# Children's temperament and parenting practices in the first five years of life and cognitive, academic and adiposity outcomes in later childhood and adolescence

**Shiau Yun Chong**

**BSc (Hons), MPH**

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

August 2016

**School of Public Health**

**Faculty of Health Sciences**

**The University of Adelaide**

**Australia**

# Contents

# Abstract

The aims of this thesis are to examine the associations between children's temperament, parenting practices and three important public health outcomes: cognitive ability, academic achievement and adiposity. While there have been decades of psychological research in this area, this thesis takes a contemporary epidemiological approach to the topic and addresses some of the methodological limitations of past studies by using more advanced methods and longitudinal data from both Australia and the UK.

There are four papers in this thesis. The first study examined whether norms in the Revised Infant Temperament Questionnaire (RITQ) were suitable for use in a population sample of UK infants. The RITQ was normed on a small group of US infants in 1978 and has never been updated. Findings showed that 15% of children would be classified as temperamentally difficult using norms empirically derived from the UK infant data, compared to 24% using RITQ's norms, suggesting that potential misclassification of infant temperament occurred from using different norms. This study highlighted the need for more recent and culturally-specific temperament norms to categorise infant temperament. Temperament categories defined using the norms in this study were used in subsequent analyses in study 3 and 4.

Children's temperament may influence parenting, which is known to affect cognitive and academic outcomes. Most studies of temperament have not adequately accounted for parenting practices when examining the effect of temperament on cognitive and academic outcomes. To properly handle parenting practices at age 4 to 5 years as an intermediate variable, the second study used a marginal structural model to examine the

controlled direct effects of temperament at 2 to 3 years on cognitive and academic outcomes at 6 to 7 years in a nationally representative sample of Australian children. Temperament dimensions measured in this study were reactivity, approach, and persistence. This study found that the controlled direct effects of temperament on cognitive and academic outcomes was small. The largest effect (0.11 SD) was for persistence on verbal ability.

Since temperament had such a small influence on children's cognitive and academic outcomes, this thesis then examined parenting as the exposure, as parenting may have a greater influence on cognitive ability than temperament. The associations between parenting practices (warmth and control) and children's IQ in the UK cohort were explored in study 3. Temperament was contextualised as an effect-measure modifier, a variable that may modify the associations between warmth, control and IQ. Low parental warmth and high parental control at 24 to 47 months were associated with lower IQ at age 8 years. Effect sizes for warmth and control were 0.03 SD and 0.15 SD, respectively. Counter to the study's hypothesis, temperamentally easier children were more susceptible to the negative effects of low warmth and high control parenting than temperamentally difficult children.

Besides cognitive and academic outcomes, there is some evidence that parenting and temperament may influence children's adiposity. The fourth study focused on two more specific dimensions of parenting, namely parental feeding control and using food to soothe a child. The associations between feeding control, using food to soothe, and body mass index (BMI) and fat mass were explored in the UK cohort. Whether these

associations differed for children with different temperaments were examined using an analysis of effect-measure modification. Contrary to some studies, higher parental feeding control at age 42 to 65 months was associated with lower BMI at ages 7 and 15 years and fat mass at age 15 years. No association between using food to soothe (42 months) and BMI (7 and 15 years) or fat mass (15 years) were found.

Using two large, longitudinal observational studies from different countries, different temperament tools, and measures of temperament at different ages, the research in this thesis indicated that the effect sizes for temperament on cognitive, academic and adiposity outcomes are at best, very small. The differential susceptibility theory suggested by previous psychological studies, that temperamentally difficult children were more vulnerable to the detrimental effects of negative parenting, was not supported in the UK cohort and using contemporary epidemiological methods. It is recommended that future studies adjust rigorously for important confounders and use large, representative samples when examining the effect-measure modification by temperament of the associations between parenting and cognitive, academic and adiposity outcomes.

# Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed .....  Date 15-08-2016 .....

# Acknowledgements

I am very grateful to my supervisors – Associate Professor Lisa Smithers, Dr Catherine Chittleborough, Dr Tess Gregory, Professor John Lynch, and Dr Murthy Mittinty. Lisa, Cathy, and Tess have always provided me the support and encouragement throughout the candidature and I am very thankful for your patience and support all these years. Thank you John for your insightful comments on chapters and publications. You are always a great teacher and I have learned many things from being part of your team. Thank you Murthy for your advice and support to get me through the process of data analysis and interpretation.

I would also like to thank all participants and staff from the Avon Longitudinal Study of Parents and Children and the Longitudinal Study of Australian Children. I would like to acknowledge and thank the University of Adelaide and the Fraser Mustard Centre for funding this project. This thesis would not be possible without the data and funding sources.

A very special thank you goes to my parents, brothers, and sister. Thank you all for your unconditional love, patience, and giving me all the best opportunities in life. I would like to thank my PhD colleagues - Amelia Maika, Ting Xia, Edilene Lopes, Angela Gialamas, Gizachew Tessema, for sharing your expertise and lending me your ears. To my friends, especially Erica Sung, for your support all these years in my PhD journey. You listened to my frustration, talked through the issues with me, and I am so grateful to know you and your family.

# Publications contributing to this thesis

- Chong SY, Chittleborough CR, Gregory T, Lynch JW, Smithers LG. How many infants are temperamentally difficult? Comparing norms from the Revised Infant Temperament Questionnaire to a population sample of UK infants. *Infant Behav Dev*. 2015; 40: 20-28.

- Smithers LG, Chong SY, Chittleborough CR, Gregory T, Lynch JW. Authors respond to the commentary on Chong et al. "How many infants are temperamentally difficult?" (40 (2015)20-28). *Infant Behav Dev*. 2015; 41: 164-166.

- Chong SY, Chittleborough CR, Gregory T, Mittinty MN, Lynch JW, Smithers LG. Parenting practices at 24 to 47 months and IQ at age 8: Effect-measure modification by infant temperament. *PLoS One*. 2016; 11(3): e0152452.

- Chong SY, Chittleborough CR, Gregory T, Lynch JW, Mittinty MN, Smithers LG. Does temperament at ages 2-3 directly affect cognitive and academic outcomes at ages 6-7? Under review.

- Chong SY, Chittleborough CR, Gregory T, Lynch JW, Mittinty MN, Smithers LG. Associations of parental feeding control and use of food to soothe with adiposity in childhood and adolescence. Under review.

# Presentations arising from this thesis

- Chong SY, Chittleborough CR, Gregory T, Mittinty MN, Lynch JW, Smithers LG. Parenting influence on the association between temperament and IQ. Infant and Early Childhood Social and Emotional Wellbeing Conference. Canberra, Australia, October 2013.

- Chong SY, Chittleborough CR, Gregory T, Mittinty MN, Lynch JW, Smithers LG. Child temperament, parenting and IQ: Findings from a population-based cohort of parents and children. Public Health Association of Australia (South Australia Branch) State Population Health Conference. Adelaide, October 2014.

- Chong SY, Chittleborough CR, Gregory T, Mittinty MN, Lynch JW, Smithers LG. Influence of parenting and child temperament on Intelligence Quotient (IQ) at 8 years. The 20th Occasional Temperament Conference. Lincoln, Nebraska, USA, November 2014.

- Chong SY, Chittleborough CR, Gregory T, Mittinty MN, Lynch JW, Smithers LG. Using marginal structural models to estimate the direct effect of temperament at 2-3 years on receptive vocabulary and academic achievement at 6-7 years. 2015 Robinson Research Institute Symposium. Adelaide, November 2015.

- Chong SY, Chittleborough CR, Gregory T, Lynch JW, Mittinty MN, Smithers LG. Effects of parental feeding practices and temperament on BMI and fat mass from childhood through adolescence. 13th International Congress on Obesity, Vancouver, May 2016.

# Abbreviations

| | |
|---|---|
| ALSPAC | Avon Longitudinal Study of Parents and Children |
| ARS | Academic Rating Scale |
| ATP | Australian Temperament Project |
| BMI | Body Mass Index |
| CDE | Controlled direct effect |
| CDC | Centers for Disease Control and Prevention |
| CFQ | Child Feeding Questionnaire |
| CSE | Certificate of Secondary Education |
| DAG | Directed acyclic graph |
| DXA | Dual energy X-ray absorptiometry |
| EAS | Emotional, Activity, and Sociability |
| EMM | Effect-measure modification |
| FFQ | Food Frequency Questionnaire |
| IOTF | International Obesity Task Force |
| ITQ | Infant Temperament Questionnaire |
| IQ | Intelligence Quotient |
| IRSD | Index of Relative Socio-economic Disadvantage |
| K6 | Kessler 6 Scale |
| LAB-TAB | Laboratory Temperament Assessment Battery |
| LCA | latent class analysis |
| MAR | Missing at random |
| MCAR | Missing completely at random |
| MNAR | Missing not at random |
| MICE | Multiple imputation by chained equation |
| LSAC | Longitudinal Study of Australian Children |

| | |
|---|---|
| MSM | Marginal structural model |
| NYLS | New York Longitudinal Study |
| PPVT | Peabody Picture Vocabulary Test |
| RERI | Relative excess risk due to interaction |
| RITQ | Revised Infant Temperament Questionnaire |
| RR | Risk ratio |
| SEM | Structural equation model |
| STSC | Short Temperament Scale for Children |
| STSI | Short Temperament Scale for Infants |
| STST | Short Temperament Scale for Toddlers |
| TBAQ | Toddler Behavior Assessment Questionnaire |
| TTS | Toddler Temperament Scale |

# List of tables

# List of figures

# CHAPTER 1. Introduction

Child development is a complex process that is influenced by many factors. Bronfenbrenner's ecological model of human development (Figure 1.1) suggests that the interrelation between the child, family, school, neighbourhood and a wider context are important influences on child development.[1] This thesis examines child and family factors, specifically temperament and parenting, as two important determinants that may influence children's cognitive ability, academic outcomes and their adiposity.

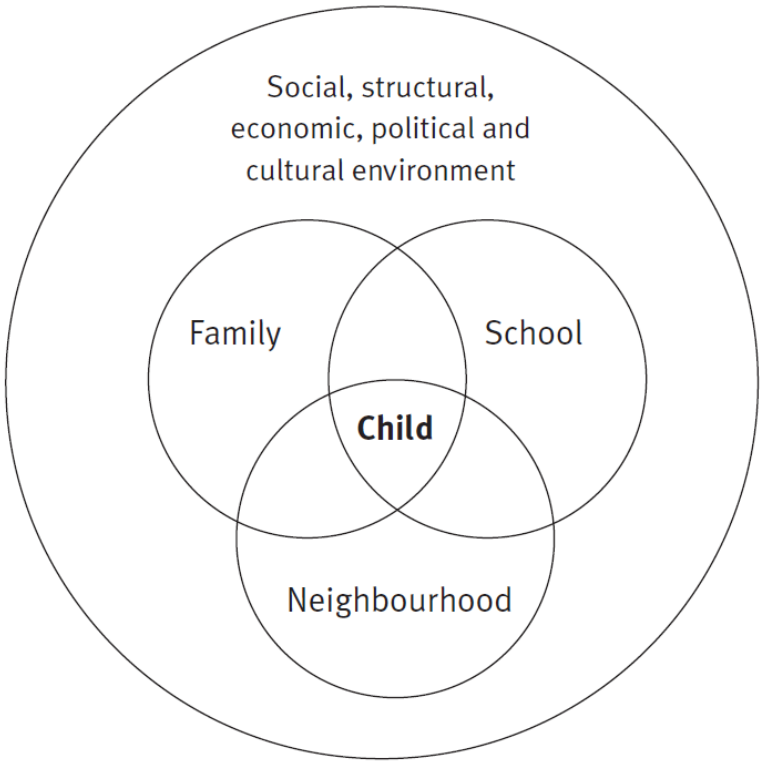


**Figure 1.1: Socio-ecological context shaping child development[2]**

## 1.1 Child temperament and parenting

Temperament refers to the observed style of an individual's behaviours.[3] In some literature, temperament is considered the same as 'personality'. However, temperament

and personality are conceptually different.[4,5] Temperament is often used to refer to the individual differences between infants, preschoolers, and school-age children. Personality often refers to individual differences between adults, as it includes a broader range of attributes than temperament, including a person's thoughts, skills, social values, morals, and beliefs.[3,6] Further definitions of temperament proposed by different researchers are included in Chapter 2.

Temperament influences how children respond to their environmental influences and has long-term effects on their adjustment in the family, at school, and in the broader environment.[7] For instance, children with temperament characterised by high levels of activity may experience difficulty in school because they are unable to sit for long periods to learn. A child's temperament can also affect their social interactions and relationships with parents, which in turn, shapes their development.[8] For example, a child with a difficult temperament may invoke negative parenting, which may exacerbate the child's temperament and increase the risk of behavioural problems. The interplay between temperament and parenting may influence children's cognitive and health outcomes.

While temperament is believed to be an individual's innate characteristics, studies have suggested that some aspects of temperament, such as persistence, are modifiable through interventions and environmental factors.[9,10] For instance, a cluster-randomized trial showed that children who received an intervention that aimed to develop their persistence, attention, and impulse control showed improvements in persistence and better academic skills than children in the control group.[11] There is also some evidence that parenting interventions are effective in improving parents' skills and behaviours and

have positive impacts on children's developmental outcomes.[12-15] For instance, a meta-analysis reported that young children who enrolled in home-visiting programmes had better cognitive and socioemotional outcomes than children in the control group.[16] Because temperament and parenting are important to children's development and both can be modified, they may be of interest for public health interventions.

## 1.2 Why study cognitive ability and overweight?

This thesis focuses on two main child health and development outcomes; cognitive ability and overweight or obesity. There are two main reasons for examining these outcomes. First, a review of literature found that there is a large body of research describing how temperament and parenting influences children's behavioural outcomes, such as externalising and internalising behaviours, but less is known about cognitive and weight outcomes. Few studies have examined both temperament and parenting as two factors that affect cognitive ability and overweight outcomes and the findings are inconsistent. This research aims to help inform current understanding of the effects of temperament and parenting on children's cognitive ability and overweight. A second reason for this focus is that cognitive ability and overweight are two important aspects of child development. Children with lower IQ have a higher risk of a range of negative life outcomes including lower occupational status, academic achievement and earnings in adulthood as well as an increased risk of mortality and morbidity.[17-20] At a population level, even a small increase in average IQ has important impacts on economic growth and productivity.[21] Childhood overweight is a major public health challenge. It is estimated that in 2013, more than 42 million children under the age of five were overweight or obese.[22] In Australia, about 17% of children aged 2 to 16 years are overweight and 6% are

obese.[23] Overweight and obesity from early life can have serious complications in later life. For instance, a systematic review of 15 observational studies found that higher Body Mass Index (BMI) from the age of 7 years is associated with increased risk of coronary heart disease in adulthood.[24] In addition, it has been shown that compared with healthy weight children, the adjusted odds ratio of being obese at age 13 years was 18.1 in children who were overweight in the first 7 years of life.[25]

## 1.3 Thesis aim and research questions

The general aim of this thesis is to examine the influence of temperament and parenting on children's cognitive ability and overweight. Four specific research questions to be addressed in this thesis are as follows:

1. Are infant temperament norms derived from the US suitable for use in a population sample of UK infants?
2. Is temperament at age 2 to 3 years directly associated with children's cognitive (verbal and non-verbal) and academic (literacy and numeracy proficiency) development at age 6 to 7 years?
3. Does infant temperament modify the effects of parenting practices (warmth and control) at age 24 to 47 months on IQ at age 8 years?
4. Is temperament at age 0 to 5 years associated with children's BMI at age 7 years? Are parental feeding practices (feeding control and use of food to soothe) at age 3.5 years associated with BMI z-scores and fat mass in childhood (7 years) and adolescence (15 years)? Does infant temperament modify the associations between feeding practices and BMI?

## 1.4 Thesis outline

This thesis is organised as follows:

Chapter 2 reviews the gaps in the research literature that motivated the four specific research questions. First, background information on temperament including history, definitions, models, stability, and contextual differences are detailed. The focus then moves to parenting, with a review of definitions and important aspects of parenting and how temperament may directly, indirectly and through "interaction" with parenting influence children's health and developmental outcomes. Finally, the current literature on temperament and parenting as individual and joint factors that influence cognitive ability and overweight is summarised.

Chapter 3 provides an overview of the two longitudinal studies that provided the datasets used in this thesis, *i.e.* the Avon Longitudinal Study of Parents and Children (ALSPAC) and the Longitudinal Study of Australian Children (LSAC). Detailed information on methodology used in this thesis, including directed acyclic graphs, latent class analysis, marginal structural models, multiple imputation, and effect-measure modification is provided.

Chapter 4, published as an academic paper in *Infant Behavior and Development* in 2015, focuses on addressing the first research question of this thesis. This paper compares temperament norms collected from a small clinical sample in the US and norms from the ALSPAC study conducted in the UK. While predominantly descriptive in nature, the findings from this chapter establish the need for updated norms for the Revised Infant

Temperament Questionnaire to categorise infant temperament. These updated norms are then used in addressing research questions 3 and 4 (Chapters 6 and 7, respectively).

Chapter 5, a submitted paper currently under review, addresses the second research question of this thesis. Using data from an Australia longitudinal study (LSAC), the controlled direct effects of temperament (dimensions of reactivity, approach, and persistence) at age 2 to 3 years on cognitive (verbal and non-verbal) and academic (literacy and numeracy) outcomes at age 6 to 7 years are examined after taking into account parenting practices at age 4 to 5 years as an intermediate variable.

Chapter 6, published as an academic paper in *PLoS One*, utilises secondary data analysis of the ALSPAC data to address the third research question. The influences of two important aspects of parenting practices, namely warmth and control, on cognitive ability (IQ) at age 8 years are explored. These effects are compared for children with different temperament profiles using an effect-measure modification analysis.

The final study is included in Chapter 7. This chapter consists of three parts. Part 1 of this chapter examines the associations between temperament and adiposity (BMI) using data from the ALSPAC and LSAC. Part 2 examines parental feeding practices (control and using food to soothe) on children's adiposity measured in childhood and adolescence (7 and 15 years) using data from the ALSPAC. The second part of Chapter 7 is a paper currently under review. Whether these effects differ for children with different temperament profiles is explored in Part 3 using an analysis of effect-measure modification.

Chapter 8 summarises key findings in this thesis, discusses strengths and limitations, potential areas requiring future research, implications and recommendations, and concluding remarks.

# CHAPTER 2. Literature review

This chapter reviews the concepts of temperament and parenting practices separately and then extending to review how temperament and parenting practices may individually and jointly influence children's cognitive ability and adiposity. Gaps relating to the design, analysis, and interpretation of the findings from previous studies are identified.

## 2.1 Temperament

### 2.1.1 History of temperament

The idea of 'temperament' dates back to the theory of four temperaments described by the ancient Greeks over 2,000 years ago.[6] According to the theory, a person's mood and behaviours resulted from the balance of four fluids in the body, which were the blood, black bile, yellow bile, and phlegm. A fourfold temperament typology (sanguine, phlegmatic, melancholic, and choleric) that represents different personality types emerged according to this theory. The sanguine person is **positive and outgoing** and is seen as having a predominance of blood; the phlegmatic person is **calm and patient** and is seen as having a predominance of phlegm; the melancholic person is **moody** with a tendency to fear and sadness and is seen as having predominantly black bile; the choleric person is **irritable and prone to aggression** and is seen as having predominantly yellow bile. This field has a come a long way since the four humours. The Greeks' view of temperament as an individual's characteristics derived from the biological and emotional processes is consistent with contemporary conceptualizations of temperament.[26]

During the 20th century, many theories of temperament emerged. One of the most well-known empirical studies of temperament is the New York Longitudinal Study (NYLS),

which was started in the 1950s by two psychologists, Alexander Thomas and Stella Chess.[27] Thomas and Chess first conducted a pilot study on 22 children aged 3 to 6 months.[28] They interviewed parents to gather information about the children's behavioural characteristics. Content analysis using data collected from the interviews was performed by Herbert Birch, a colleague of Thomas and Chess. Based on this information, nine temperament dimensions were identified: activity, rhythmicity, adaptability, approach, intensity, mood, distractibility, persistence, and threshold. Infants and children were located along a continuum from high to low on each dimension.

The researchers proceeded with a long-term study of temperament that involved a total of 141 children from 85 families and followed up the children for more than a decade. The families involved in the NYLS were predominantly middle- or upper-middle class, with 40% of the mothers and 60% of the fathers having both college education and postgraduate degrees, and less than 10% with no college education at all. In addition to the 141 children from the US, Thomas and Chess also obtained data from different groups of children, including a sample of 95 children of working class Puerto Rican parents who lived in low-income public housing projects,[29] a group of 68 children born with low birth weight (1,000 to 1,750 grams),[30] a group of 52 children of low intellectual ability,[31] and a special population of 243 children with congenital rubella.[32] The same set of temperament characteristics were found in these different populations groups. The NYLS marks an important beginning to the study of temperament. Since then, the nine temperament dimensions found in the NYLS have been widely used to examine children's temperament.

### 2.1.2 Definitions and models of temperament

Table 2.1 summarises some of the most common definitions, dimensions and models proposed by different temperament researchers. These temperament models are used for infants and young children aged less than 10 years, which is the focus age range for studies of temperament applied in this thesis.

**Table 2.1 Definitions, dimensions, and questionnaires used for temperament**

| Authors (Year) | Definition of temperament | Dimensions | Questionnaires |
|---|---|---|---|
| Alexander Thomas & Stella Chess (1963) / William Carey & Sean McDevitt (1970) | "Behavioural style relating to the *how* rather than the *what* (abilities and content) or the *why* (motivations) of behaviour"[33] (p. 9) | 9 dimensions:<br>1. Activity<br>2. Rhythmicity<br>3. Approach<br>4. Adaptability<br>5. Intensity<br>6. Mood<br>7. Persistence<br>8. Distractibility<br>9. Threshold | 1. Infant Temperament Questionnaire (4-8 months)<br>2. Revised Infant Temperament Questionnaire (4-8 months)<br>3. Toddler Temperament Scale (1-3 years)<br>4. Behavioral Style Questionnaire (3-7 years) |
| Mary Rothbart (1977) | "Constitutionally based individual differences in reactivity and self-regulation, where 'constitutional' referred to the biological bases of temperament, influenced over time by heredity, maturation, and experience"[6] (p. 100) | 3 main dimensions:<br>1. Surgency/ Extraversion<br>2. Negative affectivity<br>3. Effortful control | 1. Infant Behavior Questionnaire (3-12 months)<br>2. Early Childhood Behavior Questionnaire (18-36 months)<br>3. Children's Behavior Questionnaire (3-7 years) |
| Arnold Buss & Robert Plomin (1975) | Personality traits with an inherited component, observable and relatively stable across time and situation[34] | 4 dimensions:<br>1. Emotionality<br>2. Activity<br>3. Sociability<br>4. Shyness | 1. Emotional Activity and Sociability-III Questionnaire (toddlers and preschoolers)<br>2. The EAS and the NYLS's dimensions merged to form the Colorado Childhood Temperament Inventory (1-6 years) |
| Hill Goldsmith (1987) | "Individual differences in the probability of experiencing and expressing the primary | 5 dimensions:<br>1. Activity<br>2. Pleasure/ positive affect | 1. Toddler Behavior Assessment Questionnaire (16-36 months) |

| Authors (Year) | Definition of temperament | Dimensions | Questionnaires |
|---|---|---|---|
| | emotions and arousal"[35] (*p.*510) | 3. Social fearfulness<br>4. Anger proneness<br>5. Interest/ Persistence | |
| Jerome Kagan (1984) | Focused on behavioural inhibition, which is the hesitancy in one's approach to new or unfamiliar or unpredictability[36] | 3 dimensions:<br>1. Social novelty<br>2. Situational novelty<br>3. Physical challenges | 1. Behavioral Inhibition Questionnaire (3-5 years) |
| Ann Sanson (1986) | "Constitutionally based individual differences in behavioural style relating to affect, activity, and attention that are visible from early childhood"[37] (*p.*227) | 3 dimensions:<br>1. Reactivity<br>2. Approach<br>3. Persistence | 1. Short Temperament Scale for Infants (0-1 year)<br>2. Short Temperament Scale for Toddlers (2-3 years)<br>3. Short Temperament Scale for Children (4-5 years) |

#### 2.1.2.1 *Alexander Thomas and Stella Chess*

Temperament is the 'style' of behaviour, that is, *how* children react in a situation instead of *what* (abilities and content) and *why* (motivations) they do.[27] For instance, when showing their dislikes to a certain food, one child cries intensely, but another child may quietly and gently turn their head away.

