a mean difference of less than 4 Full Scale IQ points and this did not reach statistical significance ($p = .16$). The total problem score reported by teachers showed a weak trend towards more reported problems in the case group (median for cases 23 and for controls 14, $p = .297$).

In order to discover if the weak trend towards lower IQ and more behaviour problems could be explained by a severely affected subset of subjects, a within case group analysis was carried out using three continuous variables, height measured at school entry and converted to Z scores (FUHTZ), age at the lowest centile point (ALOW) and the number of standard deviations actual weight at the lowest centile point was below the predicted weight (THRIVE INDEX).

A significant association was found between between the THRIVE INDEX and the WIPPSI-R Full scale IQ ($r=.32$, $P=0.03$). For every standard deviation an infant was below predicted weight, their average IQ was 5.8 IQ points lower at follow up (Table 7.).

Differences between behaviour problems as rated by mothers or teachers failed to reach any of the conventional levels of significance with any of the covariates described. Neither a between groups nor within group analysis of behaviour problems reported by mothers or teachers has shown that behaviour problems, as measured by the rating scales used, are associated with growth retardation.



CHAPTER 4

4.1. DISCUSSION

The design features needed to provide a strong test of the outcomes of failure to thrive were outlined in the introduction. Firstly, since no common criterion has been agreed, a clearly defined criterion should be used to screen in cases. Secondly, the subjects should be selected by screening of the whole population, not preselected by referral processes. Thirdly, those screened in should be compared with a control group who are known to be growing normally contemporaneously, not just matched on demographic variables retrospectively. Fourthly, there should be minimal exclusions. At follow up every attempt should be made to minimise the rate of attrition, and those lost compared with those followed up to determine if those lost differ from the follow up group. Fifthly, testing should be carried out 'blind' to prevent results unwittingly being biased in favour of the hypothesis being tested. Finally, testing should be carried out at an age when IQ results have reasonably strong predictive validity.

All these design features were included in this study. For this reason this was a strong test of the outcomes based on the criterion for failure to thrive that was used. No statistically significant difference was found in behaviour problems and IQ between the case group identified in the primary care setting and the comparison group of normally growing children. It can be concluded that the screening criterion used for failure to thrive does not identify a group of infants who have impaired cognitive development or more behaviour problems in later childhood.

This conclusion needs to be treated with caution and does not mean that there is no difference to find, or that failure to thrive is not associated with lower IQ or more behaviour problems. Each of the design features will be reviewed in order to ascertain if this was indeed a strong test of outcomes of failure to thrive.

4.1.1.THE SCREENING CRITERION.

The growth criterion used was based on the idea that smallness alone was not cause for concern, as in a normal distribution of growth in a population a number of normal individuals will occupy the tails of the distribution. The salient factor in failure to thrive was a falling growth trajectory. In other words, the criterion identifies infants who had started out bigger, but their rate of weight gain had fallen relative to the population and they had shifted towards the left hand tail of the normal distribution for weight.

The problem with this criterion is that it does not take into account the phenomenon of regression to the mean in normal growth. Waddington (1957) argued that there is a species specific growth trajectory, the chreod. As a result of canalization, the process by which growth returns to the chreod, small infants, or those previously exposed to adverse environmental factors which are subsequently removed, will start to grow at a faster rate and will catch-up towards the population mean. This catch-up growth was monitored by Casey and Arnold (1985) in 10 infants with failure to thrive. For this reason it could be expected that the criterion used would underestimate the prevalence of failure to thrive in the community since it does not take into account the increase in rate of weight gain which should have occurred if infants had started out small. But, since the criterion focussed on a relative fall in rate of weight gain, it

also failed to exclude the effect of canalization in those large children whose rate of growth slows towards the mean.

Subsequently, a measure which is able to make this distinction between normal and abnormal fall in growth trajectory, the Thrive Index (Wright et al, 1994) was developed, and it was this that was used to reanalyse the original growth data for this study to provide a measure of severity of growth retardation.

It can be seen from the stem and leaf display (Figure 1. on page 85) that 6 infants in the case group remained above their predicted weight (values greater than 0) and, despite a fall in growth velocity, would now be classified as showing a normal regression towards the mean. Cases with values between 0 and -0.89 ($N=17$), although mildly growth retarded, would not now be classified as failing to thrive as their weight gain is above the 10th centile. Therefore, only 30 of the case group would now be classified as cases (below the 10th centile for weight gain) and 43% of the 'cases' would not now be regarded as such. The effect of this reduction in actual cases is a considerable weakening of a between samples comparison.

On the other hand it can be seen that the screening criterion used (Edwards, 1990) was highly sensitive and able to detect the most mildly growth retarded infants. It was also able to identify a group of infants from an annual cohort whose median value for weight gain was below the tenth centile (-1.1 standard deviations from predicted weight). Amongst these infants it was found that 16 were severely growth retarded and 23 were below the 10th centile for weight gain.

A positive outcome of this study has been the testing of a prospective criterion for diagnosis of failure to thrive. The development of such a criterion would be of considerable clinical value, enabling early intervention and prevention of long term adverse outcomes. The findings of this study show that the screening criterion was able to identify as cases those infants whose rate of weight gain was below the population mean and falling. But it also included those large infants showing a normal fall towards the population mean.

The Thrive Index, on the other hand appears to have the combined virtues of sensitivity and specificity. It is a sensitive measure as it provides a continuous variable which can be used to monitor weight gain for all infants. The rate of weight gain, in the case group, was quantified from those exceeding expected growth rate through to those infants with the poorest weight gain. It is specific since it clearly shows which infants started out large and have a normal fall in weight towards the mean.

4.1.2. THE PRIMARY CARE SETTING.

In the discussion of the screening criterion used it was pointed out that larger infants were identified as cases, whose weight gain was showing a normal fall towards the mean. It would seem obvious that something was amiss with the criterion when a prevalence of 20.9% of failure to thrive was identified in the community, as this was more than twice that of the most generous estimate found in previous studies (Mitchell et al, 1980, found a prevalence of 9.6%). However, this criticism ignores the characteristics of the area.

Not only was this study carried out in two of the poorest wards in the North (Townsend, Phillimore and Beattie, 1988) but amongst some of the poorest people in those wards (Edwards, 1987). This resulted in an expectation that prevalence of failure to thrive would be high and, although the screening criterion used was overinclusive, even the revised number of 'cases' represents 12.3% of the cohort, which is still larger than previous estimates of failure to thrive in the community.

Although Sameroff (1987) argued that risk factors which have an adverse effect on IQ have a cumulative effect, the graph of each cumulative risk score showed that the greatest effect on IQ occurred between the addition of 2-4 risk factors. After that the graph levels off. It could be reasonably argued that the addition of one more risk factor (ie. poor growth) amongst so many is only going to have a minimal effect on IQ. In order to quantify the effects of failure to thrive alone risk factors should be assessed and cases should be screened in from a broader range of social class.

It would be necessary to bear in mind, though, that the loss of 6 IQ points in a population of mean IQ 120 is relatively less vital than the loss of 6 IQ points when the population mean is only 85. It makes the difference at the lower end of the scale between coping with school and being unable to do so.

4.1.3. SELECTION OF A CONTROL GROUP.

A related point is that the controls for this study were not originally selected with a follow up of cognitive and behavioural outcomes in mind, so the cases and controls were matched on criteria which were not related to the outcome measures.

There were no significant differences between measures of deprivation. But there was a significant difference in per capita food expenditure in that controls reported spending a mean of £10.10 per head per week on food and cases £8.40 per week (Edwards, 1987). However, in such a deprived group there is no guarantee that this would remain a stable feature of the controls. Anecdotally, it can be reported that some of the families visited the second time had experienced changes of fortune which altered their circumstances noticeably from the first visit. Fathers had either been released from prison, or recently imprisoned. One family lived in a well furnished home in 1991, but by 1992 the contents of the home had been removed. There was also a high rate of broken relationships amongst the parents and the establishment of new relationships.

In the absence of data from the time of the first study in 1987 to the time of the follow up study, we cannot be sure that the difference in per capita expenditure remained the same, that control children were adequately fed or that they grew normally during the intervening period. We know only that they were slightly taller at school entry which suggests that if they experienced growth failure, it occurred later and not during a critical period for future growth. It has been shown that older children from this area of Newcastle are shorter than those in better off areas (Wright et al, 1992) which may testify to the stability of adverse environmental factors for growth. The problem then is that the control group's nutritional and growth history may also have been compromised since the first study.

4.1.3. EXCLUSIONS AND ATTRITION.

There were very few exclusions when screening in infants to the study. The exclusions were twins, Asian infants and children whose growth failure had known organic aetiology. Children with low birth weight were retained in the study. The minimal use of exclusionary criteria enables findings to be generalised to many types of growth failure.

Some exclusions were unavoidable: 24 subjects with incomplete weight data recorded. No information is available to indicate the reason for the incomplete data sets. Health professionals may have been unable to meet the demands of the added workload, or the mothers of the infants involved may have refused to attend clinics. The latter explanation might indicate that some of those excluded would have been counted as cases.

Analysis of the demographic characteristics of those lost at follow up show no significant difference between subjects lost and those followed up and it seems unlikely that the inclusion of those additional children who received no teacher rating ($n = 7$) would have produced a significant result in this outcome measure as the Teacher's Report Form correlated with IQ ($r = -.44$) and IQ failed to reach conventional levels of statistical significance ($P=.16$) in the between group comparison.