As described earlier, Thomas and Chess identified nine temperament dimensions from the NYLS. Table 2.2 summarises the definitions and examples of items for each dimension. In addition to receiving a score on each of the nine dimensions, Thomas and

Chess categorised children into having an easy, difficult, or slow-to-warm-up temperament based on their scores on five of the nine dimensions, *i.e.* rhythmicity, approach, adaptability, intensity, and mood. A difficult child was characterised by having negative mood, irregular daily functions, high intensity in response (very intense when crying, screaming and laughing), being slow to adapt to new environments, and withdrawn when exposed to a new object or person. An easy child was characterised by having positive mood, high regularity in daily functions, low or moderate intensity of response, easily adapting to a new environment, and positive approach to new situations. The slow-to-warm-up child was characterised by being slightly negative in mood, slower to adapt, withdrawn on their first exposure to a new environment, and low or moderate in intensity. Of the 141 children in the NYLS, 40% of the children were categorised as easy, 10% as difficult, and 15% as slow-to-warm-up. The remaining 35% of children who did not fit into these three categories were classified as intermediate/other. Thomas and Chess' "difficult" concept was associated with later development of behavioural problems.[38,39] This concept has been commonly used in paediatric practices and temperament research.[40]

Using information collected by Thomas, Chess and colleagues from interviews with parents in the NYLS, Carey and McDevitt developed the Infant Temperament Questionnaire to allow assessment of infant temperament in a faster manner.[41] The questionnaire consisted of 70 items and norms were standardised in a group of US infants aged 4 to 8 months (n=101).[41] The questionnaire was revised in 1978 and the number of items in the Revised Infant Temperament Questionnaire (RITQ) increased to 95.[42] Norms in the RITQ were standardised based on 203 US infants from three local paediatric

clinics.[42] Both the “difficult” temperament as conceptualised by Thomas and Chess, and the RITQ developed by Carey and McDevitt have been widely used in paediatric clinical practice and temperament research.[40]

**Table 2.2 Definitions and sample items of the nine temperament dimensions proposed by Thomas and Chess[27]**

| Dimensions | Definition | Sample items from the Revised Infant Temperament Questionnaire (for infants 3 to 6 months) |
|---|---|---|
| Activity | The level and extent of physical movement | Moves a lot during nappy change and dressing. |
| Rhythmicity | The degree of regularity of daily functioning including eating, sleeping | Eats about the same amount of solid food from day to day. |
| Adaptability | The ease with which a child adapts to changes in environment | Accepts regular procedures at any time without protest. |
| Approach/withdrawal | The response to a new object or person | Initial reaction to strangers is acceptance. |
| Threshold | The degree of sensitivity to external stimuli including noise, heat | Takes any food offered without seeming to notice the difference. |
| Intensity | The energy level of a response | Takes feeding quietly with mild expressions of likes and dislikes. |
| Mood | The amount of pleasant, joyful behaviour as contrasted with unpleasant, unfriendly behaviour | Cries when left to play alone. |
| Distractibility | The degree to which a child can be distracted or comforted when needed | Continues to cry in spite of several minutes of soothing. |
| Attention & persistence | The amount of time devoted to an activity, and the effect of distraction on the activity | Plays with a toy for less than a minute and then looks for another toy or activity. |

*Note.* Items are scored on a 6-point scale from 'almost never' to 'almost always'.

#### 2.1.2.2 *Mary Rothbart*

In the late 1970s, Rothbart, a professor of psychology, together with her colleagues reviewed the temperament model proposed by Thomas and Chess.[5] They found insufficient consistency in the nine temperament dimensions and they questioned the temperament definition proposed by Thomas and Chess. A more contemporary

temperament definition was developed by Rothbart, with an emphasis on attentional and neurobiological development. Rothbart and colleagues defined temperament as "constitutionally based individual differences in reactivity and self-regulation, in the domains of affect, activity, and attention" (*p.*100).[6] The term "constitutional" was used to refer to a biological basis of temperament, which was also influenced by environment and experiences. Reactivity refer to a person's responsiveness to the internal and external environment such as the arousal of fear and anger, whereas self-regulation refer to processes such as effortful control that function to modulate reactivity.

Rothbart examined the development of temperament dimensions from early infancy to later childhood and found three main dimensions appeared consistently across the period. The three dimensions are surgency (extraversion), negative emotionality, and a dimension labelled 'regulatory capacity' in infants and 'effortful control' in older individuals.[5,6] The three main dimensions are further divided into multiple sub-dimensions (Table 2.3). Surgency is the combined disposition of positive emotion, rapid approach and high activity level.[5] Negative emotionality, also known as distress proneness, combines a disposition toward fear, anger/frustration, sadness, irritability, anxiety, guilt, and discomfort.[5,6] Effortful control is the major form of self-regulation that enables an individual to voluntarily control emotions, behaviours and attention.[5] Effortful control can be divided into attention control, inhibitory control, low-intensity pleasure, and perceptual sensitivity. Effortful control begins to develop from the age of 10 months and undergoes rapid development between the ages of 2 and 7 years.[43] Effortful control is important for child development as it regulates positive and negative emotions, and constrains the child's actions to better fit their values and goals. These three main

temperament dimensions were also found in studies performed in the United States, China, and Japan.[44]

**Table 2.3 The main dimensions and sub-dimensions of temperament proposed by Rothbart[5]**

| Dimensions | Sub-dimension | Definition | Example of items from Children's Behavior Questionnaire (for children 3 to 7 years) |
|---|---|---|---|
| Surgency/ extraversion | Activity level | Level (rate and extent) of physical movement. | Seems always in a big hurry to get from one place to another. |
| | Approach/Positive Anticipation | Amount of excitement and positive anticipation for expected pleasurable activities. | Gets very enthusiastic about the things s/he does. |
| | High Intensity Pleasure | Amount of pleasure or enjoyment related to situations involving high intensity and novelty. | Likes going down high slides or other adventurous activities. |
| | Shyness (negative loading) | Slow or inhibited approach in situations involving novelty or uncertainty. | Seems to be at ease with almost any person. |
| | Impulsivity | Speed of response initiation. | Often rushes into new situations. |
| | Smiling and Laughter | Amount of positive affect in response | Smiles a lot at people s/he likes. |
| Negative emotionality | Discomfort | Amount of negative affect related to sensory qualities of stimulation, including intensity, light, movement, sound, texture. | Is quite upset by a little cut or bruise. |
| | Anger/Frustration | Amount of negative affect related to interruption of ongoing tasks or goal blocking. | Gets angry when told s/he has to go to bed. |
| | Falling Reactivity/Soothability (negative loading) | Rate of recovery from distress and excitement. | Is easy to soothe when s/he is upset. |
| | Fear | Amount of negative affect, including unease, worry or | Is afraid of the dark. |

| Dimensions | Sub-dimension | Definition | Example of items from Children's Behavior Questionnaire (for children 3 to 7 years) |
|---|---|---|---|
| | | nervousness related to pain, distress or potentially threatening situations. | |
| | Sadness | Amount of negative affect related to exposure to suffering, disappointment and object loss. | Cries sadly when a favourite toy gets lost or broken. |
| Effortful control | Inhibitory control | The capacity to plan and to suppress inappropriate approach responses under instructions or in novel or uncertain situations. | Can wait before entering into new activities if s/he is asked to. |
| | Attentional focusing | Tendency to maintain attentional focus upon task-related channels. | When practicing an activity, has a hard time keeping her/his mind on it. |
| | Low Intensity Pleasure | Amount of pleasure or enjoyment related to situations involving low stimulus intensity and novelty. | Enjoys just being talked to. |
| | Perceptual Sensitivity | Detection of slight, low-intensity stimuli from the external environment. | Is quickly aware of some new item in the living room. |

#### 2.1.2.3 *Arnold Buss and Robert Plomin*

According to Buss and Plomin,[45] temperament dimensions are genetically-based, observable and relatively stable across time and situation.[46] Buss and Plomin identified four temperament dimensions: emotionality, activity, sociability, and shyness. Emotionality refers to how easily a child gets distressed and upset. Activity refers to the speed and intensity of speech, bodily movements and energetic behaviours such as

running. Sociability refers to the child's preference to be with others, for example, the child prefers playing with others rather than alone. Shyness refers to the feelings of tension and distress in social situations, for instance, requiring a long time to warm up to strangers. Because Buss and Plomin's temperament model emphasised the "inherited" criteria of temperament, their model has been used in a number of behavioural genetic studies.[47,48] Such studies in twins and adoptive siblings have enhanced the understanding of the influence of genetics and environment on temperament.[49]

#### 2.1.2.4 *Hill Goldsmith*

Goldsmith defined temperament as the individual's emotional characteristics, which encompass motor activity, anger, fearfulness, pleasure and interest.[35] Goldsmith developed the Toddler Behavior Assessment Questionnaire (TBAQ) for use with children aged 16 to 36 months.[50] Goldsmith and colleagues also developed the Laboratory Temperament Assessment Battery (LAB-TAB) to examine children's emotional responses through direct observations. For instance, in the Box Empty task, the child is given a wrapped box to open, with the examiner pretending that there is an appealing toy was inside. The child is left alone with the box and soon discovers that the box is empty. This test examines the emotional sadness, anger or negative emotionality of the child in response to the failed expectations.[51] The LAB-TAB has been widely used for observational assessment of temperament in small-scale experimental studies. However, the administration of the LAB-TAB is time consuming (about 40 minutes), requires special training, and is relatively expensive compared to temperament questionnaires.[51] Therefore, its applicability in large-scale population based studies is limited.

#### 2.1.2.5 *Jerome Kagan*

Kagan focused on "behavioural inhibition", which is the tendency toward high or low reactivity to novel or unfamiliar situations, in which children express fear and tendency to withdraw in the face of stressful, novel situations.[52] Kagan and colleagues observed 4-month-old infants' responses to unfamiliar stimuli in a laboratory setting. They found that 20% could be classified as high reactive (characterised by vigorous motor activity and frequent crying), 40% classified as low reactive (minimal motor activity and crying), 25% classified as distressed (minimal motor activity but cried frequently), and 10% classified as aroused (vigorous motor activity but little crying).[53] The high reactive group was most likely to avoid unfamiliar events at young ages and be emotionally subdued and wary of new situations at the later ages.[53] By comparison, the low reactive group was the least avoidant of unfamiliarity and most sociable at later ages. The "behavioural inhibition" dimension of temperament is believed to be associated with social anxiety in later life.[54] Based on Kagan's "behavioural inhibition" concept of temperament, the Behavioral Inhibition Questionnaire was developed by Bishop *et al.*[55] and consists of three dimensions: social novelty, situational novelty, and physical challenges.

#### 2.1.2.6 *Ann Sanson*

Sanson defined temperament as "constitutionally based individual differences in behavioural style relating to affect, activity, and attention that are visible from early childhood" (*p.*227).[37] Sanson adapted the RITQ[42] developed by Carey from the temperament theory of Thomas and Chess, and collected data on child temperament and development for the Australian Temperament Project (ATP). The ATP is a large-scale longitudinal study of the development of children in Victoria, Australia. The study

commenced in 1983 with the enrolment of 2443 families with infants aged 4 to 8 months (mean age: 25 weeks) and has followed the participants for more than 30 years.[56] Using factor analyses of the adapted RITQ, Sanson and colleagues found five temperament dimensions – approach, rhythmicity, cooperation-manageability, activity-reactivity, and irritability in infants aged 4 to 8 months.[57] Based on these findings, Sanson and colleagues developed the Short Infant Temperament Questionnaire – a 30-item questionnaire for infants under 1 year of age measuring all five temperament dimensions listed above. In a similar vein to Carey and colleagues, Sanson developed a categorical indicator of easy/difficult temperament using scores on three dimensions (approach, cooperation-manageability, and irritability). Factor analysis of the Carey's Toddler Temperament Scale[58] found seven dimensions of temperament (approach, irritability, cooperation-manageability, activity-reactivity, rhythmicity, persistence, and distractibility) in Australian toddlers aged 18 to 36 months.[3] These seven dimensions were included in the Short Temperament Questionnaire for Toddlers.[59] Factor analysis of the Child Temperament Questionnaire found four dimensions of temperament (inflexibility, persistence, sociability, and rhythmicity) in children aged 44 to 52 months. In summary, Sanson identified three dimensions which were continuously found in the first five years, namely approach/sociability, rhythmicity, and cooperation-manageability/inflexibility.[59] Persistence was an emergent dimension in toddlers (2 to 3 years) and became more important in childhood (4 to 5 years).

#### 2.1.2.7 *Summary of temperament models*

Various definitions, models, and dimensions of temperament have been proposed by different temperament researchers but none have been universally accepted as the

"gold-standard". Although different names were used, many dimensions were very similar across different models. For example, a study comparing four temperament models - Thomas and Chess, Buss and Plomin, Rothbart, and Goldsmith, found temperament dimension of 'mood' in Thomas and Chess model overlapped with 'emotionality' in Buss and Plomin's model, 'negative affectivity' in Rothbart's model and 'distress and anger' in Goldsmith's model.[60] The dimension of 'approach/shyness' and 'activity/surgency' were found in all four models. The dimension of 'persistence' in the Thomas and Chess's model was found to overlap with 'effortful control' in the Rothbart's model.

This thesis examines temperament in two ways: 1) the "difficultness concept" which categorises temperament as "easy" versus "difficult" using multiple dimensions such as approach, adaptability, intensity, rhythmicity, and mood; and 2) individual dimensions of temperament, *i.e.* examining the association between temperamental dimensions of reactivity, approach, and persistence on developmental outcomes. By categorising children into different temperament profiles, groups of children with similar temperament characteristics can then be identified and parenting interventions may be recommended according to the temperament profiles of children.[61] Examining different dimensions of temperament *i.e.* reactivity, approach, and persistence is also useful because previous studies have suggested that some dimensions may have greater influence on children's cognitive and adiposity outcomes[62,63] and are more likely to be modified[64] than others. However, most past studies are limited by the use of small samples and cross-sectional design. Studies with small samples are likely to generate low statistical power and largely varying effect sizes. Cross-sectional studies are unable to

determine the temporal order of temperament and child outcomes. In addition, early life effects of temperament might wane over time, as such cross-sectional studies may find larger effects than longitudinal studies.

### 2.1.3 Approaches to measuring temperament in infants and children

The three main approaches to measuring temperament in infants and children are direct home observation, laboratory assessment and parent-reported questionnaire. Among the three approaches, parent reported questionnaires are the most commonly used to measure temperament in infants and children. Questionnaires are relatively inexpensive to administer and score, which is ideal for studies that involve large numbers of families. There have been different parent-reported questionnaires developed for measuring temperament in infants and children, Table 2.1 summarises the commonly-used temperament questionnaires developed by different researchers. Although the validity of parent-reported information about children's temperament has been questioned because parents may provide socially desirable answers,[4] recent studies have shown that there is moderate agreement between parent-reported temperament questionnaires and other measurement approaches including direct observation.[6,65,66] For instance, Asendorpf[65] showed that parent report of children's shyness at 3 to 7 years correlated with observed shyness measures, such as latency to talk to strangers (average correlations ranged from 0.43 to 0.53). The laboratory observational approach, for example the LAB-TAB developed by Goldsmith based on 408 US children aged 4.5 years,[51] allows the detection of child reactions to specific stimuli and gives information about children in carefully controlled settings. While this observation avoids limitations of parent-report methods, children may become fearful in the laboratory environment and react differently from how they would

typically respond in a familiar environment.[5] Both the home and laboratory observation approaches are also limited because they are relatively expensive and labour intensive.[5,6] Direct observations of temperament may not be practical for most large scale, population-based studies due to costs and time constraints.

### 2.1.4 Stability of temperament

Understanding the stability of temperament is relevant to public health intervention, for instance to determine whether intervention on temperament is possible, what factors influence the stability of temperament, and at what ages is best to intervene.

Temperament is considered to be less stable under age 5 but shows considerable stability in later childhood and adulthood.[67] A meta-analysis by Roberts and DelVecchio[67] from 152 longitudinal studies found moderate stability of temperament in the first three years of life with increasing stability through to later childhood. The lower stability in early childhood may be due to the development of self-regulation/effortful control/ persistence/attention during the toddlerhood and preschool years. Higher consistency at a later age may be due to a greater ability to assimilate experience.[67]

There may also be changes in temperament across time for successive cohorts of infants and children. An Australian study comparing differences in infant and child temperament styles using data from the ATP (beginning in the 1980s) and LSAC (beginning in the 2000s) showed changes in both infant and childhood temperament over the past 20 years.[68] The study found that infants in the more recent 2000s cohort (LSAC; n=5104, mean age 8.8 months) showed lower irritability than infants in the 1980s (ATP; n=2443, mean age 5.9

months). At 4 to 5 years of age, children in the 2000s showed a higher level of sociability and a lower level of reactivity than children in the 1980s. In general, children in the 2000s were 'easier' in temperament style (less irritable and reactive and also more outgoing and sociable) than children in the 1980s. These changes may be a result of secular shifts in reporting bias, possibly due to differences in parental socio-economic positions as LSAC parents have a higher educational level than parents in the ATP.[68] Additionally, the ATP is intended to be representative of the Victorian population,[59] and the LSAC sample is considered as broadly representative of the Australian population, which is likely to result in socio-demographic differences that may have impacts on children's temperament.

As temperament may change over time and differ from one generation to another, updating the temperament norms on a regular basis is necessary. In 1986, Oberklaid *et al.*[69] examined infant temperament using the RITQ in 2443 infants from the ATP and reported that there were some differences between the Australian norms and the original RITQ norms published in the 1978. Although it is unclear whether the differences were due to cultural or temporal differences, it is possible that external environmental factors such as increased parental employment and use of child care may influence the parents' reporting of temperament and shift the norms over time.

### 2.1.5 Differences in temperament by gender, socio-economic background, and culture

Understanding the differences in temperament across culture and context is important to determine whether findings on temperament from one population could be generalised

to other populations and to develop culturally-specific norms and interventions for specific population groups.

A meta-analysis of 189 studies involving children aged 3 months to 13 years examined gender differences in three main dimensions of temperament (effortful control, negative affectivity, and surgency).[51] Findings indicated that girls have higher effortful control than boys, while boys at the same age are slightly higher in activity and positive emotions and lower in shyness than girls. However, there are negligible gender differences in negative affectivity, indicating that boys and girls do not differ in the extent to which they are difficult, emotional, or soothable.

Differences in temperament by socio-economic background have also been investigated by a number of studies. Some studies found there was no difference or small differences in temperament across socio-economic background[70] while others suggested that children from lower socio-economic families were more likely to be temperamentally difficult.[3] For instance, the ATP showed that children from higher socio-economic backgrounds were less reactive and more outgoing than children from lower socio-economic backgrounds. It is unclear why children from lower socio-economic backgrounds were more difficult but the authors suggested that this may be due to differing parental perceptions about positive and negative behaviours, parenting values and practices.[71] Other studies assessing neurobiological development suggested that growing up in a socio-economically disadvantaged environment could have an impact on children's self-regulation as it may compromise the neurobiology of the brain such as

through the expression of stress hormones such as cortisol[72,73] and the exposure to prenatal smoking in infants in more disadvantaged families.[74]

There have been some cross-cultural comparisons of temperament. Differences in temperament have been found between children from Western and Eastern countries. Children from Western countries are more outgoing and emotionally intense while children from Eastern countries are more shy.[75] This may be because temperamental shyness is viewed as positive in the Eastern culture but considered as less socially desirable in Western cultures.[75] Using the RITQ, studies have found some differences between infants from the US and infants from Taiwan[76] and Australia.[69] Compared to a sample of 203 infants from the US, infants from Taiwan (n=349) were rated as more intense, less active, less rhythmic, less likely to approach to new subjects, less adaptable, more negative in mood, less distractible, and have lower sensitivity to environmental stimuli. On the other hand, the Australian infants (n=2443) were rated as less active, less rhythmic, less adaptable and less intense than the US infants. Differences in temperament rating may be due to cultural differences in these settings.

### 2.1.6 Research question 1

***Are infant temperament norms derived from the US suitable for use in a cohort of UK infants?***

Even though there are a number of temperament measures, there is a lack of context-specific norms for temperament that are suitable for use in different settings. Previous studies of temperament have generally involved small, homogenous samples of children. For instance, norms in the RITQ were developed based on data collected from 203

infants, mainly from middle-class US families who attended local private paediatric clinics. It is possible that the norms developed from the RITQ may not be applicable to other populations. The first study of this thesis examines the infant's temperament using a population representative sample of UK children (n>10000) born in the 1990s and determines whether temperament norms derived from the US are comparable and suitable for use in the UK sample. The proportions of infants who fall into different temperament categories (*i.e.* easy, difficult) resulting from using old and new norms are compared. This research question is addressed in Chapter 4 of the thesis. The temperament categories derived from this first piece of work are then used in subsequent analyses using the ALSPAC data set in Chapters 6 and 7.

## 2.2 Parenting

While temperament is considered as an individual factor that may influence a child's development, the impact of temperament needs to be interpreted within the broader context of the child's life. Bronfenbrenner's ecological model (Figure 1.1) extends from individual characteristics of a child to interacting with the whole environment around them, including families, peers, school settings, and the wider community.[1] The interplay between a child's temperament and their environment is important for shaping their development. During the first few years of life, the home environment, family and parents have a big impact on the child by providing the resources and support needed for healthy development. There is evidence that parenting and parent-child relationships are important for children's social and emotional development,[77] cognitive functioning, future education attainment and social status as well as physical health in childhood and adulthood.[78,79]

### 2.2.1 Definition of parenting practices

To date, there have been a number of studies on parenting but different definitions of parenting are used in each study. While different terms have been used to refer to parenting, this thesis focuses on 'parenting practices'- specific behaviours parents use with their children. For example, a warmth dimension that involves hugging and kissing; and a control dimension that involves reasoning and providing autonomy support.[2] Parenting practices are different from 'parenting styles', which combine the warmth and control dimensions, for example authoritarian parenting style is characterised by having low parental warmth and high parental control.