What is more important is that 2 of the most severely growth retarded infants, were lost to follow up. Of those children not followed up at all, 4 out of the 5 cases lost had rates of weight gain below the 5th centile and the 2 most severely affected were in the slowest growing 2%. This problem came to light when the original growth data was analysed using the Thrive Index. The Thrive Indices for those lost are given in Table 8 (p112).

CASES WHO DID NOT COMPLETE ANY STAGE:

	-2.58
	-2.34
	-1.64
	-1.35
(CBCL ONLY)	.21

CASES WITH NO TEACHERS REPORT:

	-1.23
	-1.04
	- .73
	- .55

Table 8. THRIVE INDEX OF THOSE CASES LOST TO FOLLOW-UP
(Number of standard deviations below predicted weight.)

Four out of the five children not receiving an IQ test had Thrive Indices below -1.26 which is the 5th centile for weight gain. Two of the most severely affected children in the case group were lost to follow up.

All the children who were not rated by their teachers were above the 5th centile for weight gain and 2 were above the 10th centile and would not now be included as cases according to the Thrive Index.

With such a low rate of attrition the overall effect of this on outcome measures in a between samples comparison is not strong, but the loss of the severely growth retarded subjects may affect a within group analysis where a large range of scores is required and where there may be too few severe cases to produce a strong result. Even so, a significant association was found between IQ and severity of growth retardation. This indicates that a further study which identifies more actual cases using the Thrive Index is needed to clarify the relationship between rate of growth and cognitive development.

4.1.4. TESTER AWARENESS.

Throughout this study several strategies were employed to prevent those gathering and analysing outcome data from knowing the status of the children tested. These included a third party scrambling the original identifiers, and unscrambling them for data analysis, and the use of a second psychologist to test those children whose identities became suspected by the first. No suspicion of 'caseness' was confirmed until final analysis was complete. This was made easier by the nature of the group as so many families had problems they could not all be 'cases'.

4.1.5. OUTCOME MEASURES.

The outcome measures were all reliable and well validated. The question is whether they are sensitive enough to detect a difference and, whether we are measuring the appropriate outcomes.

Justification for the use of the WIPPSI-R has been discussed. It can be seen now from the results that where a sensitive and specific measure of severity of growth failure is used, there is a significant association between the two measures.

The CBCL and TRF problem items showed no significant difference in the between groups analysis. The mean CBCL scores previously reported (Corbett, 1992) were identical (mean = 37.9). Neither was a significant association found between the CBCL and TRF with any of the growth variables. This may be due to a failure to report problems or lack of sensitivity of the measure.

Mothers were willing to report problems. No mother reported no problems. The mean was 38 out of a possible maximum score of 354 (no maximum score has ever been recorded, Achenbach, 1991) and there were 9 scores in excess of 70 points, the highest being 143. Teachers scored fewer problems (mean 24) also scoring 9 over 70, but with a lower maximum score of 118. Teachers seemed to be more reluctant to report problems as 11 scored 0. Of these 11, 4 children scored above the mean when rated by

their mothers, although only 1 rated as 0 by the teacher had an IQ more than 1 standard deviation below the total group mean (IQ of 63). The problem then was not that of persuading the mothers and teachers to report problems.

The second issue is whether the measure is sensitive enough to detect differences. These are not direct measures of behaviour problems. We must rely on other people's observations of behaviour and their judgements of what constitutes a behaviour problem. However, Achenbach (1991) points out that we normally rely on a mother's evaluation of her child in diagnosis of illness, in the treatment of behavioural problems and in evaluation of mental health. A teacher's report provides another rating for comparison. Achenbach (1991) reports that the CBCL and TRF ratings correctly differentiated between referred and non-referred populations of children.

In the Newcastle study it was found that there was a weak but significant relationship ($r = .22$) between teachers and mothers scores even though these data were collected one year apart. The correlation between the teacher's and mother's ratings suggests that in those items which apply to both home and school there was agreement. There is further support for agreement between the rating measures as teacher's scores correlated moderately with IQ ($r = -.44$, $P < .01$) and mother's ratings correlated with IQ ($r = -.3$). So teacher rating in particular was sensitive enough to predict IQ to some extent.

Despite significant levels of agreement across measures there was no statistically significant difference in the problems listed on the TRF (Median for cases and controls respectively 23 and 14, $p = .296$) and CBCL (Median for cases 33 and for controls 34, $p = .994$) between the case and control groups, and neither measure was found to be significantly associated with Follow up height, the Thrive Index, or age at the lowest centile point (all $p > .1$). It must be concluded that the association between behavioural problems and failure to thrive reported by the studies previously reviewed (ie. Hufton and Oates, 1977) could be explained as the outcome of a referral process that identifies small children from problem families as failing to thrive, and those with no signs of family dysfunction as just small.

The social competence measures on the CBCL and TRF were not thought to be appropriate for use with the subjects identified in Newcastle. As the problem items include items such as 'poor peer relations', then social competence would to some extent be measured by the problem items alone. On reflection, since both cases and controls could be expected to score high on problem items in such a poor environment, it would have been preferable to have some measure of positive achievement which the competence scales provide. This type of measure (The Vineland Maturity Scales) had some discriminatory power in the Oates (1984) study where failure to thrive cases not only had more behaviour problems, but were also less socially mature, than abused or control children at follow up.

4.2. CONCLUSION

Methodologically, this study has provided a strong test of cognitive and behavioural outcomes of failure to thrive as defined by a new growth screening measure (Edwards et al, 1990). No statistically significant difference was found in IQ and behaviour problems between cases identified by the criterion and a group of controls.

It is argued that the screening criterion, by focussing upon a fall in relative weight gain alone, did not allow for the tendency for growth of individuals to move towards the average, so including as cases those large infants whose weight showed a normal fall towards the mean.

In a further analysis using The Thrive Index (Wright et al, 1994), which was able to make this distinction between a normal fall in weight gain of large infants towards the mean and an abnormal fall in weight gain away from the mean, it was found that this measure correlated with IQ ($r=.32$) and the association between covariates was found to be statistically significant ($F=4.95$, $p=.03$).

A stronger test of the association between the The Thrive Index and IQ is now needed. This could be achieved by using the measure to screen a larger population with a larger range of Thrive Index and IQ scores.

CONTENTS OF APPENDICES.

APPENDIX 1
TABLES OF COMPARISONS OF DEMOGRAPHIC CHARACTERISTICS OF
CASES AND CONTROLS LOST TO FOLLOW UP WITH THOSE FOLLOWED
UP.....119-123

APPENDIX 2
INTRODUCTORY LETTER TO MOTHER.....124

APPENDIX 3
SECOND LETTER TO MOTHERS.....125

APPENDIX 4
CONSENT FORM.....126

APPENDIX 5
LETTER TO HEAD TEACHERS.....127

APPENDIX 6
LETTER TO MOTHERS.....128

APPENDIX 7
NOTELET TO CHILDREN.....129

APPENDIX 8
CONCLUDING LETTER TO HEADS.....130

APPENDIX 1

COMPARISON OF SUBJECTS LOST TO FOLLOW UP WITH SUBJECTS STUDIED

ANALYSIS OF CASES WHO DID NOT COMPLETE THE TEACHER'S REPORT FORM.

	N=53	
	BOYS	GIRLS
LOST	4	5
EXPECTED	(5.77)	(3.23)
FOLLOWED UP	30	14
EXPECTED	(27.6)	(15.4)

Table 1. SEX

K-1=1 Chi Square = 1.9 Critical value at .05 = 3.84.
Not significantly different.

In the following analysis of deprivation factors for cases data for 1 case lost and 2 cases followed up is missing. N = 50

	YES	NO
LOST	5	4
EXPECTED	(5.58)	(3.42)
FOLLOWED UP	26	15
EXPECTED	(25.42)	(15.58)

Table 2. NATURAL FATHER PRESENT

Chi Square = .19 Not significant.

	YES	NO
LOST	1	8
EXPECTED	(1.98)	(7.02)
FOLLOWED UP	10	31
EXPECTED	(9.02)	(32)

Table 3. EMPLOYED

Chi square = .76 Not significant

ANALYSIS OF CASES WHO DID NOT COMPLETE THE TEACHER'S
REPORT FORM continued.

	YES	NO
LOST	1	9
EXPECTED	(0.8)	(9.2)
FOLLOWED UP	3	37
EXPECTED	(3.2)	(36.8)

Table 4. OWNS CAR
Chi square = .0675 Not significant

	YES	NO
LOST	0	9
EXPECTED	(0.72)	(8.46)
FOLLOWED UP	3	38
EXPECTED	(3.28)	(38.54)

Table 5. OWNS HOME
Chi square = 0.92 Not significant

ANALYSIS OF CONTROLS WHO DID NOT COMPLETE THE TEACHER'S
REPORT FORM.

N = 52

LOST	5	4
EXPECTED	(5.7)	(3.46)
FOLLOWED UP	28	15
EXPECTED	(27.2)	(16.54)

Table 6. SEX
Chi square = .336 Not Significant

	YES	NO
LOST	5	3
EXPECTED	(5.8)	(2.2)
FOLLOWED UP	32	11
EXPECTED	(31)	(12)

Table 7. NATURAL FATHER PRESENT

One set of data about natural father is missing of a control lost to follow up, therefore N = 51. Chi square = .5 Not Significant.

	YES	NO
LOST	4	4
EXPECTED	(2.5)	(5.5)
FOLLOWED UP	12	32
EXPECTED	(13.5)	(30.5)

Table 8. EMPLOYED
Chi square = 1.5 Not Significant

ANALYSIS OF CONTROLS WHO DID NOT COMPLETE THE TEACHER'S
REPORT FORM continued.