### 2.2.2 Parenting practices that are important for child cognitive and adiposity outcomes

#### 2.2.2.1 *Parental warmth*

Parental warmth refers to the degree of affection between the parent and the child and is regarded as a continuum that is marked at one end with warmth/acceptance and the other end with hostility/rejection.[80-82] High parental warmth or acceptance is characterised by affectionate behaviours, a high level of positive regard, interest and involvement in the child's activities, responsiveness to the child's emotions, and positive expressions of approval and support.[83-85] On the other hand, parenting practices that involve coercive physical disciplines such as smacking, and overt communication of negative feelings such as criticism and disapproval are indicators of low parental warmth.[83]

High parental warmth is consistently associated with positive outcomes in children. For instance, a longitudinal study from New Zealand (n=1265) showed that children with low

parental warmth at 0 to 5 years had a three-fold increased risk of problems such as conduct disorder, depression, substance use and other health risk behaviours by age 15 years, than children with high parental warmth.[86] A randomised trial of an intervention to improve maternal warmth (responsiveness, emotional support, responses to infant foci of attention, and quality of language input) for infants aged 6 to 13 months found that infants in the intervention group had a greater improvement in socio-emotional competence, communication (use more words), and cognitive development than infants in the comparison group.[87]

#### 2.2.2.2 *Parental control*

Parental control refers to parent's use of power in achieving compliance and obedience in children along a continuum of "autonomy" at one end to "control" at the other end.[81,83] Parental control is reflected in the number of decisions parents make, the amount of supervision they exercise, and the number of rules they hold for their children.[84] High parental control is characterised by behaviours that intrude upon or hinder the child's individuation process,[88,89] for instance the use of pressure, solving problems for children, and making decisions for the child based on a parental perspective.[90] On the other hand, parental autonomy support is characterised by encouraging problem solving, giving choices, and allowing participation in decision making.[91]

The influence of parental control on child development is unclear because different studies measure parental control differently. While parents' provision of structure such as guidelines and limits (also known as "behavioural control") has a positive influence on development, parental use of power such as pressure and intrusion (also known as

"psychological control") is detrimental for child development.[92,93] For instance, a short-term (6 months) longitudinal study of 12-year-old adolescents from the United States (n=373) and China (n=433) showed that higher psychological support (choice making, opinion exchange) and higher behavioural control (solicitation, restriction) is weakly associated with better academic competence (0.02 SD higher in grades) while higher psychological control (authority assertion, guilt induction) is associated with emotional problems (0.04 SD lower emotional well-being).[94] There have been a number of studies that examined parental control in adolescence or late childhood but limited studies have examined the influence of parental control in early childhood (under the age of 5). In addition, although there is some evidence of parental control on children's academic and emotional development, there is a lack of research on children's cognitive outcomes.

#### 2.2.2.3 *Parent involvement in activities with children*

There is consistent evidence that parent involvement in activities such as playing and reading are important for children, particularly for cognitive development. For instance, a longitudinal study showed that engaging with children in activities such as writing and reading alphabets at age 4 years was associated with 0.14 SD better cognitive ability at age 8 years.[95] Parents engaging in free-play with their children is also important for language and cognitive development.[96,97] Parental involvement in children's activities can help foster parent-child relationships and provide emotional support, such as warmth and availability which are important for child language and cognitive development. A review by Desforges and Abouchaar[98] concluded that parental involvement such as teaching and engaging in cognitive stimulation activities at home, providing a learning home environment, and engaging in discussions and communications had an important

influence on children's cognitive and academic achievement and behavioural adjustment even after taking into account confounding factors such as social class, maternal education, income, and maternal psychosocial health.

#### 2.2.2.4 *Parental feeding practices*

In addition to general parenting practices (warmth, control and involvement), there is growing interest in food-specific parenting practices. Food-specific parenting practices such as feeding control[99-101] and using food to soothe[102,103] may influence children's eating behaviours and weight. However, there is a lack of clear definition for feeding control,[104,105] and there is no standard measure of feeding control.[109,110] However, feeding control can be thought of as one specific example of parental control and, like parental control, could have positive or negative impacts on children depending on a range of contextual factors. Most studies[106-108] conceptualised feeding control using the restriction and/or pressure to eat subscales from the Child Feeding Questionnaire (CFQ).[109] The restriction subscale assesses the extent to which parents restrict their child's access to foods while the pressure to eat subscale assesses the tendency to pressure the child to eat more food. However, this narrow conceptualisation of parental control yielded conflicting results on child eating and weight status (more details in section 2.4.2). More recently, Ogden *et al.* [110] suggested that feeding control can be conceptualised as two different aspects: overt control and covert control. Overt control refers to the use of control in ways that can be "detected by the child", for instance deliberately limiting certain food intake and keeping food out of reach, which are strategies measured in the CFQ. On the other hand, covert control refers to the use of control in more subtle ways that are "undetected by the child", for instance avoiding going to restaurants that sell

unhealthy foods and avoiding buying sweets and crisps and bringing them into the house. To date, it is not clear which strategies of control are beneficial and which are detrimental for children. These two conceptually different feeding controls may have different effects on child eating and weight outcome (more details in section 2.4.2).[111]

"Food to soothe" is defined as using food in response to a child's emotional distress due to reasons other than hunger. Parental use of food to soothe may also be associated with children's eating and weight status. There is some evidence that using food to soothe with a distressed child may be related to emotional eating,[112] lower fruit intake and more energy dense snacking.[111] However, the association between food to soothe and child weight outcomes is not clear (more details in section 2.4.2).

## 2.3 Temperament and parenting influences on child development

Developmental psychologists have suggested that temperament can influence child development through three main mechanisms: 1) a direct effect, *i.e.* temperament directly associated with developmental outcomes, 2) an indirect effect, *i.e.* the association between temperament and developmental outcomes is mediated by environmental factors such as parenting, and 3) an "interaction" effect, *i.e.* the association between parenting and developmental outcomes depends on the child's temperament.[5,37]

This section of the thesis explains how these three mechanisms may possibly influence children's development, and what methods were used by previous studies in examining

these three mechanisms. Section 2.4 and 2.5 review more specifically the literature on how these mechanisms influence children's cognitive ability and BMI or diet.

### 2.3.1 Direct and indirect effects

It is possible that temperament may have a direct influence on children's development. For instance, children with low persistence are less able to maintain attention and focus on a task; therefore more likely to perform poorly in academic achievement than children with a high level of persistence. Children with high emotional reactivity may be less able to control their anger and frustration and more likely to develop externalising problems (*e.g.* aggression and conduct problems).[113] Children's temperament may also influence the types of parenting they received.[114] For instance, parents engage in fewer playing or reading activities with a child who shows emotional distress and this could in turn influence the child's cognitive development. In addition to the direct effect of temperament on child health and developmental outcomes, there may be an indirect effect that is mediated by parenting.

Understanding the direct effect of temperament is important because there is growing interest on the influence of temperament dimensions, particularly persistence, on children's educational outcomes.[115,116] It is suggested that non-cognitive abilities such as temperamental persistence may have greater importance for future education attainment, earnings, and employment than cognitive abilities such as IQ.[117] If this is true, improving temperamental persistence may have positive implications for human capital at a population level.

To properly examine the direct effect of temperament, it is important for studies to take into account the possible mediation by parenting. Previous studies have used standard regression models and structural equation models (SEM) to account for parenting as an intermediate variable.[118,119] The standard regression approach[120] accounts for the mediation effect by 'adjustment' ('controlling for' the parenting variable in a regression model).

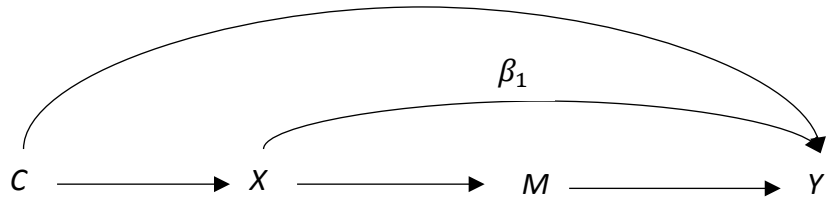


**Figure 2.1 Regression approach to estimate direct effect with exposure *X*, mediator *M*, outcome *Y* and a set of confounders *C***

Statistically, the regression approach estimates the direct effect as the following:

$$E(Y) = \beta_0 + \beta_1 X + \beta_2 M + \beta_3 C$$

where *X* refers to the exposure, *M* refers to the mediator and *C* refers to a set of confounders. $\beta_1$is taken as the *direct effect* of the exposure (Figure 2.1). However, recent advances in epidemiological and statistical research have shown that conditioning on the mediator, *M* in a regression model is subjected to limitations due to the failure to account for mediator-outcome confounders, *L* and interaction between *X* and *M*. Conditioning on *M* induces collider-stratification bias because *M* is a common effect of *X* and *L*, [121] and this could potentially lead to biased estimates of the direct effect.[122] The SEM approach is also limited because it is based on linearity assumptions and ignores the potential interaction between *X* and *M*.[123] A marginal structural model may overcome the limitations of regression and SEM approaches to estimate of the direct effect of

temperament.[124,125] The marginal structural model is not conditional on any confounders but uses weighting to account for confounders (more details in Section 3.2.2).

### 2.3.2 Theoretical "interaction" models

Developmental researchers suggest that temperament and parenting can "interact", to influence children's development.[126,127] The term "interaction" commonly used in psychological research implies that the effect of an exposure on an outcome variable is dependent on another variable.[128] In epidemiological studies, this is known as "effect-measure modification", while "statistical interaction" is referred to as the joint effect of two independent variables on a dependent variable. More details about effect-measure modification and statistical interaction are included in Section 3.2.3.

Three well-known theoretical models that explain the influence of temperament and parenting on children's development are the goodness of fit model, the diathesis stress model, and the differential susceptibility model.

#### 2.3.2.1 *The goodness of fit model*

The goodness of fit model was proposed by Thomas and Chess in the NYLS. The goodness of fit model reflects the match between the temperament characteristics of a child and the demands and expectations of the parents.[28] Thomas and Chess hypothesized that a good fit occurs when there is a match or consonance between the child's temperament and parents' demands and expectations, while a poor fit occurs as a result of mismatch between the two. The idea is that certain types of parenting will have better outcomes for children with specific temperament characteristics than for others, and that the type of parenting that is good for some children may lead to poor outcomes for children with

other temperament characteristics. For instance, a poor fit occurs when a child who is shy does not meet the expectations of parents who value positive, outgoing characteristics.

Some studies support the goodness of fit concept. For instance, a small (n=125) longitudinal study that investigated the relationship between positive and negative emotionality in 1-year old children and behavioural inhibition at age three found that parental warmth, while beneficial for most children's development, was detrimental for children high in emotional negativity.[129] The authors suggested that parents who are highly demanding tend to "push" or encourage the child to overcome their anxieties and to be less fearful while parents who are highly supportive are more likely to accept the child's inhibition. Likewise, a longitudinal study of 629 US infants found that a less positive parenting style has a better fit for slow-to-warm-up infants (infants with lower temperamental approach and adaptability) than did a more positive parenting style.[130]

#### 2.3.2.2 *The diathesis stress model and the differential susceptibility model*

Both the diathesis stress and differential susceptibility models (Figure 2.2) suggest that children with difficult temperament are more sensitive or susceptible to their environmental stimuli than children with easy temperament. In addition, the diathesis stress model hypothesizes that when parenting is negative, children with difficult temperament have poorer outcomes than children with easy temperament, but when parenting is positive, children with difficult temperament have the same outcomes as children with easy temperament.[131] The 'differential susceptibility' model proposed by Belsky[126] is distinguishable from the diathesis stress model in the sense that highly

reactive children have better outcomes than less reactive children when exposed to positive parenting.

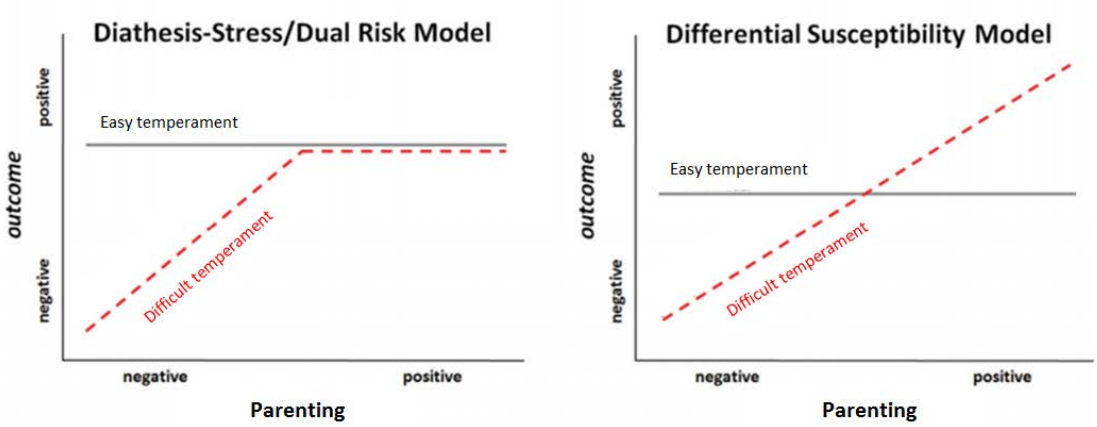


**Figure 2.2 Adapted figure for diathesis stress and differential susceptibility model[132]**

Kochanska and Kim,[133] in their sample of 186 children aged 30 to 33 months, found that when parenting was negative (low responsiveness), children with difficult temperament had more behavioural problems than their peers with easy temperament; however, there was no difference in behavioural outcomes between children with easy temperament and children with difficult temperament when parenting was positive (high responsiveness). Findings support the diathesis stress hypothesis. On the other hand, a study that examined infant temperament (measured using the RITQ) and maternal parenting (video-recorded) at age 15, 24, 36, and 54 months on children's adjustment found that children who were temperamentally more difficult had more adjustment problems than temperamentally easy children in a negative parenting environment, but showed fewer adjustment problems than children who were temperamentally easier in the presence of positive parenting, thus supporting the differential susceptibility hypothesis.[134] The findings of diathesis stress versus differential susceptibility hypothesis may be due to differences in sample characteristics, such as age, background (culture, parental socio-

economic position), the aspects of temperament and parenting that were measured and the ways they were measured (*i.e.* observational versus self-report questionnaires). In addition, most studies that examined these theoretical models are limited by small, non-representative samples and limited adjustment for confounders. Therefore, studies using population-based samples and rigorous adjustment for confounders are needed to understand whether the theoretical models are plausible.

These three theoretical models lead to the temperament and parenting as two important influences on children's development. Using an epidemiological approach, this thesis focuses on the *effect-measure modification* by temperament, *i.e.* the effects of parenting practices on children's health and developmental outcomes differ in children with different temperaments. As suggested by the differential susceptibility and diathesis stress model, the detrimental effects of negative parenting (such as low warmth) may be heightened in children with difficult temperament. If the differential susceptibility or the diathesis stress theory is supported, interventions may be better to be targeted to children with difficult temperament because they are more vulnerable to negative parenting than children with easy temperament.

To examine whether there is effect-measure modification between temperament and parenting, most studies have previously used multiple regression with a product term.[128] Statistically, the regression approach, on the linear scale, estimates the effect-measure modification as the following:

$$E(Y) = \beta_0 + \beta_1 P + \beta_2 T + \beta_3 PT$$

where *P* refers to the first exposure, *T* refers to the second exposure and *PT* refers to the product of *P* and *T* (*P* multiplied by *T*). However, the regression model does not distinguish which variable is the exposure of interest, which variable is the effect-measure modifier, or whether both variables are exposures of interest. Therefore, when using a regression with interaction terms, it is not known whether the coefficient for the product term ($\beta_3$) is interpreted as "statistical interaction", or "effect-measure modification", or both, or none.[135] Interpretations of findings are confusing, with some studies conducting effect-measure modification but interpreting results as statistical interaction, a combined effect, or joint effect and some studies doing statistical interaction but interpreting results as effect-measure modification. In addition, effect-measure modification is a scale-dependent measure. That is, effect-measure modification can be examined on two different scales: 1) a risk-difference/additive scale (for example, using a linear regression), and 2) a risk-ratio/multiplicative scale (for example, using a logistic regression).[136,137] Hence, using a regression with product terms is insufficient. The reporting of effect-measure modification on both the risk-difference and the risk-ratio scales is highly recommended in epidemiological research.[136-138] To address the methodological limitations from previous studies, this thesis uses a formally defined, contemporary epidemiological approach which provides distinctions between effect-measure modification and statistical interaction to examine effect-measure modification. Effect-measure modification on both the risk-difference and the risk-ratio scales is examined (details in Sections 3.2.3 and 6.2.4).

## 2.4 Temperament and parenting influences on cognitive outcomes

### 2.4.1 Temperament and cognitive outcomes

In this section, studies that examine the association between temperamental "difficultness" and cognitive outcomes are reviewed, followed by studies that examine the association between different dimensions of temperament (*e.g.* approach, mood) and cognitive outcomes. Finally, interventions studies that are tailored to children's temperament and aimed at improving cognitive outcomes are reviewed.

Studies that conceptualise temperament as a broad concept of "difficultness" by combining multiple dimensions have found mixed results when looking at the impact of temperament on cognitive abilities. In 1977, Carey *et al.*[139] compared cognitive ability in 51 children aged 5.5 to 7 years with different level of temperamental difficultness: easy, intermediate low, intermediate high and difficult, and found no differences in cognitive ability among the four groups. In the late 1980s, a Canadian longitudinal study involved 358 children whose temperament was measured at 4 and 8 months using five dimensions in the RITQ (adaptability, approach, intensity, mood, distractibility) and IQ was measured at 4.5 years using the Wechsler Preschool and Primary Scale of Intelligence. The study found that temperamentally difficult children had higher IQ than children with easy temperament but only in children from middle- to upper social classes.[140] However, this study did not adjust for any confounders; therefore it is likely that findings are subject to reporting bias, for instance, mothers of lower socio-economic positions and mothers with depression may be more likely to report their children as temperamentally difficult.[141] The authors suggested that children with difficult temperament may demand more

attention from, and interaction with, their parents, and these factors may positively influence their cognitive development. On the contrary, two studies have suggested that children with difficult temperament were more likely to have poorer cognitive ability.[142,143] However, findings from these two studies may not be generalised due to the use of small, non-representative samples, as one study involved 151 infants with health problems while another involved 100 children (93% Caucasians). Thus, inconsistencies in findings may be due to methodological issues such as the differences in how temperament "difficultness" was measured and conceptualised, study design, ages of study sample, and appropriate adjustment for confounders.

Studies examining the associations between individual dimensions of temperament and cognitive outcomes have found more consistent findings, with emotional reactivity/negative mood associated with poorer cognitive outcomes, and adaptability, persistence, attention and effortful control associated with better cognitive outcomes.[62,144] For instance, in a sample of 75 children, Lawson and Ruff[145] found that negative emotionality at 1 to 2 years was associated with poorer IQ at 3.5 years for boys but not girls. There are also a number of studies that examine the associations between temperament dimensions and academic achievement. Similar to cognitive outcomes, higher effortful control, attention and persistence were associated with better academic achievement.[63,146-148] The influence of temperament dimensions of approach-withdrawal and activity on academic performance is not clear. Some studies suggested that children who are temperamentally shy were more likely to perform poorly on academic tests at age 9 to 13 years (n=125).[149,150] The authors suggested that shy children were less likely to engage in classroom activities, and therefore more likely to be reported by teachers to

have poorer academic achievement.[149] Two studies that examined activity in infancy reported that higher activity was associated with better cognitive outcomes[151,152] while others examined activity of preschoolers and reported a negative association between activity and cognitive/academic outcomes.[62,63] High activity during infancy may indicate a higher level of curiosity and parents respond to their infants' high activity by engaging in more interactions. On the other hand, children who are highly active may present challenges to parents and teachers with maintaining their engagement with learning activities.

To date, the majority of studies that examined the associations between temperament and cognitive or academic outcomes are cross-sectional, involved small samples and many have not adjusted for important confounding factors. For instance, most studies have not included maternal depression as a potential confounder although evidence suggests that maternal depression is associated with child temperament[153,154] and cognitive and academic outcomes.[155] In addition, it is likely that the effect of temperament on cognitive outcomes is mediated by parenting factors. Studies have shown that parents who are more involved in teaching and childrearing are influenced by children's temperament[156-158] and higher parental involvement in cognitive stimulation activities such as reading has a positive effect on children's cognitive development.[95] Although at least two studies have accounted for parenting by regression adjustment or path analysis in structural equation models,[118,119] these methods have several limitations (see section 2.3.1 Direct and indirect effect).

Several intervention programs aimed at improving children's cognitive and academic outcomes have been designed specifically for temperamentally difficult children, recognising the additional challenges that this group sometimes face. For instance, a randomized controlled study that involved 100 infants who were temperamentally difficult found that improving parent-child interaction through play and providing more stimulation resulted in less crying and more cognitive exploratory behaviours in the intervention group than the control group.[159] Intervention programs have been designed to modify children's self-regulation abilities, emotional competence, and coping skills and to improve behaviours, cognitive, and academic outcomes.[9,116,160] The INSIGHTS into Children's Temperament intervention is designed to teach parents and teachers strategies to enhance children's self-regulation and reduce behavioural problems and to improve academic achievement.[161,162] Similarly, the 'Tools of the Mind' is a curriculum that promotes self-regulation through a comprehensive system of activities. Such temperament-based interventions have been used in studies targeting children from low income families[10] and temperamentally difficult children who are at risk for poor outcomes.[163]

### 2.4.2 Parenting influences on cognitive outcomes

Longitudinal observational studies have shown that parental warmth (such as sensitivity, positive affect, and cognitive stimulation) has beneficial impacts on children's cognitive and language development while controlling, punitive parenting is detrimental for cognitive development.[96,164] Findings from the 1946 British Birth cohort (n=1690) showed that parental coercive disciplines characterised by use of threat, discipline and disapproval at age 4 years was associated with 0.23 SD lower IQ z-score at age 8 years

even after adjusting for a range of confounders, including parental cognitive ability, education, social class, physical and mental health.[95]

There is also some evidence that interventions aimed at improving parenting practices had positive impacts on their children's cognitive development and academic outcomes. A systematic review of 50 articles showed that interventions were effective in improving maternal warmth and responsiveness and had about 0.50 to 0.75 SD improvements on children's cognitive or academic outcomes.[14] These interventions were typically provided through home visits, clinic care, skills training and adult education with strategies that focused on encouraging mother-child interactions through play and promoting psychological stimulation such as reading and telling stories.[15] However, it is difficult to tease out the specific effect of parenting because some of the interventions also include other components such as nutrition supplementation. A meta-analysis showed that the standardised effect sizes of home visiting interventions on children's cognitive outcomes and parenting behaviours were 0.18 and 0.14, respectively.[16]

General parenting interventions such as Family Nurse Partnership (UK), Nurse Family Partnership (US) and Sure Start (UK) suggest the impact on children's IQ is mixed. A randomised controlled trial of the nurse-home visiting programme provided to 743 black women in the US who were vulnerable (unmarried, less than 12 years of education, or unemployed) showed small improvements in IQ (0.18 SD), vocabulary (0.17 SD), and academic achievement (0.09 SD).[165] A quasi-experimental study of the Sure Start programme, targeted to the 20% most disadvantaged areas in the UK, showed less negative parenting such as harsh discipline and parent-child conflicts (0.44 SD) and better

home environment (for example, reading to the child, helping to learn) (0.27 SD) but no evidence of impacts on vocabulary (measured by Peabody Picture Vocabulary Test).[166] A pragmatic randomised controlled trial assessing the effectiveness of the Family Nurse Partnership provided to teenage mothers (less than 19 years) in the UK found no evidence of impacts on mother-child interactions or cognitive development.[167] Mixed findings may be due to differences in type and design of intervention (randomised versus quasi), length and intensity, population targeted (*e.g*. teenage mothers versus mothers with multiple risk factors), location (rural versus urban), and the targeted outcomes of interventions.

### 2.4.3 Temperament and parenting and influence on cognitive outcomes

There is a large body of research examining whether temperament modifies the association between parenting and children's behaviours (*e.g.* externalising and internalising problems) but few studies have focused on cognitive outcomes. Poehlmann *et al.*[168] examined the impact of parenting behaviours (communication, positive affect and connectedness) and temperamental emotional distress at 9 months on IQ at 36 months preterm infants (<37 weeks). Temperamentally distressed children had higher IQs than less distressed children when parenting was positive but had lower IQ when parenting was negative, supporting the differential susceptibility hypothesis (Figure 2.1). The findings of this study may not apply to infants born full-term due to differences in parenting behaviour.[169] Conway and Stifter[170] observed maternal attention-directing behaviours and children's temperamental inhibition at 24 to 25 months on executive function (*i.e.* ability to engage in goal-directed thought or action through inhibitory control, attention shifting, and working memory processes)[10] at age 4.5 years and found

that attention-directing behaviours were associated with executive function only among temperamentally inhibited children. However, this study was limited by its small (n=68), non-representative sample (predominantly White, educated, middle-class families), thus findings cannot to be generalised to the whole population.

The two studies described above examined whether the association between parenting and cognitive outcomes differed across children with different temperament, *i.e.* the "effect-measure modification" by temperament. However, results were interpreted as "interaction" (joint effect of both temperament and parenting). Chapter 3 includes a deeper discussion of analytical approaches needed to examine effect-measure modification and interaction.

### 2.4.4 Research question 2

***Is temperament at age 2 to 3 years directly affecting children's cognitive (verbal and non-verbal) and academic (literacy and numeracy proficiency) development at age 6 to 7 years?***

There is wide interest in the influence of temperament dimensions on cognitive/academic outcomes.[171,172] However, the vast majority of the studies are cross-sectional and hampered by methodological weaknesses including failure to control for important confounding factors, and small samples. Importantly, most studies have not accounted for parenting practices as an intermediate variable; parenting is influenced by child temperament and is a major influence of children's cognitive and academic achievement. The few studies that accounted for parenting by adjusting for parenting in the regression models or SEM, but there are several limitations for using these methods, as described in

section 2.3.1. Using a national representative of Australian children, the second study of this thesis uses a marginal structural model to examine whether temperament at ages two to three years is associated with children's cognitive and academic outcomes at age 6 to 7 years by properly taking into account parenting practices at age 4 to 5 years as an intermediate variable and a range of potential confounding factors. This research question is addressed in Chapter 5, in the form of an academic manuscript currently under review.

### 2.4.5 Research question 3

***Does infant temperament modify the effects of parenting practices (warmth and control) at age 24 to 47 months on IQ at age 8 years?***

Few studies have examined whether the associations between parenting and cognitive abilities differ by infant temperament. The regression approach used by previous studies has not been able to distinguish effect-measure modification from statistical interaction. While effect-measure modification is of interest in some studies, results are often interpreted as statistical interaction (more details are included in Sections 3.2.3 and 6.2.4). The third study of this thesis examines whether parenting practices (warmth and control) at age 2 to 4 years are associated with IQ at age 8 years and whether the associations are modified by infant temperament (*i.e.* the effect-measure modification by infant temperament). This research question is addressed in Chapter 6, in the form of an academic manuscript that is published at *PLoS One.*

## 2.5 Temperament and parenting influences on adiposity and diet

### 2.5.1 Temperament influences on adiposity and diet

With obesity being one of the biggest challenges to public health of modern times,[22] the literature on how temperament influences children's eating behaviours and risk of obesity is growing. A 2014 systematic review of 18 papers (5 cross-sectional, 13 longitudinal) found some evidence of an impact of temperament on BMI and weight gain in infants and preschool-aged children.[173] However, effect sizes (Cohen's d) of these studies varied from very small (*d*=0-0.1) to large (*d*>1.0), possibly due to variation in study design, sample size, measures of temperament, age when outcomes were measured and analytical approach. In this section, studies are reviewed that conceptualised temperament as "difficult" followed by studies that examined specific dimensions of temperament (*e.g.* approach, reactivity).

The earliest study of temperament on children's adiposity was conducted by Carey[174] in 1985 and involved 200 middle-class US infants aged 6 to 12 months. Of the infants who had higher weight gain from 6 to 12 month, a large proportion were temperamentally difficult (defined as low on rhythmicity, approach, and adaptability and high on emotional intensity and negative mood) at 6 months. This finding supported by a large study involving 30000 Norwegian infants.[175] Paradoxically, Wu *et al.*[176] reported an association between temperamental difficulty and infant weight gain after stratifying by maternal warmth. When maternal warmth was low, children who were temperamentally difficult (measured using the RITQ) had higher BMI percentile at age 5 to 12 years than children who were easy. Thus, increasing speculation that the association between temperament difficultness and weight may depend on the level of maternal warmth.[177]

Differences in findings may also be due to the use of different temperament measures and concepts of "difficult" temperament. For instance, Carey[174] and Wu[176] defined difficult temperament from the RITQ, whereas Niegel *et al.*[175] defined difficult temperament using seven items from the Infant Characteristics Questionnaire which assessed levels of fussiness. Positive associations in studies by Carey[174] and Wu *et al.*[176] may also be due to limited confounding adjustment.

When different dimensions of temperament are examined rather than a broad concept of "difficultness", there is some consistency in findings, with negative emotionality (*e.g.* fussing and crying) associated with increased body weight in children and better self-regulation associated with lower body weight.[173] A longitudinal study with 206 infants from low income African families found that children with higher negative emotionality had a 0.24 SD increase in skinfold fatness and about 0.04 SD increase in weight-length z-score at 12 months.[178] Francis and Susman[179] reported that children in the poorest self-regulation group (low inhibitory control and delay of gratification) at age 3 years had higher gains in BMI z-score (0.57±0.05) from age 3 to 12 years than groups of children with better self-regulation. There are several possible explanations for why children with a difficult temperament or negative emotionality may have a higher weight status. Parents may respond to a child's negative emotions by attempting to calm using sweets foods or drinks,[180] television exposure,[181] or by feeding solid food to infants.[182] On the other hand, children with poor self-regulation may be less able to resist the temptations of palatable foods and have higher intake of obesogenic food than children with better self-regulation.[179]

### 2.5.2 Parenting influences on child adiposity and diet

The influence of parental warmth and control on children's risk of obesity is not clear. A longitudinal study of 850 American children showed that low maternal warmth (low sensitivity, positive regard, supportive presence respect for autonomy, and high intrusiveness and hostility) assessed using direct observation at 15, 24, and 36 months was associated with 42% higher risk of obesity at 12 to 15 years of age, after adjusting for socio-economic position, birth weight, and maternal obesity.[183] On the other hand, a cross-sectional study by Wake *et al.*[184] using data from the Longitudinal Study of Australian Children (LSAC) found no association between maternal and paternal warmth and control on BMI at age 4 to 5 years (n=3040). A more recent study using LSAC data reported that parental consistency (*e.g.* setting and enforcing explanations and limits) is protective of children's BMI.[185] Differences in findings may be due to discrepancies in study design (longitudinal versus cross-sectional), context (America versus Australia), sample sizes (a chance association may have occurred in the smaller study), measures of warmth, sample ages (*e.g.* younger children need higher warmth than older children), or measurement approach (observation versus self-reported).

Current literature also suggests that parental feeding control and using food to soothe may influence children's BMI. Most studies conceptualised feeding control as "overt" (directive strategies that are detectable by the child), for example, restricting children's access to specific foods or amounts of food. A number of studies have suggested that food restriction may increase children's preference for the restricted foods, increasing children's eating in the absence of hunger and weight gain.[99,100] However, much of the

evidence of the positive association between restriction and weight status comes from analyses of a longitudinal cohort of American girls (about 5 years) from White, middle to high income, well-educated families.[100,101,186] Other studies using samples from different populations found no association between food restriction and eating and weight outcomes.[106,187-189] On the other hand, fewer studies have examined "covert" control. These are less directive strategies that are undetected by the child, such as hiding or not purchasing unhealthy foods, which may have a different influence on children's eating and weight status. A cross-sectional study examined parental covert and overt control on snacking behaviours in a sample of 61 children aged 4 to 7 years and reported that covert control was associated with a decreased intake of unhealthy snacks, and interestingly, overt control was associated with an increased intake of healthy snacks. The authors suggested that parents who used overt control to manage the child's food intake but did not use food as a reward or promote unhealthy food intake may be beneficial for the child.[110]

Another parenting practice that has been argued as contributing to children's weight problems is the use of food to soothe a distressed or unhappy infant or child. There have been at least four published studies reporting food to soothe and children's weight or BMI outcomes.[102,103,188,190] These studies vary in design (cross sectional,[103] longitudinal[102]) and size (ranging 78 to 438 participants), with two studies suggesting that food to soothe is associated with higher weight gain or BMI at 6 to 18 months[102,103] and two reporting food to soothe was not associated with weight at preschool years.[188,190] One explanation for these differences may involve the way in which food to soothe is measured.[102] When completing questionnaires, parents may be less willing to report their actual use of food

to soothe, whereas parents may be less likely use food to soothe their distressed child in a laboratory setting.[102] It is difficult to understand why there are such differences between studies, and more research on whether food to soothe is convincingly linked to children's adiposity is warranted.

Parents have been the primary target of interventions to reduce childhood overweight and obesity. However, three large randomised controlled trials (Infant Feeding Activity and Nutrition Trial, NOURISH, Healthy Beginnings Trial) in Australia covering different aspects of parenting behaviours have found little or no effect on children's dietary intake and BMI despite adopting different approaches, covering a broad range of messages and assessing various outcomes. For instance, the Infant Feeding Activity and Nutrition Trial, involved 542 first-time parents from representative backgrounds was conducted for 20 months to improve parenting skills to support healthy eating and to promote physical activity. The trial lowered sweet foods consumption (assessed by three days 24 hour diet recall) in children but had no effect on BMI at 20 months of age.[191] The NOURISH trial is a 6-month obesity-prevention intervention with 698 first-time mothers and their healthy term infants aimed to encourage parents to use repeated exposure to unfamiliar foods and responsive feeding, and to promote parenting (warmth, autonomy encouragement, and self-efficacy). The NOURISH intervention group used less feeding control and food to soothe, but no difference in BMI z-score was found.[192] These high-quality interventions showed that parental feeding practices were modifiable; however, it remains unclear how to intervene on childhood overweight effectively. Providing education and support to parents may not be sufficient to change children's eating behaviours and weight outcomes. Individual factors such as child temperament may need to be considered, for

instance the association between parental feeding practices on BMI may differ according to child temperament.

### 2.5.3 Temperament and parental feeding practices influence on adiposity

Guided by the temperament and parenting theoretical "interaction" models, more recent studies have started to look at whether temperament may put some children at higher risk of overweight than others, particularly in the presence of negative feeding practices such as high restriction. No study has yet examined whether children with different levels of "difficultness" are differentially influenced by parental feeding practices, although there are a few studies that have looked at individual dimensions of temperament. For instance, Rollins[193] examined whether the effect of food restriction differs by child temperamental approach, *i.e.* the level of excitement or joy in response to pleasurable activities such as eating. This experimental study was conducted over a five week period with 37 children aged 3 to 5 years. Children were given free access to unlimited sweet snacks for 2 weeks. Based on the children's preferences, experimenters decided the type of restricted and unrestricted food to be used in subsequent sessions. In week three and four, children underwent a period of restriction where the restricted food was available for only five minutes while the unrestricted food was available throughout the whole session. This study found that children with high temperamental approach (*i.e.* level of excitement experienced in respond to pleasurable activities such as food) had a higher intake of restricted snacks after the two-weeks restricted period compared to baseline while children with low approach had no increase in intake of restricted snacks. This study suggests that children with high temperamental approach may be more susceptible to the influence of food restriction than children with low temperamental approach. In a

study of 197 non-Hispanic White girls, Anzman-Frasca[194] found that children with low inhibitory control (a dimension of temperament that overlaps with impulse control, self-regulation, and executive function) and high restrictive parenting had a higher BMI and weight gain from age 7 to 15 years than children with high inhibitory control and high restrictive parenting after adjusting for family income, parental education and parental BMI. A similar finding was found in a study by Rollins[186] (n=180) with children whose parents allowed unlimited access to snacks having higher weight gain than children whose parents set limits but did not restrict snacks. In addition, the detrimental effect of unlimited access to snacks on weight gain was larger in children with low inhibitory control than children with high inhibitory control. This study suggests that children with lower inhibitory control may be less able to control their intake in the presence of palatable foods.

There are several methodological limitations in these previous studies, including small sample size (range from 37 to 180) and problems with analysis and interpretation of effect-measure modification. As described earlier in section 2.3.2, analyses of effect-measure modification were poorly conducted and interpreted in all these studies.

### 2.5.4 Research question 4

***Is temperament at age 0 to 5 years associated with children's BMI at age 7 years? Are parental feeding control and use of food to soothe at age 3.5 years associated with BMI z-scores and fat mass in childhood (7 years) and adolescence (15 years)? Does infant temperament modify the associations between feeding practices and BMI?***

Findings on the influence of feeding control and use of food to soothe on children's adiposity is mixed. Most of the studies examining parental feeding practices on children's adiposity are cross-sectional or short-term longitudinal studies (less than three years) and involved small samples. The fourth research question of this thesis is to examine whether parental feeding control and use of food to soothe when children were at the age of 3.5 years are associated with BMI at 7 and 15 years and fat mass at 15 years. Additionally, this study examines whether the effect of feeding control differs by infant temperament (*i.e.* effect-measure modification by temperament). This study aims to fill the gap identified in the literature by using data from a population-based cohort study (ALSPAC) with a larger sample size (about 8000 children), longer follow-up, using multiple indicators of adiposity, and more robust analysis of effect-measure modification (more details in Section 3.2.3). This research question is addressed in Chapter 7.

# CHAPTER 3. Methods

This chapter describes: 1) the two datasets - the Avon Longitudinal Study of Parents and Children (ALSPAC) from the United Kingdom and the Longitudinal Study of Australian Children (LSAC) from Australia - used to address the research questions, 2) the temperament, parenting, cognitive ability, academic achievement and adiposity measures in these two datasets, and 3) the methodological and analytical approach used to answer the four research questions.

## 3.1 Data sources

Data from two longitudinal, population-based observational studies – The Avon Longitudinal Study of Parents and Children (ALSPAC) and the Longitudinal Study of Australian Children (LSAC) - were used in this thesis because they contain rich information on child temperament and parenting practices, which were measured repeatedly and consistently across time, and information on important developmental and health outcomes (cognitive ability, academic achievement, adiposity). Information on a wide range of confounders such as socio-economic position and parents' health were also collected. These two datasets adopted different measures of temperament and parenting practices. As there may be differences in temperament and parenting across cultural context and time, the use of these two datasets allows accumulation of evidence across two different populations to better understand the overall concept of "temperament". Moreover, ALSPAC has a longer follow-up than LSAC, this allows the investigation of the influence of parenting and temperament on later developmental outcomes (adiposity at adolescence).

Table 3.1 summarises the research questions addressed in the thesis and the data used to answer each research question. Research question 1 examines the temperament profiles of UK infants using data from ALSPAC. Temperament norms for the Revised Infant Temperament Questionnaire were originally developed for US infants, and norms for Australian infants were published in the 1980s from the Australian Temperament Project but temperament norms in UK infants had not been published.[69] Therefore, ALSPAC was used to fill in the gaps in the literature. LSAC was used to answer research question 2 because the association between temperament and IQ in the ALSPAC sample has been explored in a previous Master's degree dissertation project.[195] Research question 3 examines the associations between parental warmth and control on IQ and the effect-measure modification by temperament. ALSPAC was used to answer research question 3 because the parental warmth and control variables in LSAC are highly skewed with approximately 80% parents reporting the highest possible score. In research question 4, we focus on two specific aspects of parenting (parental feeding control and use of food to soothe) and adiposity outcomes. ALSPAC was used because there are parental feeding control and use of food soothe items in ALSPAC but not in LSAC.

**Table 3.1 Outline of data used to answer each research question**

| Research question | Data source |
|---|---|
| 1. Are infant temperament norms derived from the US suitable for use in a cohort of UK infants? | ALSPAC |
| 2. Is temperament at age 2 to 3 years directly affecting children's cognitive (verbal and non-verbal) and academic (literacy and numeracy proficiency) development at age 6 to 7 years? | LSAC |
| 3. Does infant temperament modify the effects of parenting practices (warmth and control) at age 24 to 47 months on IQ at age 8 years? | ALSPAC |
| 4. Is temperament at age 0 to 5 years associated with children's BMI at age 7 years? Are parental feeding control and use of food to soothe at age | ALSPAC |

| | | |
|---|---|---|
| | 3.5 years associated with BMI z-scores and fat mass in childhood (7 years) and adolescence (15 years)? Does infant temperament modify the associations between feeding practices and BMI? | |

Overall, this research roughly followed a triangulation design by combining evidence from different data sources to gain insight into the topic.[196] In this research, triangulation was achieved through the use of two population-based studies from different countries, measuring different aspects of temperament and parenting using different questionnaires, and analysing data using different approaches. This design led to a better understanding of the interplay between parenting and children's temperament.

### 3.1.1 Avon Longitudinal Study of Parents and Children (ALSPAC)

ALSPAC is a prospective observational study investigating the influence of genetic and environmental characteristics on health and development across the life course. All pregnant women residing in a defined part of the former county of Avon in the Southwest of England with expected delivery date between 1st April 1991 and 31st December 1992 were eligible to be participate in ALSPAC (Figure 3.1).[197] The study has been described in detail elsewhere.[197,198]


**Figure 3.1 ALSPAC eligible study areas. Pregnant women from three District Health Authorities (DHA) were recruited[198]**

Media information was used to encourage participation of women as early in pregnancy as possible. Antenatal care centres and maternal health services were used to promote the study.[197] Eligible women were given "expression of interest" cards which allowed them to request further information or to decline participation. Women who requested further information were sent a study information booklet followed by an initial questionnaire a week later. The eligible sample has been defined retrospectively, based on ALSPAC recruitment records and maternity, birth and child health records. The recruitment campaign identified 20248 eligible pregnancies. Mothers with no known contacts or opted out via expression of interest card were not enrolled. A total of 14541 pregnant mothers (71.8% of all pregnancies at that time) were recruited. Of these 14541 pregnancies, 68 had no known birth outcome, 195 were twins, 3 were triplets and 1 was quadruplet, resulting in 14676 known foetuses. These pregnancies resulted in 14062 live-born children of whom 13988 infants survived to 1 year of age.

Compared to the 1991 National Census based on women with infants of less than 1 year of age, ALSPAC mothers were more likely to live in owner-occupied accommodation (79.1% in ALSPAC, 63.4% in UK and 68.7% in Avon), have a car (90.8% in ALSPAC, 75.6% in UK, and 83.7% in Avon), be married (79.4% in ALSPAC, 71.8% in UK and 71.7% in Avon) and less likely to be non-White (2.2 % in ALSPAC; 7.6% in UK, 4.1% in Avon).[198] Although mothers in the ALSPAC sample had broadly higher socio-economic position indicators than mothers in Avon and the UK, mothers in ALSPAC were more likely to live in a crowded household (more than one persons per room) (33.5% in ALSPAC, 30.8% in UK and 26.0% in Avon).

Information about the mothers themselves, their partners, and the study child were collected from the pregnant mothers through self-completed questionnaires since enrolment. Starting from the age of 7 years, all children were invited to attend the annual clinical assessment, known as 'Focus clinic'.[197] The 'Focus clinic' included assessments for cognitive ability (IQ, reading ability), anthropometry (weight, height, waist circumference, fat mass) and collected biological samples (saliva, urine, and blood).

#### 3.1.1.1 *Cognitive ability*

Cognitive ability was measured at age 8 years in the 'Focus clinic' using the Weschler Intelligence Scale for Children (WISC-III $^{UK}$) administered by the ALSPAC psychology team. The WISC-III consisted of two subscales: verbal and performance, with five subtests in each subscale. The five subtests in the verbal subscale were Information (general knowledge), Similarities (verbal abstract reasoning), Arithmetic (numeric reasoning), Vocabulary (understanding of different words), and Comprehension (social

comprehension and judgement). The five subtests in the performance subscale were Picture Completion (attention to visual detail), Coding (visual-motor skills, processing speed), Picture Arrangement (attention to visual detail, sequential reasoning), Block Design (visual abstract ability), and Object Assembly (visual analysis and construction of objects).

To reduce response burden, a short form of the measure with alternate items from each of the ten subtests was administered. Individual items within each subtest were summed and multiplied by 2 for picture completion, information, arithmetic, vocabulary, comprehension, and picture arrangement; multiplied by 5/3 for similarities; and multiplied by 3/2 for object assembly and block design. This made the raw scores comparable to those that would have been achieved had the full test been administered. Raw scores were converted to age-scaled scores according to standard procedures.[199] The scores on verbal and performance subscales were combined to form the Full-scale IQ. The IQ scores were standardised on a normal British population in the early 1990s to have a mean of 100 and standard deviation of 15. In research question 3 (Chapter 6), continuous IQ scores were used in linear regression models to examine the associations between parental warmth, hostility and IQ. IQ was dichotomized into 'low' (less than 85) and 'high' (85 and above) in analyses of effect-measure modification by temperament.

#### 3.1.1.2 *Adiposity*

This thesis used BMI data collected at age 7 and 15 years as measures of adiposity in childhood and adolescence. BMI was calculated as:

$$BMI = \frac{weight\ (kg)}{height\ (m^2)}$$

Weight was measured to the nearest 0.1 kg using Tanita scales and height was measured to the nearest 0.1 cm with a Harpenden stadiometer. Age and sex specific BMI z-scores were created based on the 1990 British Growth reference.[200] Overweight and obesity were defined according to the International Obesity Task Force (IOTF) standard (Table 3.2).[201]

**Table 3.2 International Obesity Task Force (IOTF) age and sex specific cut points for BMI used to define overweight/obesity[201]**

| Age (years) | BMI cut points ($kg/m^2$) | |
|---|---|---|
| | Males | Females |
| 6.5 | 17.71 | 17.53 |
| 7.0 | 17.92 | 17.75 |
| 7.5 | 18.16 | 18.03 |
| 14.5 | 22.96 | 23.66 |
| 15.0 | 23.29 | 23.94 |
| 15.5 | 23.60 | 24.17 |

BMI is a convenient measure of adiposity, but it does not differentiate lean mass from fat mass. Therefore, fat mass was used as another indicator of adiposity. Fat mass was measured at age 15 years in the 'Focus clinic'. A Lunar prodigy fan beam densitometer was used to perform a whole body dual energy X ray absorptiometry (DXA) scan to measure fat mass. The procedure was clearly explained to the parent and the child and parental consent was obtained before proceeding. BMI and fat mass were used as outcome measures in research question 4 (Chapter 7).

#### 3.1.1.3 *Temperament*

Temperament at age 6 months was reported by mothers using the adapted Revised Infant Temperament Questionnaire (RITQ). The RITQ was developed by Carey and colleagues based on Thomas and Chess's temperament model (Chapter 2) and consists of

95 questions in nine temperament dimensions– activity, rhythmicity, approach, adaptability, intensity, mood, distractibility, persistence, and threshold.[42,58] Definitions and examples of items in each dimension were provided in section 2.1.2.1.

The adapted RITQ used in ALSPAC consists of 88 items. There were 12 items in activity (1 omitted), 10 items in rhythmicity (2 omitted), 10 items in approach (1 omitted), 10 items in adaptability (1 omitted), 10 items in intensity (no item omitted), 9 items in mood (1 omitted), 7 items in persistence (1 omitted), 10 items in distractibility (no item omitted), and 10 items in threshold (no item omitted). Items were omitted as they were shown to have non-response of at least 10% in a pilot test prior to ALSPAC because they were considered by parents as irrelevant to the child. For example, a question on the baby's reactions when having an infection was irrelevant to those mothers whose baby had never had an infection, and hence this question was omitted. The internal consistency of the adapted RITQ ranged from 0.40 to 0.73 for subscales, and 0.79 for the composite score. This is consistent with the internal consistencies of the original questions included in the RITQ.[42,202] Response for items ranged from 1 (almost never) to 6 (almost always). Higher scores indicated a more 'difficult' temperament, *i.e.* high activity, low rhythmicity (consistency in daily functions), low approach to new situation, low adaptability, high emotional intensity, negative mood, easy distractible, low persistence and low threshold to external stimuli. Based on the categorisation algorithm suggested by Carey and McDevitt,[41] infant temperament was categorised into easy, intermediate low, intermediate high, and difficult (more details in Chapter 4).