	YES	NO
LOST	2	6
EXPECTED	(0.9)	(7)
FOLLOWED UP	4	40
EXPECTED	(5.1)	(38.9)

Table 9. OWN HOME
Chi square = 2.8 Not Significant

	YES	NO
LOST	2	6
EXPECTED	(1.8)	(6.1)
FOLLOWED UP	10	34
EXPECTED	(10.1)	(33.8)

Table 10. OWN CAR
Chi square = .02 Not Significant

APPENDIX 2 - Introductory letter to mother

The Parkin Project
3rd floor
Shieldfield Health Centre
Clarence Walk
Newcastle-Upon-Tyne

Dear

Research Project on the Growth of Children

I expect you will remember that some years ago a researcher, Adrian Edwards, visited you to discuss your child as part of a study of growth in childhood. Charlotte Wright may also have written to you more recently concerning this project.

I am working with Dr. Wright, and now that your little girl is a little older I would very much like to meet you to find out how you think she is getting on, and to discuss the possibility of gathering some other information that would be helpful to us.

My own children will be back at school in September, and I hope you won't mind if I visit you then. We are very grateful for your help with this study, which is providing us with very important information on the growth of children.

I look forward to meeting you,

Yours sincerely,

(Mrs) Sally Corbett.

APPENDIX 3 - Second letter to mothers

The Parkin Project
3rd floor
Shieldfield Health Centre
Clarence Walk
Newcastle-Upon-Tyne

Dear

Research Project on the Growth of Children

All the children Dr. Adrian Edwards saw a few years ago as part of our growth project are old enough to go to school now, and we are writing to ask whether you would be willing to agree to our collecting some information about progress at school.

We would like to ask one of their school teachers to let us have details of their work and behaviour at school. The information will be provided directly to us in written form, and will not be seen by anybody else in school or outside.

Then we would like to work with them on some simple tests involving puzzles, drawing etc. Most children find these tests quite enjoyable, and they are carefully designed to be suitable for all children, whether they normally do well at school or not.

If you would like to discuss this with us first, or would like more information, please phone or send us a message at the above address and we will get in touch with you. Otherwise, if you are happy for us to approach the school for this purpose we would be grateful if you would return the enclosed form, so that the school will know we have your permission.

With best wishes,

Sally Corbett
on behalf of the Parkin Project.

APPENDIX 4 - Consent Form

Name of parent or main carer

Address

Name of child

School (from September 1992)

I am willing for a member of the Parkin Project to approach my child's school teacher for information concerning their work and behaviour at school, and to carry out the necessary tests.

I understand that the information will be given confidentially, and will not be shown to anybody outside the project team.

Signed

Date

APPENDIX 5 - Letter to Head Teachers.

Dear,

I am writing to ask if you could allow us to visit your school to see some children who are enrolled in a research project. The project is a collaboration between this department and the Department of Child Health at Newcastle University, where Dr. Charlotte Wright is the paediatrician responsible. It has been approved by the Newcastle Health Authority Ethics Committee and by the City of Newcastle Director of Education, and we have written consent for each child's participation from their family.

We are studying a group of children some of whom gave grounds for medical concern during the first year of life. (number) of the group are at your school, and we would like to see them each individually for about an hour, to assess their cognitive development using the Wechsler Preschool and Primary Scale of Intelligence. This is carefully designed to be suitable for children of this age, whether they do well at school or not. We would also like to ask the child's teacher to rate the child's behaviour in class, using a standardised checklist. This takes only a few minutes to complete.

I am sorry to have to trouble you with this, but it is important that we see every child in the group, and we would do our best to cause you as little inconvenience as possible. I would be glad to discuss the project with you personally, and will telephone you in a few days time.

Yours sincerely,

Mrs. Sally Corbett

for project research group

Dr Charlotte Wright
Dr Robert Drewett
Mrs Alice Thams
Mrs Sally Corbett

The Parkin Project,
3rd floor,
Shieldfield Health Centre,
Clarence Walk,
Newcastle Upon Tyne
Tel: 091-261-0789

Dear

Research project of the growth of children

You may remember that I called earlier this year to ask if you would mind if I visited your child at school to carry out a few simple tests.

My colleague and I have now started to see all the children involved with the project at school and I am writing to let you know that we saw _____ at school today.

We are most grateful and would like to thank you for helping us once again.

With best wishes,

Sally Corbett

on behalf of the Parkin Project

APPENDIX 7 - Notelet to children

Dear

Thank-you for helping me today,

Love,

Sally.

APPENDIX 8 - Concluding letter to Heads.

Dear...

You may recall that my colleague Alice Thams and I visited your school last year to test the cognitive development of a number of children and to ask their teachers for a report on their behaviour. We are now in a position to give you some information on the outcome of our research.

The aim of the project was to discover if there was a relationship between poor growth in infancy and cognitive deficits and behaviour problems in later childhood. A recent study in London found that children whose growth was poor in infancy had substantial cognitive deficits in comparison with a control group. This finding would indicate that early growth failure may be a preventable cause of later learning difficulties.