Temperament at age 38 months was measured using the Buss and Plomin's Emotional, Activity, and Sociability (EAS) questionnaire. The EAS consists five items in each of the four dimensions:

1. Emotionality - tendency to get distressed and upset. Examples of emotionality items included 'child cries easily' and 'child reacts intensely when upset'.
2. Activity - speed and intensity of speech, bodily movements and energetic behaviours, for instance, 'active as soon as wake up' and 'always on the go'.
3. Shyness - feelings of tension and distress in social situations. Items in the dimension of shyness included 'child tends to be shy' and 'child takes a long time to warm up with strangers'.
4. Sociability - preference to be with others, for instance, 'likes to be with people' and 'feels isolated when alone'.

Responses for items in EAS questionnaire ranged from 1 (not at all like him/her) to 5 (exactly like him/her). Higher scores indicate higher emotionality, activity, shyness, and sociability. Each dimension in the EAS questionnaire demonstrated acceptable to good internal consistency.[203] The EAS questionnaire was used in Section 7.2 to examine the associations between temperament and BMI.

Temperament measured in ALSPAC at the age of 6 months were used to address research question 1, 3, and 4. In research question 1 (Chapter 4), temperament norms (means and standard deviations) were compared for the US and the UK infants. In research question 3 (Chapter 5), temperament was conceptualised as an effect-measure modifier of the parental warmth-IQ and parental control-IQ associations. In research question 4 (Chapter

6), temperament was used as a confounder of the associations between parental feeding control and adiposity and parental use of food to soothe and adiposity.

#### 3.1.1.4 *Parenting*

As described in Chapter 2, parental warmth and control are two important aspects of parenting that impact child development. There are a number of items relating to parenting in ALSPAC questionnaires but no specific instrument of parenting warmth and control. Based on the warmth and control definitions in the literature, 17 items measured from 24 to 47 months from the ALSPAC questionnaires that were related to parental warmth and control were identified. Items selected as indicators of parenting warmth included kissing, hugging, cuddling, shouting, and smacking (Chapter 6). An example of a parental warmth item was 'mother smacks child during tantrums'. Responses for this item included never, rarely, once a month, once a week, daily. Indicators of parenting control included reasoning with the child, degree of choices with food, clothes and battle of wills between mothers and their study child. An example of a parental control item was 'mother reasons with child during tantrums'. Responses for this item were often, sometimes, never. All parenting items were reported by mothers on three to five point Likert scales. Parental warmth and control were conceptualised as two exposures in research question 3 (Chapter 6).

To examine the associations between parenting practices and adiposity, two food-specific parenting dimensions were used, *i.e.* feeding control and using food to soothe. Parental feeding control and use of food to soothe were measured at 42 months (3.5 years of age) from items selected from the ALSPAC questionnaire. Parental use of control was

measured using an item 'how much choice do you allow him in deciding what foods he eats at meals?'. Mothers responded to this item as 'he can choose from any food available', 'he is given a choice from a few alternatives that I select', and 'I decide what he will eat'. Mothers' use of food to soothe was measured using one item 'how often do you use sweets or other foods to stop his crying or fussing?'. Response to this item was 'frequently (once a day or many times a week)', 'infrequently' and 'never'. Parental feeding control and use of food to soothe were conceptualised as two independent exposures in research question 4 (Chapter 7).

### 3.1.2 Longitudinal Study of Australian Children (LSAC)

LSAC is a population-based longitudinal study following children and families across Australia that commenced in 2004. The study recruited two cohorts - infants in their first year of life (B cohort, born between March 2003 and February 2004) and children in their fourth year of life (K cohort, born between March 1999 and February 2000) at study commencement.[204] The Medicare enrolment database was used as the sampling frame to identify eligible children. Postcodes were stratified by state/territory and by capital city statistical division/remainder of state to ensure the sample was geographically representative. A two-stage clustered sampling design was used to select participants.[205] This included a random selection of postcodes at stage one and then a random selection of children within each selected postcode at the second stage. For both cohorts, children were selected from the same 311 postcodes. Some remote parts of Australia were excluded from the study design. On most characteristics, the LSAC sample is considered broadly representative of Australian children.[206]

This thesis used data collected from the infant cohort. A total of 5107 infants (0 to 1 year) were recruited into LSAC in 2004, and they were followed up at age 2 to 3 (n=4606), 4 to 5 (n=4386), and 6 to 7 (n=4242) years. Data were collected biennially since 2004 through face-to face interviews with the primary caregivers (at recruitment 97% were mothers), questionnaires for the parents and teachers, and direct assessments with the child.

#### 3.1.2.1 *Cognitive ability*

Cognitive ability in LSAC was measured using an adapted version of the Peabody Picture Vocabulary Test (PPVT-III) for verbal ability and the Matrix Reasoning test from the Wechsler Intelligence Scale for Children, 4$^{th}$ edition for non-verbal ability. The adapted PPVT was administered by a trained interviewer to children aged 6 to 7 years during home interviews. There are 40 items in the adapted PPVT. The child was asked to point to the picture that best represents the meaning of a word spoken by the examiner. Scores were created using Rasch modelling to enable comparison of scores across ages.[207] Scores on the adapted PPVT in the LSAC sample ranged from 36 to 92 with higher scores representing better vocabulary. The Matrix Reasoning test comprised 35 items of increasing difficulty. The child was presented with an incomplete set of diagrams and was asked to select the picture that completes the set from five different options (Figure 3.2). Scores from the Matrix Reasoning test were reported as standard scores. The score of Matrix Reasoning test in LSAC sample ranged from 3 to 19.

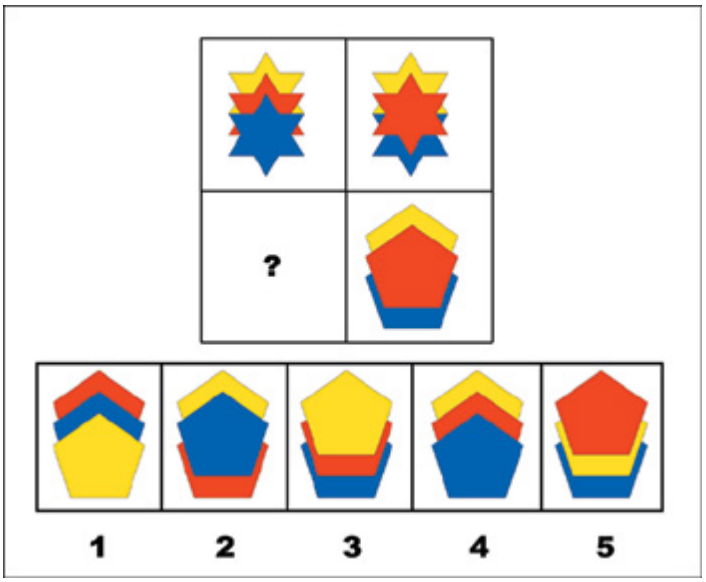


**Figure 3.2 An example in the Matrix Reasoning test**

#### 3.1.2.2 *Academic achievement*

Academic achievement was reported by teachers at ages 6 to 7 years using the adapted Academic Rating Scale (ARS).[208] The ARS consists of ten items in the literacy subscale (alpha=0.91) and nine items in the numeracy subscale (alpha=0.94). Teachers were asked to rate the child's skills and knowledge in literacy and mathematical understanding in relation to other children of the same age. Examples of items in the literacy subscale included 'reads age-appropriate books fluently' and 'writes sentences with more than one clause'. Items in the numeracy subscale included 'uses a variety of strategies to solve math problems' and 'makes reasonable estimates of quantities'. Teachers rated each item from 1 (not yet demonstrated skill) to 5 (demonstrates skill competently and consistently). Total scores ranging from 1 to 5 were created using Rasch modelling for both subscales with higher scores indicating higher proficiency.[209]

#### 3.1.2.3 *Adiposity*

Adiposity in LSAC was measured using BMI. Body weight and height were measured at age 6 to 7 years by LSAC staff during home interviews. Weight was measured in light clothing to the nearest 0.01 kg using a digital scale (Salter Australia, Code 79985) and height was measured without shoes to the nearest 0.10 cm using a portable rigid stadiometer (Invicta, Code IPO955). Two measurements of weight and height were taken. The averages of two weight and two height measurements were used. Age and sex-specific BMI z-scores in the LSAC sample were created based on the 2000 Centers for Disease and Prevention (CDC) Growth Charts.[210] The association between temperament and BMI z-score was presented Section 7.2.

#### 3.1.2.4 *Temperament*

At age 2 to 3 years, temperament characteristics of the LSAC sample was measured using the Short Temperament Scale for Toddlers (STST). The STST was developed from factor analyses of the Toddler Temperament Scale (TTS), which consists of 97 items in nine dimensions.[58] The TTS was first administered to a sample of 397 children aged 1 to 3 years old in the Australian Temperament Project. The content of the TTS was found to be suitable for use in Australian toddlers.[211] The STST used in LSAC consists of four items in each of the three dimensions (alpha=0.99):[113,212]

1. Approach - degree of comfort when encountering new situations or new people, for instance, 'pleasant (smiles, laughs) when first arriving in unfamiliar places' and 'smiles when an unfamiliar adult plays with him/her'.
2. Persistence - capacity to self-regulate and see tasks through to completion. Examples of persistence items included 'plays continuously for more than 10

minutes at a time with a favourite toy' and 'stops to examine objects thoroughly (5 minutes or more)'.

3. Reactivity - degree of intensity and emotional volatility. Items measured in reactivity dimension included 'responds to frustration intensely (screams, yells)' and 'reacts strongly (cries, screams) when unable to complete a play activity'.

The primary caregiver responded to each item from 1 (almost never) to 6 (almost always). A mean score was calculated for each temperament dimension. Higher scores indicate higher approach, persistence, and reactivity.

#### 3.1.2.5 *Parenting*

At age 0 to 1, parenting practices were measured through face-to-face interviews between trained professional interviewers and primary caregivers. At other ages, parenting measures were obtained through mothers' and fathers' self-completed questionnaires. Parenting dimensions measured in the infant cohort in LSAC include warmth, hostility, inductiveness, overprotectiveness, and self-efficacy. Although warmth and hostility dimensions are of interest in this thesis, the distribution of warmth and hostility dimensions in LSAC data were highly skewed (*e.g.* about 77% parents received the highest possible score in warmth dimension). While there has been continued research on the LSAC parenting measures,[85] the recommendations to improve measures were not available for use in this thesis.

Besides these general parenting dimensions, the LSAC contained several items related to parental involvement. At age 4 to 5, the primary caregivers were asked how often they

engaged in the following activities with their child: read to the child, tell stories, draw pictures or other craft activities, play with toys or games indoors such as board or card games, play music, involve child in daily activities such as cooking or pet care, and play outdoor games including walking or cycling. Items were rated from 0 (never) to 3 (everyday). Scores on these seven items were summed into a total parenting score that ranged from 0 to 21. These items were used as a measure of parenting practices in research question 2 (Chapter 5).

## 3.2 Methodological & analytical approach

The following section describes the methods and analyses used for each research question. This is also described in detail in Chapters 4 to 7.

### 3.2.1 Research question 1

***Are infant temperament norms derived from the US suitable for use in a cohort of UK infants?***

The first study presented in this thesis is a cross-sectional comparison of the published temperament norms for the Revised Infant Temperament Questionnaire (RITQ) and norms from the ALSPAC sample. The RITQ norms were standardized based on a sample of 203 US children in the 1970s, and have not been updated since. The appropriateness of these norms for UK children was unclear due to temporal and context differences. The RITQ comprised nine temperament subscales (described in Section 2.1.2.1). By following the scoring approach in the manual,[202] a score between 1 and 6 was generated for each of the nine temperament subscale. The mean and standard deviation (SD) of each temperament subscale derived from the ALSPAC sample was compared with the mean

and SD published in the RITQ manual. By following Carey's algorithm,[41] infants were categorised into easy, intermediate low, intermediate high and difficult temperament groups based on the ALSPAC norms and the norms in the RITQ manual. The proportion of infants in each group using the two different sets of norms was compared using a chi-squared test.

### 3.2.2 Research question 2

***Is temperament at age 2 to 3 years directly affecting children's cognitive (verbal and non-verbal) and academic (literacy and numeracy proficiency) development at age 6 to 7 years?***

The second study in this thesis used data from the LSAC to examine the controlled direct effect of temperament (reactivity, approach, and persistence) at age 2 to 3 on cognitive ability and academic achievement at age 6 to 7 years. In Figure 3.3, a directed acyclic graph (DAG) was used to depict the hypothesized effect of temperament and parenting practices on cognitive ability and academic outcomes. DAGs are commonly used in epidemiologic research to describe the causal relationships between variables. Variables in DAGs are linked by single headed arrows that represent the direct causal effect.[213] DAGs help to identify confounding and collider bias. Confounding is the bias resulting from the presence of common causes of exposure and outcome. Controlling for confounders is important so that the only difference in outcome between the exposed and the unexposed groups is due to the exposure history.[214] On the other hand, a collider is a common effect of two or more variables. Adjusting for a collider may introduce bias as it induces an association between the exposure and the outcome even if there is no

true association between the exposure and the outcome.[121] In this thesis, all DAGs were drawn using the open-access DAGitty 2.3 software developed by Textor *et al*.[215]

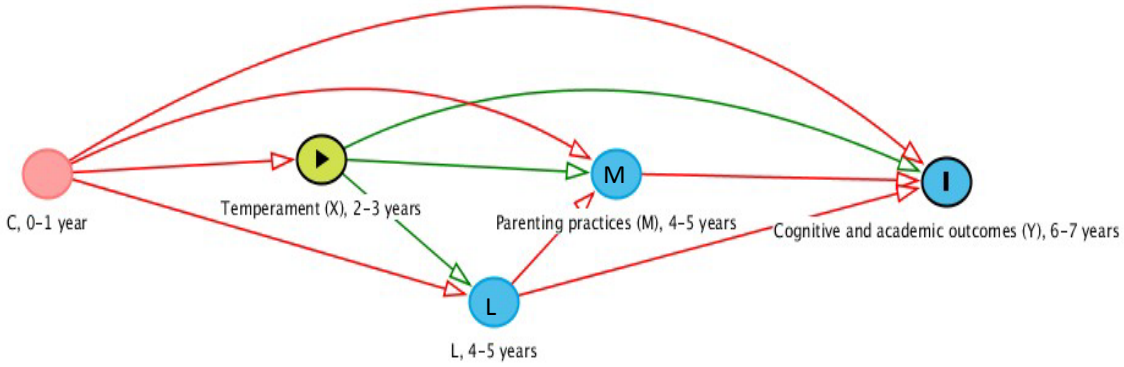


**Figure 3.3 DAG of the hypothesized effect of temperament and parenting pratices on cognitive and academic achievement**

**Exposure**: Temperament reactivity, approach, and persistence. **Outcome**: verbal (PPVT) and nob-verbal (Matrix Reasoning) cognitive abilities, literacy and numeracy (ARS). M **Intermediate variable**: parenting practices **Confounders of exposure and outcomes**: maternal education, financial hardship, housing tenure, Aboriginal or Torres Strait Islander status, neighbourhood disadvantage, sex, birth weight for gestational age z-score, duration of breastfeeding, maternal age, maternal country of birth, maternal psychological distress, mother and partner argumentative relationship, single-parent family, gestational hypertension, gestational diabetes, smoking and alcohol intake during pregnancy. L **Confounders of parenting and outcomes**: maternal psychological distress, number of siblings, mothers' working status, household income, and financial hardship Causal path. Biasing path.

The common approach to estimate the direct effect of an exposure (*X*) on an outcome (*Y*) is fitting a regression model of the form:

$$Y = \beta_0 + \beta_1 X + \beta_2 M + C$$

In this model, outcome, *Y* is regressed on the exposure, *X,* intermediate variable, *M,* and some confounding factors, *C*. $\beta_1$ is taken as the direct effect of *X* on *Y*. However, this approach suffers from limitations when there is confounding of the intermediate variable and the outcome, *L* (Figure 3.3).[216] Conditioning on *M* would induce an association between *X* and *L*, opening a backdoor path from *X*⟶*L*⟶*Y*. A backdoor path is a non-causal path from *X* to *Y*.[213] On the other hand, because *L* is influenced by *X*, conditioning on *L*

would block the path from $X \rightarrow L \rightarrow Y$, which would also block part of the effect of *X* on *Y* that is not mediated through *M*. Conditioning on *L* would also induce collider bias because *L* is a common effect of *C* and *X*.[121]

Another approach used in psychological studies to examine direct effects is path analysis, which is part of structural equation modelling (SEM). SEM allows the estimation of the direct and indirect effect, but it has been criticised for not adequately addressing issues of confounding in making causal inferences.[217] In addition, the SEM approach assumes that all the associations between the variables in the model are linear, it does not account for any nonlinearities or interactions between any of the variables in the model.[123]

To overcome the limitations of the traditional regression approach and SEM, a marginal structural model with inverse probability of treatment weights (IPTW) was used. Marginal structural models are based on the potential outcome or counterfactual approach, which is the outcome that would have been observed had the value of the exposure, *x* been set to a counterfactual level of *x**. For example, if the exposure is binary, *i.e.* a child has either a high or low level of temperamental persistence, for a child who has high persistence their potential outcome can be defined as what would be the outcome if the child had low persistence and vice versa.

Marginal structural models handle parenting practices as an intermediate variable by using IPTW rather than "adjusting" for variables in regressions. IPTW predicts the probability of each individual receiving the level of exposure (*i.e.* temperament) conditioned on past confounding factors and then each individual is weighted by the

inverse of this conditional probability.[124,218] The weighting method creates a pseudo population in which the association between the exposure and the measured confounding factors is removed, and is analogous to direct standardisation of the exposure.[125,219] Two sets of weights were created, one accounted for the confounding between temperament and cognitive ability and another one accounted for the confounding of the parenting and cognitive ability association.[124,220] A stabilised weight was generated by multiplying the two sets of weights together. To avoid excessively upweighing some individuals, weights were truncated at the 1st and the 99th percentile.[219]

The creation of weights was based on four assumptions:[219]

1. Consistency - the potential outcome for every individual depends on his/her exposure history.
2. Conditional exchangeability - the outcome *Y* is independent of the exposure *X,* given covariates. This assumption also known as 'no unmeasured confounding'. The conditional exchangeability is not testable in observational data but could be explored using sensitivity analyses.
3. Positivity - both exposed and unexposed individuals are present at every level of the confounders.
4. Correct model misspecification - the model used to create the weights was appropriate (*e.g.* linear relationship, interaction term included if appropriate, and sufficient confounding factors). Stabilised weights with a mean of one and small range indicate no model misspecification.[219]

A marginal structural model was used to estimate the controlled direct effects of temperament (reactivity, approach, and persistence) on cognitive and academic outcomes after accounting for parenting practices as a mediator/intermediate variable. The controlled direct effect compares the effect of temperament at level *x* to a counterfactual level of *x** by setting the value of the intermediate variable *M* to *m*. In addition to the four assumptions above, two more assumptions are required to examine the controlled direct effect: 1) no unmeasured confounding of the effect of exposure on outcome, 2) no unmeasured confounding of the effect of intermediate variable on outcome.

To examine whether there was any unmeasured confounding of the parenting-outcomes relationship that is present to explain away the controlled direct effects of temperament on cognitive and academic outcomes, sensitivity analyses was conducted. Following VanderWeele,[221] the bias for the controlled direct effect was estimated using two parameters: $\delta$, denotes the difference in the prevalence of the unmeasured confounder between the exposed and the counterfactual group; $\gamma$, denotes the effect size of the binary unmeasured confounder on the outcomes. The bias in the estimated controlled direct effect was then estimated as the product of $\delta$ and $\gamma$, under several conditions with plausible level of $\delta$ and $\gamma$.

### 3.2.3 Research question 3

***Does infant temperament modify the effects of parenting practices (warmth and control) at age 24 to 47 months on IQ at age 8 years?***

Study 3 examines the associations between parenting warmth and control at age 24 to 47 months and cognitive ability (IQ) at age 8 years, adjusting for infant temperament at 6 months (created from Study 1) and other confounding factors. Effect-measure modification by infant temperament of the associations between parenting warmth, control, and IQ was explored.

*Latent class analysis*

Eight items on parental warmth and nine items on parental control across 24 to 47 months were selected from the ALSPAC questionnaires based on definitions in the literature. Items included in 'warmth' were related to the use of acceptance (*e.g.* kissing and hugging) and hostility (*e.g.* slapping and shouting). Items included in 'control' were related to the use of reasoning and autonomy (*e.g.* allowing choices). Children who were given choices may be more likely to have better emotional control and less likely to have behavioural problems.[222] While parental control may also include behavioural control such as setting guidelines and limits, these items were not available in the ALSPAC dataset and therefore could not be included. To identify classes of parents with distinct characteristics, latent class analyses (LCA) were undertaken. LCA is a model-based cluster analysis. LCA has been widely used in a variety of health, social and behavioural sciences including parenting,[223] temperament,[224] substance use,[225] and health-risk behaviours to identify individuals with distinct profiles.[226] LCA has the advantages over other traditional cluster analysis methods such as *k*-mean cluster analysis and hierarchical cluster analysis

in that it takes into account the uncertainties when allocating individuals into classes and it is based on the Expectation Maximum (EM) algorithm which allows the handling of incomplete data in a sensible way.[227-229]

LCA were performed separately for parental warmth and control (Figure 3.4). The LCA is based on the underlying theory that the latent variable (parenting warmth or control) is unobservable but can be indicated from multiple observed items or indicators. The latent class models assume that the items used as indicators within a latent class are independent of each other, *i.e.* the "local independence" or "conditional independence" assumption.

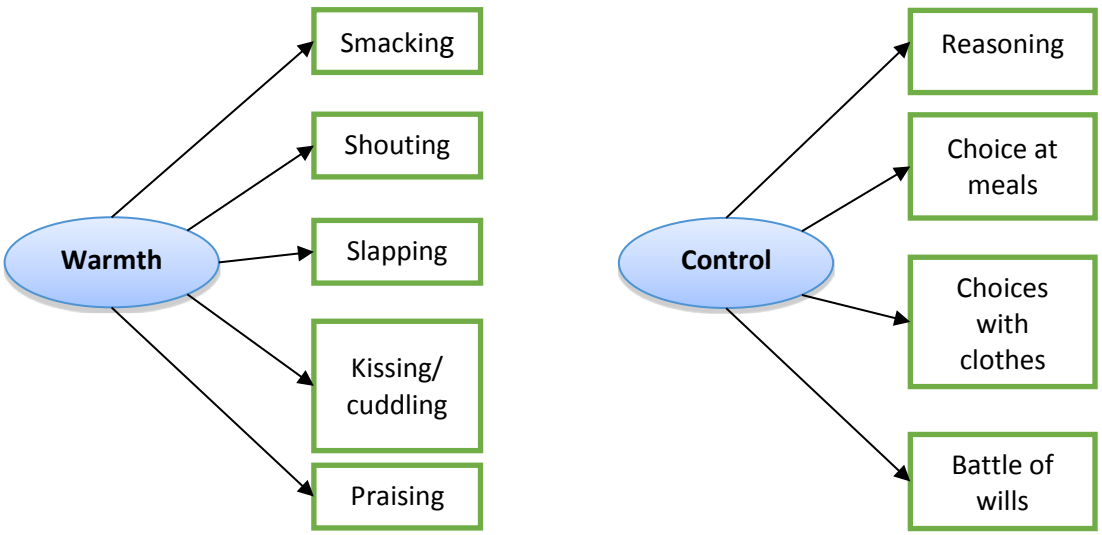


**Figure 3.4 Latent class model of parenting warmth and control. Latent variable (oval), latent class indicators (rectangles), associations between latent variable and indicators (arrows).**

The latent classes were created based on two parameters: the gamma parameter (γ) which represents the latent class membership probabilities and the rho parameter (ρ)

which represents the conditional item-response probabilities.[230] The latent class membership probabilities estimate the proportion of the population that fall into a given class. The conditional item-response probabilities are probabilities of responding to each item, given the class membership.[227] The relations between the observed items and the latent variable were indicated from the item-response probabilities. Item-response probabilities near 0 or 1 represent a strong association between the observed items and the latent construct.[227] At least two categorical items were needed as indicators of the latent class model. Individuals who responded to at least one of the items were assigned to a class. Individuals were assigned to the class that they had the highest probability of membership.

Latent class models with different number of classes (two, three, four, and five) were tested. Model selection was based on several criteria such as log-likelihood, Bayesian Information Criterion (BIC), Aikaike Information Criterion (AIC), and entropy. BIC and AIC are penalized log-likelihood model information criteria to assess model fit. A smaller AIC and BIC indicates the trade-off between model fit and parsimony is preferable.[231] Entropy refers to the precision of membership assignment across all individuals. Entropy of 1 indicates perfect assignment of all individuals to all classes. Model interpretability, that is, how each class differs from the others on the basis of the item-response probabilities, was also considered. Two underlying classes were found for warmth dimensions, labelled as 'high warmth' and 'low warmth'. Both 'high' and 'low' warmth classes were characterised by frequent cuddling and kissing, but the 'high' warmth class involved less shouting, slapping and smacking as opposed to the 'low' warmth class. A two-class solution was chosen for the parental control dimension, labelled as 'high control' and 'low

'control'. High' parental control class was characterised by less reasoning, more often battle of wills, and less choices for clothing and food than the 'low' parental control class. Latent class analyses were performed using the PROC LCA command in SAS.[231,232]

*Effect measure modification*

Research question 3 aimed to examine the effect-measure modification by temperament of the associations between parenting warmth and IQ and parenting control and IQ. Effect-measure modification concerns whether the effect of a primary exposure on an outcome differs across strata of a secondary exposure.[135,137] In this study, parenting warmth and control are two primary exposures of interest, and temperament is the secondary exposure. Effect-measure modification examines whether the effect of parenting warmth and control on IQ differ across different temperament subgroups (*i.e.* easy versus difficult).

The terms "effect-measure modification" and "interaction" are often used interchangeably.[217] However, there are some distinctions between effect-measure modification and interaction which are important especially when considering potential intervention and policy recommendations. Only intervention on the primary exposure is considered in effect-measure modification, while in interaction, intervention on both primary and secondary exposures is considered. To examine effect-measure modification, adjustment for one set of confounding factors (*i.e.* confounding factors of the primary exposure-outcome association) is sufficient, however, to examine interaction, adjustment for two sets of confounding factors (*i.e.* the confounding factors of the primary exposure-

outcome association and the confounding factors of the secondary exposure-outcome association) is required.

Traditionally, effect-measure modification and interaction are tested in regression analyses by including a product term. However, it is not clear whether the coefficient of the interaction term should be interpreted as effect-measure modification or interaction, or both, or neither.[135] A number of epidemiological researchers have argued that the presentation of effect-measure modification and interaction is inadequate.[136,137] In epidemiological studies where logistic regressions are commonly used, effect-measure modification and interaction are often reported on the risk-ratio scale. However, modern epidemiology research has suggested that the risk-difference scale is of greater policy relevance.[217] For example, if the effect of the primary exposure is greater in one subgroup than the other, then intervention on the subgroup with a greater risk may be considered. The presence or absence of effect-measure modification and interaction also depends on which scale (risk-difference or risk-ratio) it was assessed. Essentially, if both exposures have causal effects on the outcome, the absence of effect-measure modification on the risk-difference scale implies the presence of effect-measure modification on the risk-ratio scale, and vice versa.[233] In other words, if both exposures have causal effects on the outcome, and there is no effect on one scale, then mathematically, there must be an effect on the other scale. For these reasons, modern epidemiology research has recommended that effect-measure modification and interaction to be examined using additive (risk-difference) and multiplicative (risk-ratio) scales.[137]

This study examines effect-measure modification by temperament of the association between parenting warmth and IQ and parenting control and IQ using the approach outlined by Knol and VanderWeele.[138] Effect measure modification on the risk-difference scale is examined using the Relative Excess Risk due to Interaction (RERI). RERI greater than 0 indicates a positive effect-measure modification (the effect of the parenting and temperament operating together is greater than the individual effect of each added together), RERI less than 0 indicates a negative effect-measure modification (the effect of the parenting and temperament operating together is less than the individual effect of each added together), RERI of 0 indicates the absence of effect-measure modification on the risk-difference scale.[217] Effect-measure modification on the risk-ratio scale was examined using the ratio of relative risks. In the absence of effect-measure modification on the risk-ratio scale, the ratio of relative risks equals 1. This approach is recommended in epidemiological methodology to fully explore the effect-measure modification.[137,138,217]

### 3.2.4 Research question 4

***Is temperament at age 0 to 5 years associated with children's BMI at age 7 years? Are parental feeding control and use of food to soothe at age 3.5 years associated with BMI z-scores and fat mass in childhood (7 years) and adolescence (15 years)? Does infant temperament modify the associations between feeding practices and BMI?***

Research question 4 is divided into three parts. In the first part, the association between temperament and BMI was examined using data from both ALSPAC and LSAC. Linear regression models were used to estimate the total effect of temperament dimensions (*e.g.* approach, mood) from age 0 to 5 years on BMI at age 7 years, adjusted for a range of confounders.

The second part of research question examines the associations between parental feeding control and using food to soothe at age 3.5 years) on children's BMI z-scores at age 7 and 15 years and fat mass at 9 years using data from the ALSPAC. As shown in Figure 3.5 and Figure 3.6, DAG were used to demonstrate the *a priori* assumptions regarding the causal associations among the exposure, outcomes and covariates. Linear regression models were used to estimate the total effect of parental feeding control and use of food to soothe on adiposity outcomes, adjusting for a range of confounders.

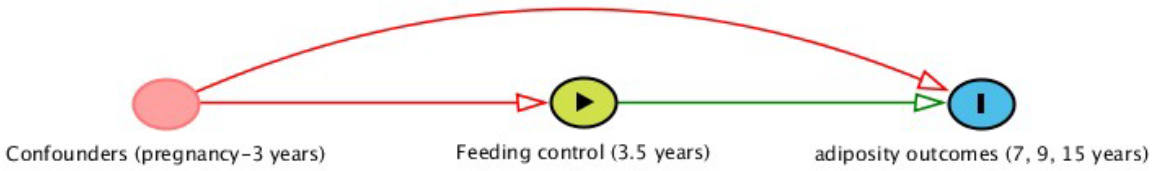


**Figure 3.5 DAG representing the effect of feeding control on child adiposity outcomes**

**Exposure**: Parental feeding control. **Outcome**: Age and sex-specific BMI z-scores at 7 and 15 years, fat mass at 9 years. **Confounders of exposure and outcome**: child temperament, dietary patterns at 3 years, birth weight z-score, eating problems (not eaten enough, refused right food, choosy with food, difficulty to establish eating routine), maternal characteristics (age, BMI, social class, home ownership, household crowding, financial difficulties, sole parenting, depression, smoking, alcohol consumption, number of children, ethnicity) and food to soothe at 3.5 years. Causal path. Biasing path.

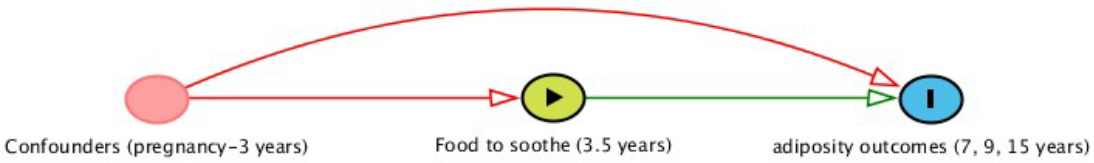


**Figure 3.6 DAG representing the effect of food to soothe on child adiposity outcomes**

**Exposure**: Parental use of food to soothe. **Outcome**: Age and sex-specific BMI z-scores at 7 and 15 years, fat mass at 9 years. **Confounders of exposure and outcome**: child temperament, dietary patterns at 3 years, birth weight z-score, eating problems (not eaten enough, refused right food, choosy with food, difficulty to establish eating routine), maternal characteristics (age, BMI, social class, home ownership, household crowding, financial difficulties, sole parenting, depression, smoking, alcohol consumption, number of children, ethnicity) and feeding control at 3.5 years. Causal path. Biasing path.

The third part of research question 4 examines whether the association of parental feeding control with BMI and fat mass differs for children with different temperaments using data from ALSPAC. An analysis of effect-measure modification was conducted using the approach outlined by Knol and VanderWeele.[138]

## 3.1 Dealing with missing data

Missing data is a common problem in longitudinal studies due to participants' attrition and nonresponse to items, questionnaires, or clinic visits.[234] The mechanisms of data being missing can be classified into three types: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR).[235] Missing completely at random means that the probability of data being missing does not depend on observed or unobserved data.[236] Missing at random means that the probability of data being missing depends on the observed data, but not the unobserved data. Missing not at random means that the probability of being missing depends on the unobserved and observed data.

Traditionally, methods used to deal with missing data include complete-case analysis (list wise deletion) and single imputation. In complete case analyses, only participants with complete data are included in the analyses. Complete-case analysis may cause biased estimates as it assumes that data are MCAR.[237] Single imputation strategies such as mean imputation, fail to account for the uncertainties about the missing values and often results in standard errors that are too small.[238] To overcome the limitations by complete-case analysis and single imputation, multiple imputation is recommended as a better way to deal with missing data.[234]

Multiple imputation takes into account the uncertainty in missing values by creating multiple copies of the dataset in which missing values are replaced by imputed values.[234] There are two common types of multiple imputation - multivariate normal imputation (MVNI) and multiple imputation using chained equation (MICE). The MVNI replaces missing values of variables using multivariate normal regressions. The MVNI is useful when variables are normally distributed and there are few variables to be imputed but have some issues when imputing variables that are not normally distributed such as binary and categorical variables.[239] MICE is more flexible as it allows imputation of different variable types, for example linear regression, predicted mean matching, or truncated normal regression are used to impute continuous variables (*e.g.* PPVT, IQ, BMI), ordered logistic regression is used to impute ordinal variables (*e.g.* maternal education), and logistic regression is used to impute binary variables (*e.g.* smoking yes/no).[240] MICE is useful when there are different types of variables and when there are many missing values. Nevertheless, a study showed that both MVNI and MICE produce similar results even in the presence of categorical variables.[239] Multiple imputation by chained equation

(MICE) was used in research question 2, 3, and 4 to deal with missing data. Variables included in the multiple imputation model include all variables in the analysis model (exposures, outcomes, covariates) and additional auxiliary variables that are not part of the analysis model but were included in the imputation model to help predict the missing values. Interaction terms (*e.g.* temperament*parenting) were also included in the imputation model because they will be tested in the subsequent analysis models.[236]

MICE was performed using the 'mi impute chained' command implemented in STATA. Twenty imputed datasets were generated with fifty cycles of "regression switching".[240,241] In each chained equation cycle, missing values in each variable were imputed based on a predictive distribution derived from a regression on all other variables in the imputation model. At the end of 50 cycles, one imputed dataset was generated. The process was repeated 20 times to generate 20 imputed datasets. The overall estimates were obtained by averaging the results from each of these 20 datasets using Rubin's rules.[235] The procedure takes into account both within-imputation variability (uncertainty in the one imputed dataset) and between-imputation variability (uncertainty due to missing value).

There have been some discussions about whether or not individuals with imputed outcomes should be excluded from analysis. Von Hippel[242] suggested to use the "multiple imputation then deletion" method where all individuals were used for imputation, however, following imputation, individuals with missing outcome were excluded from the analysis. It was suggested that individuals with missing outcome provided no any further information to improve regression estimates. Retaining these individuals in the analysis only adds noise to the estimation process. However, the "multiple imputation, then

deletion" approach has been questioned by a later publication by Sullivan *et al.*[243] who recommended to use the "multiple imputation" approach - individuals with imputed outcomes were retained in the analysis. "Multiple imputation, then deletion" approach was used in research question 3 because it has conducted before findings of the Sullivan *et al.* paper were available. The "multiple imputation" approach was used in research question 2 and 4.

# CHAPTER 4. How many infants are temperamentally difficult? Comparing norms from the Revised Infant Temperament Questionnaire to a population sample of UK infants

## 4.1 Preface

This chapter contains the first of four articles contributing to this thesis. This article was published in *Infant Behavior and Development* in 2015.[244]

Temperament researchers face some difficulties in comparing and classifying children's temperament across countries and populations. One of the reasons is that many temperament scales do not have published norms, and where they do exist, they are often not based on large representative samples of children. This article focuses on one of the most frequently used temperament questionnaires – the Revised Infant Temperament Questionnaire (RITQ) – and examines whether the norms published in the manual are suitable for use with a large cohort of UK infants born in the 1990s. Despite its popularity for use in paediatric practices and research, the norms for the nine subscales of the RITQ were developed based on a small, non-representative sample of US infants in the 1970s. This study addresses a significant gap in the literature by comparing norms from the RITQ with norms derived from a large, representative population sample of UK infants. Scores on each of the nine temperament sub-scales and the proportion of infants in each of the temperament categories (*i.e.* 'difficult', 'easy') were compared using the two different sets of norms. Based on the findings of this paper, norms from the UK

population rather than the manual were used to categorise children's temperament in the two subsequent research questions in this thesis that use the RITQ (question 3 in Chapter 6) and question 4 in Chapter 7).

The updated norms published from this study will be useful for UK paediatricians and clinicians to assess children's temperament for primary care purposes, and may be useful for temperament researchers to compare infant temperament across time, contexts and cultures.

## 4.2 Publication: How many infants are temperamentally difficult? Comparing norms from the Revised Infant Temperament Questionnaire to a population sample of UK infants

### 4.2.1 Statement of Authorship

| | |
|---|---|
| Title of Paper | How many infants are temperamentally difficult? Comparing norms from the Revised Infant Temperament Questionnaire to a population sample of UK infants |
| Publication Status | ☑ Published<br>☐ Accepted for Publication<br>☐ Submitted for Publication<br>☐ Unpublished and Unsubmitted w ork w ritten in manuscript style |
| Publication Details | Chong SY, Chittleborough CR, Gregory T, Lynch JW, Smithers LG. How many infants are temperamentally difficult? Comparing norms from the Revised Infant Temperament Questionnaire to a population sample of UK infants. *Infant Behav Dev*. 2015;40:20-28. |

### Principal Author

| | | | |
|---|---|---|---|
| Name of Principal Author (Candidate) | Shiau Yun Chong | | |
| Contribution to the Paper | Conceived and designed the study, analysed and interpreted the data, drafted the manuscript. | | |
| Overall percentage (%) | 85 | | |
| Certification: | This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper. | | |
| Signature | | Date | 15-08-2016 |

### Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

| | | | |
|---|---|---|---|
| Name of Co-Author | Catherine Chittleborough | | |
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| | | | |
|---|---|---|---|
| Name of Co-Author | Tess Gregory | | |
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| | | | |
|---|---|---|---|
| Name of Co-Author | John Lynch | | |
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| | | | |
|---|---|---|---|
| Name of Co-Author | Lisa Smithers | | |
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

### 4.2.2 Abstract

The original norms for the Revised Infant Temperament Questionnaire (RITQ) were published in 1978 and were based on a small sample from the US. The aim of this study is to compare temperament scores from the original RITQ against scores from a large population-based cohort of infants from the UK. This study consists of 10937 infants from the Avon Longitudinal Study of Parents and Children (ALSPAC) born between April 1991 and December 1992 in the southwest of England. Infant temperament at 6 months of age was reported by parents using the adapted RITQ. Responses were scored according to the RITQ manual and then categorised into temperament groups (easy, intermediate low, intermediate high, and difficult) using either the RITQ norms or norms derived from the data. The scores for each temperament subscale and the proportion of children in each temperament group were compared across the two methods. Subscale scores for the ALSPAC sample were higher (more "difficult") than the RITQ norms for rhythmicity, approach, adaptability, intensity, and distractibility. When RITQ norms were applied, 24% infants were categorised as difficult and 25% as easy, compared with 15% difficult and 38% easy when ALSPAC norms were used. There are discrepancies between RITQ norms and the ALSPAC norms which resulted in differences in the distribution of temperament groups. There is a need to re-examine RITQ norms and categorization for use in primary care practice and contemporary population-based studies.

## Highlights

- RITQ norms from the US differ from a UK population cohort
- Different norms alter proportions of infants categorised as difficult
- Contemporary norms derived from relevant representative samples are needed

### 4.2.3 Introduction

Infant temperament assessment is recommended for health practitioners including paediatricians, physicians, and paediatric nurse practitioners, as part of their routine screenings. The Infant Temperament Questionnaire (ITQ) is one of the well-established tools for assessment of infant's temperament. The ITQ was published in 1970 by the paediatrician, Dr William Carey[41] based on results published by Thomas, Chess and colleagues from the New York Longitudinal Study (NYLS).[27] Thomas and Chess identified nine dimensions of temperament from extensive observations and qualitative interviews with the parents of 22 children in the NYLS and these form the nine subscales of temperament in the ITQ (activity, rhythmicity, adaptability, approach, intensity, mood, persistence, distractibility, and threshold). The ITQ was revised in 1978 by Carey and McDevitt, and the Revised Infant Temperament Questionnaire (RITQ) has shown moderate internal consistency (0.49 to 0.71 for subscales, 0.83 for composite) and good test-retest reliability (0.66 to 0.81).[42]

The RITQ is used extensively by health practitioners in primary care settings[40,245,246] as it is clinically derived and is useful for identifying childhood clinical conditions such as behavioural problems.[245,247,248] The concept of "goodness of fit" introduced by Thomas and Chess is particularly useful in clinical interventions to help parents understand the importance of the consonance between the child's temperament and the expectations of the parents to the development of the child.[40] Parenting advice on how to manage the child's behaviour can then be given according to the temperament profile of the child.[249,250] For example, in *Bright Futures in Practice: Mental Health*, a set of paediatric guidelines for promoting socio-emotional wellbeing of children from birth through

adolescence, the RITQ temperament subscales were explained and strategies on how to improve the "fit" for the children were provided.[251] A number of parent education materials have also been developed based on the RITQ temperament subscales.[252,253]

The RITQ is also used in many large-scale longitudinal studies such as the Millennium Cohort Study,[254] the National Institute of Child Health and Human Development Study of Early Child Care,[255] and the Helsinki Longitudinal Temperament Study.[256] Longitudinal studies have provided evidence that temperament is associated with later development of mental disorder,[257,258] behavioural problems,[3] as well as cognitive,[140] language,[259] and academic performance outcomes.[260] Assessing infant temperament in large-scale community or population samples helps identify infants who may be at increased risk for later cognitive, academic or behavioural problems and may assist policy makers to better target groups of children for interventions. For example, parenting programs may be considered to provide targeted support to families with temperamentally difficult children.

As part of the RITQ, a profile sheet provides means and standard deviations (SDs) for each subscale that can be used to identify an infant's temperament profile.[202] These means and SDs were derived from a standardization on 203 infants (104 boys and 99 girls) aged 4- to 8-months old, predominantly from middle-to-upper class US families in 1978.[42] When examining an infant's temperament, clinicians compare the infant's scores on each temperament subscale with the normative scores on the RITQ defined by the 1978 sample. Infants can also be categorised into different temperament groups (easy, intermediate low, intermediate high, and difficult) based on where their scores sit in

relation to the normative sample. This categorization is then used for subsequent investigations and interventions, and parenting advice can be provided to parents according to the temperament profile of their child.

Some studies have used their own sample norms to categorize infants into temperament groups. For instance, in a study of 985 infants in the United States, infant temperament was categorised into three groups (easy, average, and difficult) using the study sample means and SDs as cut-offs.[255] Other studies have used the RITQ norms to categorize infants into temperament groups but have shown significant differences between the means and SDs observed in their sample and those established based on the 1978 sample. For example, in a study of 349 infants aged 4 to 8 months in Taiwan, RITQ items were translated into Chinese and then translated back into English. The study found that infants scored significantly higher than the RITQ standardization sample in approach, adaptability, mood, intensity, distractibility, and threshold.[76] Another study using a Japanese version of the RITQ had means that were higher than the RITQ means on all subscales except activity and threshold.[261] However, the Japanese version of the RITQ has not been back translated, so it is difficult to tease apart whether the differences are due to translation or context-specific perceptions of difficult temperament. While using different means and SDs could potentially result in inaccurate identification of infant temperament, previous research using the RITQ has not addressed this issue.

As stated in the manual,[202] RITQ norms published based on the standardization sample may not apply to other populations. Norms may vary across cultures and populations. However, to date, there are limited studies with large, representative samples that have

published population-specific norms using the RITQ items. This is a problem for RITQ users because this could lead to misclassification of temperament if population-specific norms differ from the RITQ norms.

This study aims to compare the original norms (means and SDs) published for the RITQ using the 1978 sample[42] with norms derived from a large population sample of UK infants, and the resulting categorization of temperament (easy, intermediate low, intermediate high, and difficult) from these two different norms.

### 4.2.4 Methods

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a geographically based prospective study investigating influences on children's health and development. A total of 14541 pregnant women who resided in the Southwest of England, with expected delivery date between 1st April 1991 and 31st December 1992 were recruited to ALSPAC. The 14541 pregnancies represent 72% of the eligible pregnancies in the region during this period.[197] The ALSPAC sample is broadly representative of the population living in Avon at the time although ethnic minorities and unmarried couples were slightly underrepresented when compared with the 1991 National Census based on women with an infant of less than 1 year of age.[197] Ethical approvals were obtained from the ALSPAC Law and Ethics committee and local Research Ethics committees. Informed consent was obtained from participants.

Temperament was reported by mothers when their infant was 6 months old using an adaptation of the RITQ.[202] RITQ is a valid and reliable measure of temperament.[6,42] The

adapted RITQ used in ALSPAC comprised 88 of the 95 original questions in the RITQ. There were 12 items in activity (1 omitted), 10 items in rhythmicity (2 omitted), 10 items in approach (1 omitted), 10 items in adaptability (1 omitted), 10 items in intensity (no item omitted), 9 items in mood (1 omitted), 7 items in persistence (1 omitted), 10 items in distractibility (no item omitted), and 10 items in threshold (no item omitted). Items were omitted as they were shown to have non-response of at least 10% in a pilot test prior to ALSPAC because they were considered by parents as irrelevant to the child. For example, a question on the baby's reactions when having an infection was irrelevant to those mothers whose baby had never had an infection, and hence this question was omitted. The internal consistency of the 88-item RITQ ranged from 0.40 to 0.73 for subscales, and 0.79 for the composite score. This is consistent with the internal consistencies of the original questions included in the RITQ.[42]

Caregivers rated their perceptions about the infant's temperament on the nine subscales from 1 (almost never) to 6 (almost always). Scores for each subscale were derived according to the standard scoring procedure introduced by Carey for the RITQ.[202] A higher score on each subscale indicated a more "difficult" temperament. Following Carey,[202] participants with more than 20% missing items were excluded from analysis. T-tests were used to compare the differences between the RITQ norms and ALSPAC norms. Analyses were conducted using STATA version 12.0 (Stata Corp, College Station, Texas).