The problem with identify early growth failure is that it is necessary to distinguish between those infants who are naturally small and those who are small as a result of poor growth. A new growth criterion which makes this distinction was used to screen all the infants born in Scotswood and West City between November 1985 and November 1986 whose mothers attended the two health clinics in the area. Fifty two children were identified as failing to grow adequately. It was these children and a control group of normally growing children of the same age and sex who were followed up last year.

It was found that poor growth was significantly related to cognitive deficits, placing some of the children in the range where they would have educational difficulties. Our paediatric colleagues are now devising ways of helping families of slow growing children.

This study has been important in supporting the use of a new screening criterion and identifying a pattern of growth which is associated with cognitive deficits. It would not have been possible to carry out this research without your help and that of your teaching staff. I would like to thank you for your help and hope you will convey my thanks to the teachers involved.

Yours Sincerely,

Mrs. S.S.Corbett

For project research group

Dr. Charlotte Wright
Dr. Robert Drewett
Mrs. Sally Corbett
Mrs. Alice Thams

REFERENCES

- Achenbach, T.M., Verhulst, F.C., Baron, G.D. and Aakerhuis, G.W. (1987) Epidemiologic comparisons of American and Dutch children: I. Behavioural and emotional problems and competencies reported by parents for ages 4-16. *Journal of the American Academy of Child Psychiatry*, 26, 317-325.
- Achenbach, T.M., Verhulst, F.C., Edelbrock, C., Baron, G.D. and Aakerhuis, G.W. (1987) Epidemiologic comparisons of American and Dutch children: II. Behavioural and emotional problems and competencies reported by teachers for ages 6-11. *Journal of the American Academy of Child Psychiatry*, 26, 326-332.
- Achenbach, T.M. (1991) Manual for the child behaviour checklist/4-18 and 1991 profile. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T.M. (1991) Manual for the teacher's report form and 1991 profile. Burlington, VT: University of Vermont, Department of Psychiatry.
- Aylward, G.P., Hatcher, R.P., Stripp, B., Gustafson, N.F., and Leavitt, L.A., (1985) Who goes and who stays: subject loss in a multicenter longitudinal follow-up study. *Journal of Developmental Pediatrics*, 6, 3-8.
- Ayoub, C.C. and Milner, J.S. (1985) Failure to thrive: parental indicators, types and outcomes. *Child Abuse and Neglect*, 9, 491-499.
- Batchelor, J., and Kerslake, A. (1990) Failure to find failure to thrive: The case for improving screening, prevention and treatment in primary care. Whiting and Birch Ltd.: London.
- Bayley, N. (1970) Development of mental abilities. In P.H. Mussen (Ed.) *Carmichael's Manual of Child Psychology*. (Vol. 1, pp. 1163-1209) New York: Wiley.
- Beck, S., Collins, L., Overholser, J. and Terry, K. (1984) A comparison of children who receive and who do not receive permission to participate in research. *Journal of abnormal child psychology*, 12, 573-580.
- Berglund, G. and Rabo, E. (1973) A long-term follow-up investigation of patients with hypertropic pyloric stenosis-with special reference to the physical and mental development. *Acta Paediatrica Scandinavia*, 62, 125-129.

Berwick, D.M. (1980) Non-organic failure to thrive. *Pediatric Reviews*, 1, 265-270.

Bithoney, W.G. (1986) Elevated Lead Levels in Children with Non-Organic Failure to Thrive. *Pediatrics*, 78, 891-895.

Bowlby, J. (1953) *Child care and the growth of love*. Penguin.

Casey, P.H. and Arnold, W.C. (1985) Compensatory growth in infants with severe failure to thrive. *Southern Medical Journal*, 78, 1057-1060.

Chase, H.P. and Martin, H.P. (1970) Undernutrition and child development. *The New England Journal of Medicine*, 282, 17, 933-939.

Chatoor, I., Schaefer, S., Dickson, L., and Egan, J. (1984) Non-organic failure to thrive: A developmental perspective. *Pediatric Annals*, 13, 829-843.

Cohen, J. (1992) A Power Primer. *Psychological Bulletin*, 112, 155-159.

Corbett, S.S. (1992) 'Behavioural outcomes of failure to thrive in infancy'. (University of Durham. Unpublished undergraduate dissertation).

Davies, D.P. and Williams, T. (1983) Is weighing babies in clinics worth while? *British Medical Journal*, 286, 860-863.

Dewey, K.G., Heinig, J., Nommsen, L.A., Peerson, J.M. and Lonnerdal, B. (1992) Growth of breast fed and formula fed infants from 0-18 months: The DARLING Study. *Pediatrics*, 89, 1035-1041.

Dobbing, J. and Sands, J. (1973) Quantitative growth and development of human brain. *Archives of disease in childhood*: 48, 757-767.

Dobbing, J. and Smart, J.L. (1974) Vulnerability of the developing brain and behaviour. *British Medical Bulletin*, 30, 164-168.

Dobbing, J. (1984) Infant nutrition and later achievement, *Nutrition Reviews*.: 42, 1-7.

Dowdney, L., Skuse, D., Hepinstall, E., Puckering, C., Zur-Szpiro, S. (1987) Growth retardation and developmental delay amongst inner city children. *Journal of Child Psychology and Psychiatry*, 28, 529-540.

Drotar, D. and Sturm, L., (1988) Prediction of intellectual development in young children with early histories of non-organic failure to thrive. *Journal of Paediatric Psychology*, 13, 2, 281-296.

Drotar, D., Malone, C.A., Devost, L., Brickell, C., Mantz-Clumpner, C., Negray, J., Wallace, M., Woychik, J., Wyatt, B., Eckerle, D., Bush, M., Finlon, M.A., El-Amin, D., Nowak, M., Satola, J., Pallotta, J., (1985) Early preventive intervention in failure to thrive: Methods and preliminary outcome. In *New directions in Failure to thrive: implications for research and practice*. Drotar (Eds.) New York: Plenum Press.

Drotar, D. (1990) Sampling issues in research with non-organic failure-to-thrive children. *Journal of Pediatric Psychology*, 15, 2, 255-272.

Edwards, A.G.K. (1987) 'Failure to thrive in Newcastle children'. (University of Newcastle Upon Tyne. Unpublished thesis).

Edwards, A.G.K., Halse, P.C., Parkin, J.M. and Waterston, A.J.R. (1990) Recognising failure to thrive in early childhood. *Archives of Disease in Childhood*, 65, 1263-1265.

Elmer, E., Gregg, G., and Ellison, P. (1969) Late results of the 'failure to thrive' syndrome. *Clinical Pediatrics*, 8, 10, 584-589.

Feshback, S., Adelman, H. and Fuller, W. (1977) Prediction of reading and related academic problems. *Journal of Educational Psychology*, 69, 299-308.

Field, M., (1984) Follow-up developmental status of infants hospitalized for non-organic failure to thrive. *Journal of Pediatric Psychology*, 9, 2, 241-256.

Frank, D.A. and Zeisel, S.H. (1988) Failure to thrive. *The Pediatric Clinics Of North America*, 35, 6, 1187-1206.

Glaser, H.H., Heagarty, M.C., Bullard, D.M. and Pivchik, E.C. (1968) Physical and psychological development of children with early failure to thrive. *Journal of Pediatrics*, 73, 5, 690-698.

Gunston, G. D., Burkimsher, D., Malan, H. and Sive, A.A. (1992) Reversible Shrinkage in Kwashiorkor: an MRI study. *Archives of disease in childhood*: 67, 1030-1032.

Hall, D.M.B. (1991) *Health for all children* (2nd edition) Oxford University Press. (Oxford, New York & Tokyo).

Homer, C. and Ludwig, S. (1981) Categorization of aetiology of failure to thrive. *American Journal of Disease in Childhood*, 135, 848-851.

Hufton, I.W. and Oates, R.K. (1977) Non-organic failure to thrive: A long-term follow-up. *Pediatrics*, 59, 73-77.

Largo, R.H., Graf, S., Kundu, S., Hunziker, U., and Molinari, L. (1990) Predicting developmental outcome at school age from infant tests of normal, at-risk and retarded infants. *Developmental Medicine and Child Neurology*, 32, 30-45.

Lloyd-Still, J.D., Hurwitz, I., Wolff, P. and Shwachman, M.D. (1974) Intellectual development after severe malnutrition in infancy. *Pediatrics*, 54, 306-311.

Lynch, A., Mitchell, L.B., Vincent, E.M., Trueman, M. and MacDonald, L. (1982) The McCarthy Scales of Children's Abilities: A normative study on English 4-year-olds. *British Journal of Educational Psychology*, 52, 133-143.

Maxwell, A.E. (1972) The WIPPSI: A marked discrepancy in the correlations of the subtests for good and poor readers. *British Journal of Mathematical Statistical Psychology*, 25, 283-291.

McCall R.B., (1979) The development of intellectual functioning in infancy and the prediction of later IQ. In J.D. Osofsky (Ed.) *Handbook of infant development*. New York: Wiley.

Mitchell, W.G., Gorrell, R.W. and Greenberg, R.A. (1980) Failure to thrive: A study in the primary care setting. *Epidemiology and follow-up. Pediatrics*, 65, 971-977.

Oates, R.K. and Hufton, I.W. (1977) The spectrum of failure to thrive and child abuse - a follow up study. *Child Abuse and Neglect*, 1, 119-24.

Oates, K., Peacock, A. and Forrest, D. (1984) Development in children following abuse and nonorganic failure to thrive. *American Journal of Diseases in Children*, 138, 764-767.