By following the categorisation algorithm suggested by Carey,[41] the study sample were categorised into easy, intermediate low, intermediate high, and difficult according to their scores on five subscales: rhythmicity, approach, adaptability, intensity and mood (Figure

4.1). Infants were categorised into these four temperaments using the means and SDs from RITQ and the ALSPAC sample, and the proportion of infants in each group were compared using chi-squared tests.

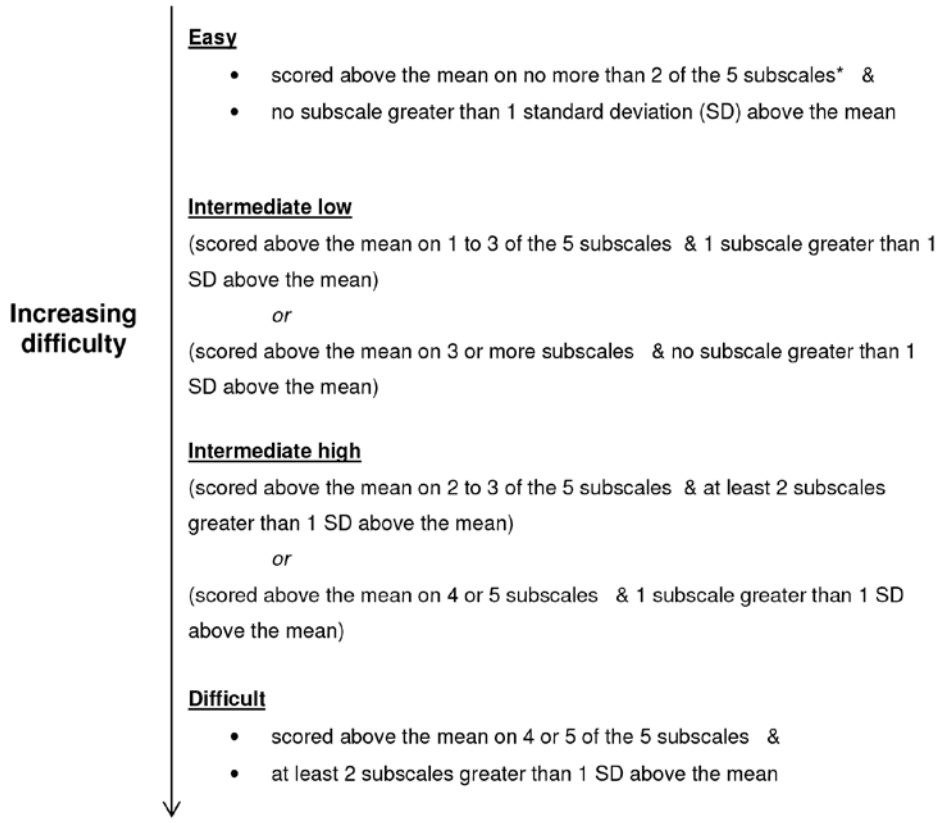


**Figure 4.1 Categorisation algorithm for each temperament group.**

The ALSPAC sample was randomly divided into four groups to examine the stability of the subscale means, SDs and categorisation.

### 4.2.5 Results

Table 4.1 shows the socio-demographic characteristics of the response sample and the sample who had complete data on all nine temperament subscales (n=10937). Participants with complete temperament data had similar characteristics as the response

sample. Of the infants with complete temperament data, 14% had a mother with a degree or higher education qualification, and 24% had a mother with an A level. About 57% of the sample had at least one parent who was in a professional or managerial role. The majority (96%) of infants were White. The average age of mothers was 28 years, 9% had financial difficulties, and 19% had depression. Approximately 5% of mothers reported living in crowded households.

**Table 4.1 Socio-demographic characteristics of the Avon Longitudinal Study of Parents and Children (ALSPAC) sample**

| | Response sample | | Temperament data available sample (*N*=10937) |
|---|---|---|---|
| | ***N*** | **Mean (SD) or %** | **Mean (SD) or %** |
| ***Infant characteristics*** | | | |
| Birth weight (grams) | 13978 | 3392.0 (559.3) | 3414.2 (548.9) |
| Gestational age (weeks) | 13976 | 39.4 (1.9) | 39.5 (1.8) |
| Ethnicity, non-white | 12083 | 5.0 | 4.1 |
| Sex, female | 13976 | 48.3 | 48.4 |
| ***Parent characteristics*** | | | |
| Maternal education[a] | | | |
| None / CSE / vocational | 3728 | 30.0 | 27.0 |
| O level | 4296 | 34.6 | 35.5 |
| A level | 2794 | 22.5 | 23.9 |
| Degree or higher | 1600 | 12.9 | 13.7 |
| Paternal education | | | |
| None / CSE / vocational | 4124 | 34.5 | 31.8 |
| O level | 2540 | 21.3 | 21.7 |
| A level | 3105 | 26.0 | 27.1 |
| Degree or higher | 2171 | 18.2 | 19.3 |
| Parental highest social class | | | |
| Professional / managerial | 6342 | 55.1 | 56.9 |
| Skilled manual / non-manual | 4481 | 38.9 | 37.9 |
| Unskilled / semiskilled manual | 682 | 5.9 | 5.2 |
| Maternal age (years) | 13978 | 28.0 (5.0) | 28.4 (4.8) |
| Financial difficulties[b] | 12088 | 10.0 | 9.1 |
| Maternal depression[c] | 10929 | 10.2 | 9.5 |
| Household crowding, >1 person/room | 12084 | 6.9 | 5.3 |

*Note*. CSE= Certificate of Secondary Education

[a] Maternal and paternal education was reported as the highest completed level. CSE, O- and A-levels are completed at secondary school. O-levels are usually studied at age 16 and A-levels at age 18.

[b] Financial difficulties were measured using five items asking how difficult the mother found it to afford food, clothing, heating, rent or mortgage, and things they will need for their babies. A total score of 0 represented no financial difficulties and 15 represented maximum financial difficulties. Mothers with a score of ≥9 were defined as experiencing financial difficulties.

[c] Maternal depression was assessed using the Edinburgh Postnatal Depression Scale at 8 weeks postpartum. Mothers with a score of ≥12 were considered as displaying depressive symptoms.

Table 4.2 shows the mean and SD scores for each of the nine temperament subscales among ALSPAC infants, alongside the mean and SD scores published using the original RITQ sample.[42] ALSPAC infants had higher mean scores than the RITQ sample on the rhythmicity, approach, adaptability, intensity, and distractibility subscales. Of the five subscales used for categorizing temperament, three (rhythmicity, approach, and adaptability) showed mean differences from the RITQ norms. The largest difference of 0.40 (95% CI 0.32, 0.48) was observed for adaptability with children in the ALSPAC sample showing significantly lower adaptability than children in the RITQ sample. Examples of items in adaptability subscale are “adjust within 10 minutes to new surrounding” and “still wary of strangers after 15 minutes”.

**Table 4.2 Comparison of temperament subscale means and SDs (norms) in the Revised Infant Temperament Questionnaire (RITQ) and Avon Longitudinal Study of Parents and Children (ALSPAC) samples**

| Temperament subscale | RITQ norms (US)[42], *N*=203 | | ALSPAC norms (UK), *N*=10937 | | Mean difference (95% CI) |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | |
| Activity (1=low; 6=high) | 4.40 | 0.56 | 4.35 | 0.53 | -0.05 (-0.12, 0.02) |
| Rhythmicity (1=regular; 6=irregular) | 2.36 | 0.68 | 2.60 | 0.69 | 0.24 (0.14, 0.34) |
| Approach (1=approach; 6=withdraw) | 2.27 | 0.78 | 2.53 | 0.64 | 0.26 (0.17, 0.35) |
| Adaptability (1=high; 6=low) | 2.02 | 0.59 | 2.42 | 0.57 | 0.40 (0.32, 0.48) |
| Intensity (1=mild; 6=intense) | 3.42 | 0.71 | 3.51 | 0.56 | 0.09 (0.01, 0.17) |
| Mood (1=positive; 6=negative) | 2.81 | 0.68 | 2.78 | 0.66 | -0.03 (-1.12, 0.06) |
| Persistence (1=high; 6=low) | 3.03 | 0.82 | 2.95 | 0.73 | -0.08 (-0.18, 0.02) |
| Distractibility (1=high; 6=low) | 2.23 | 0.60 | 2.42 | 0.56 | 0.19 (0.11, 0.27) |
| Threshold (1=high; 6=low) | 3.79 | 0.76 | 3.75 | 0.61 | -0.04 (-0.12, 0.04) |

*Note*. Subscales used for categorisation are rhythmicity, approach, adaptability, intensity, and mood

Table 4.3 shows that the proportion of infants in each temperament group differed depending on whether the RITQ or ALSPAC norms were applied ($\chi^2(9) = 109.60$, $p<0.001$). Using ALSPAC means and SDs, fewer children were categorised as difficult (15%) and more were categorised as easy (38%), compared with RITQ norms (24% and 25%, respectively).

As shown in the Appendix A (Table 4.4 and Table 4.5), the subscale means, SDs, and categorisation of temperament were similar across all four random subsamples.

**Table 4.3 Temperament categorisation of the Avon Longitudinal Study of Parents and Children (ALSPAC) sample based on means and SDs from the Revised Infant Temperament Questionnaire (RITQ) and ALSPAC samples (n=10937)**

| Temperament categorisation | Using RITQ norms | | Using ALSPAC norms | |
|---|---|---|---|---|
| | *n* | % | *n* | % |
| Easy | 2691 | 24.6 | 4101 | 37.5 |
| Intermediate low | 3593 | 32.9 | 3673 | 33.6 |
| Intermediate high | 2050 | 18.7 | 1524 | 13.9 |
| Difficult | 2603 | 23.8 | 1639 | 15.0 |

### 4.2.6 Discussion and conclusion

Compared with the RITQ sample of 203 US children from the 1970s, ALSPAC infants in the 1990s scored considerably higher in rhythmicity, approach, and adaptability indicating higher scores on these subscales is normal for infants in this population sample. As such, when ALSPAC norms were used, fewer infants were categorised as having a difficult temperament and more infants were easy. The differences between the RITQ and ALSPAC norms can potentially be explained in several ways. First, the differences might be due to real changes in infant temperament across time, with infants in the 1990s being more difficult than infants in the 1970s, or at least that mothers in the 1990s were more likely to report their infants as difficult than mothers in the 1970s. However, a study of approximately 1000 US infants born in 1991 showed that the means and SDs of their sample were similar to those in RITQ.[134] Second, it is possible that there are cross-cultural differences, in which parents' perceptions of infant temperament varied according to cultural beliefs and values.[75] In this case, higher scores in the ALSPAC sample may reflect British mothers as more likely than American mothers to perceive their infants as difficult. However, there is a lack of evidence to support this. Although the ALSPAC population is more representative than the RITQ sample, there are still limitations to the ALSPAC sample in that it represents a predominantly White study sample, which may limit its

generalizability to population groups of different cultural backgrounds. Future studies into cross-cultural examination of temperament may be needed. Third, the difference in means and SDs may be due to the nature of the RITQ standardization sample and the ALSPAC sample. ALSPAC is a large, representative sample of UK infants, whereas the RITQ standardization sample is relatively small (n=203), and was recruited from three private paediatric practices from the eastern US, with primarily Euro-American families from the middle to upper socio-economic position.[42,202] Therefore, it is likely that the RITQ standardization sample may not be representative of the wider population. When RITQ items were used in a population sample of UK infants, we found some differences between norms from the large, geographically representative sample of UK infants and norms from the RITQ standardization sample published in the manual. ALSPAC norms may be more suitable for use with UK infants than RITQ norms because they are derived from a local sample and may be more specific to the UK context. Users of RITQ need to consider using norms that are most relevant to their own contexts.

There are a variety of temperament measurement tools available and many different ways to define temperament. The RITQ was developed based on the child psychiatric approach that emphasizes the importance of temperament and the "goodness of fit"[27] of children's temperament with parental expectations. This "goodness of fit" concept is useful to provide advice to parents,[262] and teachers,[263] and for studying parent-child relationships.[130] Community interventions that applied the "goodness of fit" concept have been effective in improving aspects of child temperament (improved attention, emotional control, exploration and sociability),[159,160,162] mathematics and reading achievement,[162] as well as reducing behavioural problems.[162,163] From a public health perspective, this

"goodness of fit" concept is important for use in interventions to help parents to understand their child's temperament and to better manage their child. There are other temperament measurement tools developed based on a child neurobehavioral development that are useful for investigating the emergence and development of temperament attributes.[40] It is beyond the aim of this paper to compare the usefulness of different measurement tools and we do not attempt to advocate the use of the RITQ's nine subscales or the five-subscale composite. Because the RITQ remains widely used in primary care and research, we would like to draw attention to the importance of the RITQ norms and how these might affect the categorisation of infants' temperament.

In this study, we used five variables (rhythmicity, approach, adaptability, intensity, and mood) to define the concept of difficult. Even though these five variables are often used to define difficult temperament, other studies have found that these five variables failed to cluster together.[5] Therefore, researchers have defined and used the concept of difficult temperament differently. For example, Rothbart focuses on poor self-regulation and high reactivity as the central definition of temperament[6] while Bates' concept of difficult was only focused on negative emotionality.[264] In the current study, adaptability may be the main contributor to the difference in proportions of infants categorised as difficult since it showed the biggest difference between the RITQ sample and the ALSPAC sample. Nevertheless, we used the five-variable concept in categorizing infant temperament simply to illustrate the problem of potential misclassification of temperament from using different norms.

There are some limitations in this study. First, as the ALSPAC data were collected in 1991-1992 they may not reflect the most contemporary measures of temperament among UK infants. More up to date information from large-scale community or population studies is needed. Large scale users of RITQ are encouraged to develop and utilize their more up to date norms based on the samples collected for their specific purposes. Second, there may be minor differences between the ALSPAC norms observed here and the norms had all RITQ items been used. However, means in each subscale were calculated based on the total scores divided by the number of items with responses in the subscale. Participants with more than 20% missing items in each subscale were excluded from analysis, as recommended in the manual.[202] Given that subscale means are calculated by taking the mean of 7 to 12 items, it is unlikely that the one or two items omitted in each subscale will have a meaningful impact on the means and SDs observed in this study.

This study casts some doubt on the appropriateness of using the 1978 normative data to categorize the temperament of infants in the 21st century. ALSPAC norms may be more appropriate for use with UK infants as they are more recent and derived from a larger, representative sample than the RITQ standardization sample. However, more up to date norms are still needed. Given that temperament profiles and categorisation are used to inform interventions and parenting practices, it is important that temperament researchers re-examine the appropriateness of RITQ norms to more accurately identify infant temperament. The RITQ norms may not be generalizable to other populations or other more contemporary studies, and hence care should be taken when using the RITQ norms so that infant temperament is not misidentified.

### 4.2.7 Online Appendices

**Appendix A**

**Table 4.4 Norms from the Avon Longitudinal Study of Parents and Children (ALSPAC) and the four randomly split subsamples**

| | **ALSPAC norms (UK), *N*=10937** | | **ALSPAC random subsample 1, *N*=2706** | | **ALSPAC random subsample 2, *N*=2761** | | **ALSPAC random subsample 3, *N*=2753** | | **ALSPAC random subsample 4, *N*=2717** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| Activity (1=low; 6=high) | 4.35 | 0.53 | 4.35 | 0.53 | 4.35 | 0.53 | 4.36 | 0.53 | 4.35 | 0.53 |
| Rhythmicity (1=regular; 6=irregular) | 2.60 | 0.69 | 2.60 | 0.70 | 2.61 | 0.67 | 2.58 | 0.70 | 2.61 | 0.69 |
| Approach (1=approach; 6=withdraw) | 2.53 | 0.64 | 2.53 | 0.64 | 2.52 | 0.63 | 2.53 | 0.64 | 2.54 | 0.64 |
| Adaptability (1=high; 6=low) | 2.42 | 0.57 | 2.42 | 0.57 | 2.42 | 0.56 | 2.41 | 0.57 | 2.43 | 0.58 |
| Intensity (1=mild; 6=intense) | 3.51 | 0.56 | 3.52 | 0.57 | 3.52 | 0.57 | 3.52 | 0.55 | 3.51 | 0.55 |
| Mood (1=positive; 6=negative) | 2.78 | 0.66 | 2.79 | 0.66 | 2.78 | 0.66 | 2.78 | 0.66 | 2.78 | 0.65 |
| Persistence (1=high; 6=low) | 2.95 | 0.73 | 2.96 | 0.73 | 2.94 | 0.72 | 2.94 | 0.72 | 2.95 | 0.72 |
| Distractibility (1=high; 6=low) | 2.42 | 0.56 | 2.42 | 0.56 | 2.43 | 0.55 | 2.41 | 0.56 | 2.43 | 0.57 |
| Threshold (1=high; 6=low) | 3.75 | 0.61 | 3.75 | 0.60 | 3.76 | 0.61 | 3.76 | 0.51 | 3.74 | 0.61 |

*Note.* Subscales used for categorisation are rhythmicity, approach, adaptability, intensity, and mood.

**Table 4.5 Temperament categorisation in the Avon Longitudinal Study of Parents and Children (ALSPAC) sample and the four randomly split subsamples**

| | **ALSPAC norms (UK), *N*=10937** | | **ALSPAC random subsample 1, *N*=2706** | | **ALSPAC random subsample 2, *N*=2761** | | **ALSPAC random subsample 3, *N*=2753** | | **ALSPAC random subsample 4, *N*=2717** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **n** | **%** | **n** | **%** | **n** | **%** | **n** | **%** | **n** | **%** |
| Easy | 4101 | 37.5 | 987 | 36.5 | 1055 | 38.2 | 1049 | 38.1 | 1010 | 37.2 |
| Intermediate low | 3673 | 33.6 | 935 | 34.6 | 915 | 33.1 | 919 | 33.4 | 904 | 33.3 |
| Intermediate high | 1524 | 13.9 | 374 | 13.8 | 389 | 14.1 | 380 | 13.8 | 381 | 14.0 |
| Difficult | 1639 | 15.0 | 410 | 15.2 | 402 | 14.6 | 405 | 14.7 | 422 | 15.5 |

# CHAPTER 5. Does temperament at ages 2-3 directly affect cognitive and academic outcomes at ages 6-7?

## 5.1 Preface

This chapter contains the second article contributing to this thesis. This article was submitted for publication in April 2016 and is currently under review.

In Chapter 4, differences in temperament between US and UK infants were identified, and in Chapter 5 the focus shifts from infants to toddlers. Temperament in younger children (infancy) is thought to reflect the innate characteristics of an individual, whereas temperament in older children (toddlerhood and preschool age) is thought to be more strongly influenced by environmental factors such as parenting and the rapid development of self-regulation which makes temperament modifiable through interventions. Within the literature there is widespread interest in modifying aspects of temperament as a means of improving children's skills and abilities across a wide range of developmental domains, including cognitive, social, emotional, language and academic skills.[117] However, most studies have not accounted for parenting practices when examining the direct effect between temperament and cognitive and academic outcomes, and a majority of studies are limited by the use of small samples and cross-sectional design.

Chapter 5 examines temperament in toddlerhood (2 to 3 years) and the influence of temperament on cognitive outcomes and academic achievement in childhood (6 to 7 years). Data from a nationally representative sample of Australian children (LSAC) was

used. There are some advantages in using the LSAC data over ALSPAC data for this particular research question because LSAC collected data on temperament using culturally appropriate temperament questionnaires that were suitable for use in Australian children. The temperament questionnaires used in LSAC were developed from the Australian Temperament Project (ATP). In addition, the temperament questionnaires used in LSAC consists of three main dimensions: reactivity, approach, and persistence which were consistent with temperament models suggested in contemporary temperament research.[6] While the associations between temperament with cognitive and academic outcomes have been investigated in many psychological studies, this study applies a larger population-based sample and contemporary epidemiological analyses that account for a wide range of confounders to make more confident causal inferences.

## 5.2 Publication: Does temperament at ages 2-3 directly affect cognitive and academic outcomes at ages 6-7?

### 5.2.1 Statement of authorship

| | |
|---|---|
| Title of Paper | Does temperament at ages 2-3 directly affect cognitive and academic outcomes at ages 6-7? |
| Publication Status | ☐ Published<br>☐ Accepted for Publication<br>☑ Submitted for Publication<br>☐ Unpublished and Unsubmitted w ork w ritten in manuscript style |
| Publication Details | This manuscript is currently under review. |

### Principal Author

| | | | |
|---|---|---|---|
| Name of Principal Author (Candidate) | Shiau Yun Chong | | |
| Contribution to the Paper | Conceived and designed the study, analysed and interpreted the data, drafted the manuscript. | | |
| Overall percentage (%) | 85 | | |
| Certification: | This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper. | | |
| Signature |  | Date | 15-08-2016 |

### Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

iv. the candidate's stated contribution to the publication is accurate (as detailed above);

v. permission is granted for the candidate in include the publication in the thesis; and

vi. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

| Name of Co-Author | Catherine Chittleborough | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, data acquisition, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | Tess Gregory | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | John Lynch | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | Murthy Mittinty | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | Lisa Smithers | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

### 5.2.2 Abstract

**Aim:** There is widespread interest in temperament and its impact upon cognitive and academic outcomes. We examined the controlled direct effects of temperament (2 to 3 years) on cognitive and academic outcomes (6 to 7 years) after accounting for parenting practices (4 to 5 years) because parents may respond differently to children's temperament and parenting practices are an important influence on children's learning.

**Methods:** Participants were from the Longitudinal Study of Australian Children (n=5107). Cognitive abilities were measured by the Peabody Picture Vocabulary Test (verbal ability) and the Matrix Reasoning test (non-verbal ability). Academic outcomes (literacy and numeracy) were reported by teachers using the Academic Rating Scale. Temperament was reported by mothers using the Short Temperament Scale for Toddlers with 3 subscales: reactivity, approach, and persistence. Parenting practices were measured using 7 items about the frequency parents engaged in activities with their children. Marginal structural models with inverse probability of treatment weights were used to estimate the controlled direct effects of temperament.

**Results:** Of the 3 subscales, persistence showed the largest effects on verbal (β=0.58; 0.27, 0.89) and non-verbal (β=0.19; 0.02, 0.34) abilities. There were positive effects of persistence on literacy (β=0.08; 0.03, 0.13) and numeracy (β=0.08; 0.03, 0.13) but negative effects of reactivity on literacy (β=-0.08; -0.11, -0.05) and numeracy (β=-0.07; -0.10, -0.04). Little evidence of approach on literacy and numeracy was found.

**Conclusion:** The controlled direct effects of temperament on cognitive and academic outcomes, after accounting for parenting practices, were small.

## What is already known on this topic

- Temperament may influence cognitive and academic performance.
- Children's temperament may also influence parenting practices, which could then affect their outcomes.
- Past studies have not taken into account parenting practices when examining the direct effect of temperament.

## What this paper adds

- Temperament reactivity and persistence had direct effects on verbal, non-verbal cognitive abilities and literacy and numeracy.
- Temperament approach had direct effect on verbal and non-verbal cognitive abilities but no effect on literacy and numeracy.
- Effect sizes of temperament reactivity, approach and persistence were small after accounting for parenting.