Oates, K., Peacock, A. and Forrest, D., (1985) Long term effects of nonorganic failure to thrive. *Pediatrics*, 75, 36-40.

Pollitt, E. and Eichler, A., (1976) Behavioural disturbances among failure to thrive children. *American Journal of Diseases in Children*, 130, 24-29.

Rasbury, W.C., McCoy, J.G. and Perry, N.W. (1977) Relationship of scores on WIPPSI and WISC-R at a one year interval. *Perceptual and Motor Skills*, 44, 695-698.

Rosenn, D.W., Stein-Loeb, L. and Bates, M. (1980) Differentiation of organic from non-organic failure to thrive syndrome in infancy. *Pediatrics*, 66, 698-704.

Rosenthal, R. (1966) *Experimenter Effects in Behavioural Research*. New York. Meredith Publishing Company.

Rutter, D.R. and Quine, L. (1990) Inequalities in pregnancy outcome: A review of psychsocial and behavioural mediators. *Social Science Medicine*, 30, 5, 553-568.

Rutter, M. (1972) *Maternal deprivation reassessed*. Penguin.

Sameroff, A.J., Seifer, R., Barocas, R., Zax, M. and Greenspan, S. (1987) Intelligence quotient scores of 4 year-old children: Social environment Risk Factors. *Pediatrics*, 79, 3, 343-350.

Scarr, S. (1981) *Testing For Children: Assessment and the many determinants of Intellectual Competence*. *American Psychologist*, 36, 1159-1166.

Schaffer, R. (1977) *Mothering*. Fontana.

Segerstrom, R.R. (1976) A study in the prediction of school readiness and school failure. *Dissertation Abstracts International*, 7050A.

Sherrod, K.B. O'Connor, S. and Vietze, P.M. (1984) Child Health and maltreatment. *Child Development*, 55, 1174-1183.

Sills, R.H. (1978) Failure to thrive: The role of clinical laboratory evaluation. *American Journal Of Disease in Childhood*, 132, 967-9.

Singer, L.T. and Fagan, J.F. (1984) Cognitive development in the failure to thrive infant: A three year longitudinal study. *Journal of Pediatric Psychology*, 9,3, 263-283.

Singer, L.T. (1986) Long-term hospitalization of failure to thrive infants: Developmental outcome at three years. *Child Abuse and Neglect*, 10, 479-486.

Skuse, D. H. (1985) Non-organic failure to thrive: a reappraisal. *Archives of Disease in Childhood*, 60, 173-178.

Skuse, D. (1987) The psychological consequences of being small. *Journal of Child Psychology and Psychiatry*, 28, 5, 641-650.

Snow, R.E. and Yalow, E. (1981) Education and Intelligence. In R.J.Sternberg (Ed.), Handbook of Human Intelligence. Cambridge:Cambridge University Press.

Spitz, R.A. (1945) Hospitalism. *Psychanalytic Studies of Children*, 1, 53-74.

Stern, D., (1977) The first relationship. Cambridge:Harvard University Press.

Taitz, L.S. and King, J.M.(1988) Growth patterns in child abuse. *Acta Paediatrica Scandinavia (Supplement)*, 343, 62-72.

Tanner, J.M., Whitehouse, R.H. and Takaishi, M. (1966) Standards from Birth to Maturity for Height, Weight, Height Velocity, and Weight Velocity: British Children, 1965. Part I. *Archives of Disease in childhood*, 1966, 41, 454-471.

Tanner, J.M., Whitehouse, R.H. and Takaishi, M. (1966) Standards from Birth to Maturity for Height, Weight, Height Velocity, and Weight Velocity: British Children, 1965. Part II. *Archives of Disease in childhood*, 1966, 41, 613-627.

Tanner, J.M. and Whitehouse, R.H. (1976) Clinical longitudinal standards for height, weight, height velocity and weight velocity and the stages of puberty. *Archives of Disease in Childhood*, 51, 170-179.

Tanner, J.M. (1989) *Foetus into Man: Physical Growth from Conception to Maturity*. 2nd Edition. Castlemead Publications. Ware.

Tizard, B., (1977) *Adoption:A second chance*. Open Books.

Townsend, P., Phillimore, P. and Beattie, A. (1988) *Health and deprivation: Inequality in the North*. Routledge.

Verhulst, F.C. and Aakerhuis, G.W. (1989) Agreement between parent's and teacher's ratings of behavioural/emotional problems of children aged 4-12. *Journal Of Child Psychology and Psychiatry*, 30, 123-136.

Waddington, C.H. (1957) *The strategy of the genes*. London:Allen and Unwin.

Waterlow, J.C. and Payne, P.R. (1975) The protein gap. *Nature*, 258, 113-117.

Weschler, D. (1974) Cognitive, Conative, and non-intellective intelligence. In D. Weschler, *Selected Papers of David Weschler*. New York:Academic Press.