### 5.2.3 Introduction

There is widespread interest in the influence of children's temperament on cognitive and academic outcomes.[63,144] Temperament is the individual characteristics in behavioural styles that are biologically-based, but also shaped by environmental context.[37] Temperament dimensions found in young children include reactivity, approach, and persistence, and all three dimensions may impact learning and cognitive development.[56] Temperamental reactivity encompasses the child's emotional intensity and volatility, while persistence refers to the ability of the child to stay on task and maintain their attention.[56] Less is known about the influence of temperamental approach, which is the degree of comfort when encountering new situations or people.[56] Temperament reactivity and persistence are modifiable through positive interactions with parents and teachers.[11,265] Increasing children's persistence and reducing their reactivity may be a mechanism to improve children's cognitive and academic outcomes, provided that there are direct effects of temperament on these outcomes.[11,117,265]

To estimate the direct effect of temperament, analyses need to consider that children's temperament may also influence parenting practices,[157,266] which are known to influence children's cognitive and academic outcomes.[95] For example, Maccoby *et al.*[157] showed that temperamentally difficult children received less teaching effort from parents than temperamentally easy children. On the other hand, Dixon *et al.*[266] found that mothers engaged in more high quality play with temperamentally difficult children than easy children. Parental engagement in activities is known to have positive impacts on children's cognitive and academic outcomes,[95] and differential parental engagement for

temperamentally easy and difficult children might explain some of the effects of temperament on children's cognitive and academic outcomes.

Studies examining the direct effect of temperament on cognitive and academic outcomes involve limited adjustment for potential confounders.[63,267,268] The few studies[118,119] that accounted for parenting practices by adjusting for parenting in regressions or path analyses have reported some evidence of a direct effect of temperament on outcomes. However, adjusting for parenting practices could introduce bias when parenting practices are affected by temperament and when there are confounders of parenting and outcomes (Appendix B).[124] In this study, we use marginal structural models (MSMs) to estimate the controlled direct effects (CDEs) of temperament (reactivity, approach and persistence) at ages 2 to 3 years on cognitive and academic outcomes at ages 6 to 7 years after accounting for parenting practices at ages 4 to 5 years using a weighting approach. We study both cognitive and academic abilities because cognitive ability measures the verbal and non-verbal skills needed to succeed in life, and academic achievement provides an indication of child performance in real-life settings, which is predictive of later educational achievement and labour market outcomes.[269]

### 5.2.4 Methods

***Study design and sample***

This study used data from the infant cohort of the Longitudinal Study of Australian Children (LSAC), which is a population-based longitudinal study that commenced in 2004. A two-stage clustered sampling process was used to recruit participants.[205] This included a random selection of postcodes at stage one and then a random selection of children

within each selected postcode using the Medicare enrolment database as a sampling frame. At study commencement, 5107 infants in their first year of life were recruited into the study, and they were followed up at ages 2 to 3 (n=4606), 4 to 5 (n=4386), and 6 to 7 (n=4242) years. The LSAC sample is considered broadly representative of Australian children.[205] LSAC was approved by the Australian Institute of Family Studies ethics committee. Written informed consent was obtained from all participants.

***Cognitive ability***

Verbal ability (receptive vocabulary) was measured using an adapted Peabody Picture Vocabulary Test (PPVT), and non-verbal ability (fluid reasoning) was measured using the Matrix Reasoning test from the Wechsler Intelligence Scale for Children, 4th edition.[270] The adapted PPVT-III[207] was administered by a trained interviewer to children aged 6 to 7 years during home interviews. The child was asked to point to the picture that best represents the meaning of a word spoken by the examiner.[271] The adapted PPVT-III was comparable to the full PPVT-III.[207] Scores were created using Rasch modelling.[207] The matrix reasoning test comprised 35 items of increasing difficulty. The child was presented with an incomplete set of diagrams and was asked to select the picture that completes the set from 5 different options. Scores were reported as standard scores, based on age appropriate norms.[270]

***Academic achievement***

Academic achievement was reported by teachers at ages 6 to 7 years using the adapted Academic Rating Scale (ARS).[208] The ARS consists of two subscales: literacy (10 items) and numeracy (9 items). Teachers were asked to rate the child's skills and knowledge in

literacy and mathematical understanding in relation to other children of the same age. An example of items in literacy subscale was 'reads age-appropriate books fluently'. An item in the numeracy subscale included 'makes reasonable estimates of quantities'. Teachers rated each item from 1 (not yet demonstrated skill) to 5 (demonstrates skill competently and consistently). Total scores ranging from 1 to 5 were created using Rasch modelling with higher scores indicating higher proficiency.[209] The ARS is a reliable measure of children's academic performance (α=0.91 for literacy; 0.94 for numeracy).[272]

***Temperament***

Temperament was measured at ages 2 to 3 years using the Short Temperament Scale for Toddlers (STST).[211] The STST is an adapted version of the Toddler Temperament Scale.[58] The STST used in LSAC consisted of 3 subscales: reactivity, approach, and persistence (4 items in each subscale, α=0.99 for each subscale), rated by mothers from 1 (almost never) to 6 (almost always). An average score was calculated for each subscale. Higher scores indicate higher reactivity, approach, and persistence.

***Parenting practices***

Parenting practices were assessed at ages 4 to 5 years using 7 items measuring how often the mothers engaged in the following activities with the child: read to the child, tell stories, draw pictures or other craft activities, play with toys or games indoors such as board or card games, play music, involve child in daily activities such as cooking or pet care, and play outdoor games including walking or cycling. These seven items have been used as important indicators of early childhood development in the UNICEF Multiple Cluster Index Surveys.[273] Items were rated from 0 (none) to 3 (everyday). A total score

ranging from 0 to 21 was calculated from scores of these items. A higher score indicated more positive parenting practices.

***Confounders of the association between temperament and outcomes***

Factors that might confound the associations between temperament and cognitive and academic outcomes were decided *a priori* using a directed acyclic graph (Appendix C). Confounders included maternal education, financial hardship, housing tenure, Aboriginal or Torres Strait Islander, neighbourhood disadvantage, gestational hypertension, gestational diabetes, smoking and alcohol intake during pregnancy, sex, birth weight z-score, breastfeeding, maternal age, country of birth, psychological distress, mother and partner argumentative relationship, single-parent family. All these confounders were reported by mothers during home interviews when children were aged 0 to 1 year. Details about how these confounders were measured were included in Appendix C.

***Confounders of the association between parenting practices and outcomes***

To estimate the CDEs, we need to account for confounding associated with parenting and cognitive or academic outcomes.[124] Specifically, this set of confounders were reported by mothers at ages 4 to 5 years and included: 1) variables that were affected by temperament and in turn confound the parenting-outcome association; including maternal psychological distress and number of siblings, and 2) variables that were not affected by temperament but confound the parenting-outcome association; including maternal working status (full-time, part-time, and not working), household income (total resources from adult family members), and financial hardship.

***Analysis***

Using MSMs, we estimated the CDEs of temperament (reactivity, approach, and persistence) (*X*), on cognitive and academic outcomes (*Y*) after accounting for parenting practices *(M),* potential confounders of the association between temperament and cognitive and academic outcomes *(C)*, and confounders of the association between parenting practices and cognitive and academic outcomes (*L*). The MSM is based on a counterfactual or potential outcome framework, and it allows the estimation of the CDE by comparing the effect of temperament in the exposed group *x* and the counterfactual group *x** while setting the *M* to a uniform level of *m*.[124]

The MSM takes into account confounders that affect the association between parenting practices and outcomes to give an estimate of the CDE of temperament.

The CDEs were estimated from linear regression models of the form:

$$E[Y_i|\, X_i, M_i] = \beta_0 + \beta_1 x + \beta_2 m$$

Potential confounding was accounted by fitting the model above with stabilised inverse probability of treatment weights of the form $= w_i^X \times w_i^M$ , where

$$w_i^X = \frac{f(X_i)}{f(X|C_i)}$$

and

$$w_i^M = \frac{f(M_i|X_i)}{f(M_i|X_i,C_i,L_i)}.$$

The weight $w_i^X$ accounted for the confounding of the association between *X* and *Y* by conditioning on *C*. The weight $w_i^M$ accounted for the confounding of the association

between *M* and *Y* by conditioning on *X*, *C*, and *L*. Weights were obtained from probability density functions, estimated using linear regressions.[124] Weights were truncated at the 1st and 99th percentile to deal with outliers (Appendix B). MSMs were performed separately for each aspect of temperament (reactivity, approach, and persistence).

We performed sensitivity analysis to determine the extent to which an unmeasured confounder *U* might affect the association between *M* and *Y*. We estimated the bias for the CDEs under a range of conditions based on the conceivable prevalence and the effect size of *U* (Appendix D).[221]

***Multiple imputation***

Multiple imputation was used to account for potential bias from missing data, under the missing at random assumption. Imputed datasets were generated using the 'mi impute chained' command in STATA (StataCorp, College Station, TX). The imputation model included temperament, parenting practices, cognitive and academic outcomes, confounding variables and auxiliary variables that predicted missingness. Twenty imputed datasets were generated with 50 cycles of regression switching. All analyses were performed on the complete case sample and the imputed sample and the results were similar (data not shown). Results from the imputed sample (n=5107) are reported.

### 5.2.5 Results

Table 5.1 Characteristics of response, complete case, and imputed samples shows the characteristics of LSAC response, complete case (*n*=1647) and imputed (*n*=5107) samples. The characteristics of the imputed sample were similar to the response sample.

**Table 5.1 Characteristics of response, complete case, and imputed samples**

| | Response sample[a] | | Complete case sample,[b] n=1647 | Imputed sample, n=5107 |
|---|---|---|---|---|
| | **N** | **M (SD) or %** | **M (SD) or %** | **M (SD) or %** |
| **Outcomes, *Y* (Cognitive ability and academic achievement)** | | | | |
| PPVT | 4185 | 74.4 (5.2) | 75.2 (4.9) | 74.2 (5.2) |
| Matrix reasoning | 4180 | 10.7 (3.0) | 11.0 (3.0) | 10.7 (3.0) |
| ARS-Literacy | 3408 | 3.4 (0.8) | 3.5 (0.7) | 3.4 (0.8) |
| ARS-Numeracy | 3357 | 3.3 (0.8) | 3.5 (0.8) | 3.3 (0.8) |
| **Exposures, *X* (Temperament)** | | | | |
| Reactivity | 3530 | 3.0 (1.0) | 2.9 (0.9) | 3.0 (0.9) |
| Approach | 3533 | 3.9 (1.0) | 4.0 (1.0) | 3.9 (1.0) |
| Persistence | 3532 | 4.3 (0.7) | 4.3 (0.7) | 4.3 (0.7) |
| **Intermediate variable, *M*** | | | | |
| Parenting practices | 4385 | 11.8 (3.9) | 12.0 (3.8) | 11.7 (3.9) |
| **Confounders of *X*-*M* or *X*-*Y*** | | | | |
| Mother's highest education, % | | | | |
| Tertiary | 1677 | 32.9 | 41.4 | 32.9 |
| Diploma/certificate | 1766 | 34.6 | 34.2 | 34.6 |
| Schooling only | 1656 | 32.5 | 24.4 | 32.5 |
| Hardship score | 5089 | 0.9 (1.3) | 0.7 (1.1) | 0.9 (1.3) |
| Housing tenure, rented or other, % | 5100 | 35.8 | 23.0 | 35.8 |
| Aboriginal or Torres Strait Islander, % | 5107 | 4.5 | 2.1 | 4.5 |
| Index of relative socio-economic disadvantage (IRSD) | 5107 | 1008.8 (60.2) | 1017.2 (57.2) | 1008.8 (60.2) |
| Child is male, % | 5107 | 51.1 | 48.5 | 51.1 |
| Birth weight for gestational age z-score | 4999 | 0.0 (1.1) | 0.1 (1.0) | 0.0 (1.1) |
| Duration of breast feeding, % | | | | |
| Never breastfed | 420 | 9.2 | 5.8 | 9.2 |
| <1 month | 518 | 11.4 | 8.4 | 11.4 |
| <3 months | 473 | 10.4 | 8.7 | 10.3 |
| <6 months | 692 | 15.2 | 14.0 | 15.2 |
| 6 months or more | 2461 | 53.9 | 63.2 | 54.0 |
| Maternal age | 5106 | 31.0 (5.5) | 32.2 (4.6) | 31.0 (5.5) |
| Mother is born in Australia, % | 4997 | 80.0 | 83.0 | 79.4 |
| Mother's psychological distress | 4308 | 4.4 (0.8) | 4.5 (0.5) | 4.4 (0.6) |
| Mother and partner argumentative relationship | 3931 | 2.2 (0.6) | 2.1 (0.6) | 2.2 (0.6) |
| Single-parent household, % | 5103 | 9.7 | 0.2 | 9.7 |
| Mother had gestational hypertension, % | 4238 | 7.8 | 7.8 | 8.1 |
| Mother had gestational diabetes, % | 4223 | 5.7 | 5.2 | 5.9 |
| Any alcohol during pregnancy, % | 4227 | 38.9 | 44.1 | 37.3 |
| Smoking during pregnancy, % | 4239 | 16.7 | 10.6 | 18.2 |
| **Confounders of *M*-*Y*** | | | | |
| Mother's psychological distress | 3818 | 4.5 (0.6) | 4.5 (0.5) | 4.5 (0.6) |
| Number of siblings | 4386 | 1.5 (1.3) | 1.5 (0.9) | 1.5 (1.0) |
| Mother's working status, % | | | | |

| | Response sample[a] | | Complete case sample,[b] n=1647 | Imputed sample, n=5107 |
|---|---|---|---|---|
| | N | M (SD) or % | M (SD) or % | M (SD) or % |
| Working full-time | 957 | 21.8 | 19.8 | 21.1 |
| Working part-time | 1804 | 41.2 | 46.6 | 40.7 |
| Currently not working | 1622 | 37.0 | 33.6 | 38.2 |
| Household income per week, $ | 3668 | 1977.8 (1343.4) | 2054.5 (1365.3) | 1893.6 (1297.7) |
| Hardship score | 4365 | 0.3 (0.7) | 0.2 (0.6) | 0.3 (0.8) |

PPVT, Peabody Picture Vocabulary Test; ARS, Academic Rating Scale; *X-M*, exposure-mediator; *X-Y*, exposure-outcome; *M-Y*, mediator-outcome

[a] Response sample is the number of participants who responded to specific assessment for each child exposure, outcome, or confounder.

[b] Complete case sample includes participants with complete data on all exposures, outcomes, and confounders.

Table 5.2 displays the CDEs of temperament subscales on child outcomes after accounting for all confounding factors using stabilised weights. High temperamental reactivity had negative effects on all outcomes, particularly verbal ability (β=-0.37 95% CI -0.59, -0.14). High approach had positive effects on verbal (β=0.45 95% CI 0.22, 0.67) and non-verbal abilities (β=0.11 95% CI 0.01, 0.21) but little or no effect on literacy (β=0.03 95% CI -0.01, 0.07) or numeracy (β=0.01 95% CI -0.02, 0.05). High persistence had a positive impact on all outcomes. However, effect sizes were small for all temperament subscales. For instance, 1 unit higher persistence (range: 1 to 5) was associated with 0.58-unit (0.11 SD) higher verbal ability and 0.19-unit (0.06 SD) higher non-verbal ability.

**Table 5.2 Controlled direct effects (CDEs) of temperament reactivity, approach, and persistence at ages 2 to 3 years on child outcomes at ages 6 to 7 years (n=5107)**

| Child outcomes | Reactivity | | | Approach | | | Persistence | | |
|---|---|---|---|---|---|---|---|---|---|
| | β | (95% CI) | ES[a] | β | (95% CI) | ES | β | (95% CI) | ES |
| PPVT | -0.37 | (-0.59, -0.14) | 0.07 | 0.45 | (0.22, 0.67) | 0.09 | 0.58 | (0.27, 0.89) | 0.11 |
| Matrix reasoning | -0.11 | (-0.21, -0.01) | 0.04 | 0.11 | (0.01, 0.21) | 0.04 | 0.19 | (0.02, 0.34) | 0.06 |
| ARS-Literacy | -0.08 | (-0.11, -0.05) | 0.10 | 0.03 | (-0.01, 0.07) | 0.04 | 0.08 | (0.03, 0.13) | 0.10 |
| ARS-Numeracy | -0.07 | (-0.10, -0.04) | 0.09 | 0.01 | (-0.02, 0.05) | 0.00 | 0.08 | (0.03, 0.13) | 0.10 |

CI, confidence interval, ES, effect size, PPVT, Peabody Picture Vocabulary Test, ARS, Academic Rating Scale.

[a] ES were calculated by dividing the beta coefficients by the standard deviations.

Our sensitivity analysis (Appendix D) showed that the CDEs were generally robust in the presence of a binary unmeasured confounder. The observed CDEs would be explained by the unmeasured confounder in scenarios when there is a large difference (for example, 80%) in the prevalence of the unmeasured confounder *U* between the exposed and the counterfactual group and when the effect sizes of the *U* on outcomes were moderate to strong (for example, β=0.60).

### 5.2.6 Discussion

This study found some evidence of CDE of temperament at 2 to 3 years on children's verbal and non-verbal cognitive abilities and literacy and numeracy outcomes at ages 6 to 7 years, but these effects were small.

Of the three temperament dimensions, persistence had the largest CDE on children's cognitive abilities and academic achievement, which is consistent with other studies that have measured persistence, or other aspects of self-regulation using different questionnaires[144,146] for children at different ages.[144] For instance, a cross-sectional study showed that temperamental effortful control at 3 to 5 years measured using the Children's Behavior Questionnaire (CBQ) was associated with 0.29 SD higher letter knowledge and 0.17 SD higher math achievement.[146] Rudasill *et al.*[268] found that temperamental attention measured using CBQ at 4.5 years was associated with 0.18 SD higher reading scores and 0.14 SD higher mathematic score in children aged 8 to 10 years. The effect sizes of persistence in the current study (0.11 SD for verbal cognitive, 0.06 SD for non-verbal cognitive, 0.10 SD for literacy and numeracy) were smaller than in previous studies that did not account for parenting practices and other important

confounders.[146,268] Without accounting for parenting, we would expect the effect sizes in this sample would be larger and similar to previous research.

This current study found that temperamental reactivity in toddlerhood had negative effects on cognitive and academic outcomes, with the largest effect of 0.10 SD on literacy, which is similar to other studies that measured temperamental mood and negative emotionality in older children.[144,267] Studies that measured emotional reactivity during infancy have reported somewhat different findings. Karass and Braungart-Rieker[274] found no evidence of infants' frustration and anger on cognitive ability at age 3 years. Lawson and Ruff[145] found that mother-reported negative mood at 1 year was associated with lower IQ at age 3.5 years. It is possible that younger children are less able to regulate their emotion and attention as self-regulation develops after the age of 2.[275]

Consistent with a previous study that also used the approach subscale in the STST,[276] this current study found that higher approach was associated with higher verbal (0.09 SD) and non-verbal (0.04 SD) cognitive abilities. However, the previous study did not account for parenting, and the effect of approach on vocabulary ability was smaller (0.01 SD).[276] This smaller effect of approach could be because parents spend more time with shy children to help them cope with unfamiliar or new environments and everyday tasks. Therefore, with positive parenting, shy children may not demonstrate poorer outcomes than outgoing children. While we found little evidence of temperament approach on literacy and numeracy, a previous study found that shy children were more likely to be reported by teachers as having poorer academic achievement.[149] The difference in findings may be

because the previous study did not account for parenting practices or measured shyness at later ages when children were expected to have more developed sociability skills.

This study used a large, nationally representative sample to strengthen the generalizability of the findings to the whole population. Previous studies used smaller (n<500), non-representative samples, and cross-sectional or short term longitudinal design (1 to 2 years).[146,267] We are able to account for parenting practices, which are an important influence of children's cognitive and academic outcomes and a wide range of potential confounders to estimate the CDEs of temperament. Previous studies have not accounted for parenting or only examined narrow aspects of parenting.[118,119] In this current study, parenting practices were defined as the frequency of parents engaging in activities with their children because there is evidence that this aspect of parenting is important for children' cognitive development.[95] However, it is possible that a broader definition of parenting practices might have an impact on the estimated direct effects.

There is some evidence of a controlled direct effect of temperamental persistence, approach and reactivity at ages 2 to 3 years on cognitive and academic outcomes at ages 6 to 7 years, but the effects are small. The largest direct effect observed was for persistence on verbal ability but this effect corresponded to only a 0.11 SD increase in verbal ability for every 1 point increase in persistence. Future research may be useful to establish the direct effect of persistence on a broader range of outcomes such as school completion and grades, tertiary entrance rankings, and labour market outcomes.

### 5.2.7 Online appendices

**Appendix B The marginal structural model**

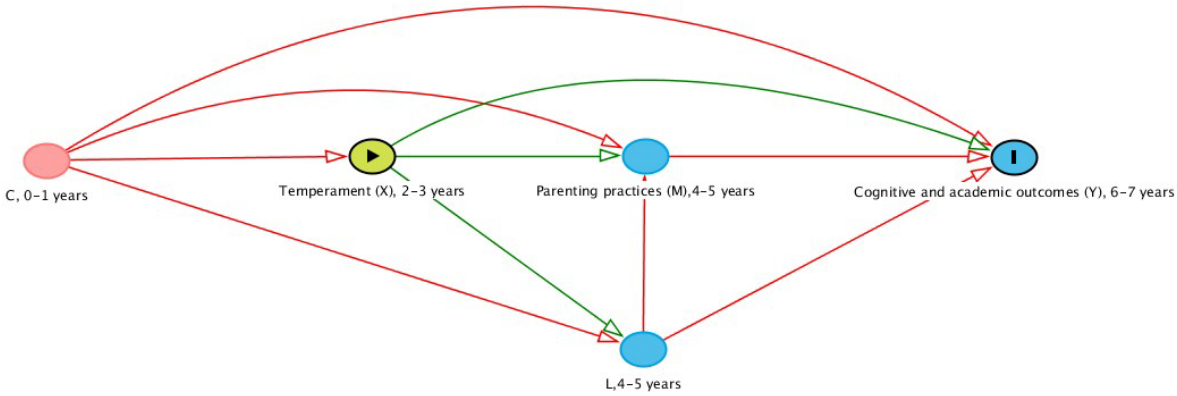


**Figure 5.1 Causal diagram of the hypothesized effects of temperament at 2 to 3 years and parenting practices at 4 to 5 years on cognitive and academic outcomes at ages 6 to 7 years**

Figure 5.1 depicts the causal diagram for our study. *X* (temperament subscales of reactivity, approach, and persistence) represents the exposure, *M* (parenting practices) represents the intermediate variable, and *Y* (cognitive and academic outcomes) represents the outcome. *C* represents confounders of the association between temperament (*X*), parenting practices (*M*), and cognitive and academic outcomes (*Y*) measured at ages 0 to 1 year (maternal education, financial hardship, housing tenure, Aboriginal or Torres Strait Islander, neighbourhood disadvantage, sex, birth weight for gestational age z-score, duration of breastfeeding, maternal age, maternal country of birth, maternal psychological distress, mother and partner argumentative relationship, single-parent family, gestational hypertension, gestational diabetes, smoking and alcohol intake during pregnancy). *L* represents confounders of the effect of parenting practices *M* on cognitive and academic outcomes *Y* measured at ages 4 to 5 years (maternal psychological distress, number of siblings, mothers' working status, household income, and financial hardship).

The standard approach to estimate the direct effect is by regressing the outcome *Y* on the exposure *X* and some exposure-outcome confounders *C* and then considering whether the coefficient for *X* changes when controlling for the intermediate variable *M*. The difference in coefficients of *X* is a measure of the effect that is going through by *M*.[120] However, using the standard regression approach to assess the direct effect can lead to biased estimates when there are confounders of the *M*-*Y* association, *L*.[216,277] For example, the number of siblings *L* affects parenting practices *M* at age 4 to 5 years and child outcomes *Y* at age 6 to 7 years. If we adjusted for *M,* as in the regression approach, we induce an association between *X* and *L*.[216] If we additionally adjusted for *L*, we block part of the direct effect of *X* on *Y* that is not through *M*.[278]

To overcome limitations of standard regression, marginal structural models have been recommended for better estimation of the direct effect.[124] The marginal structural model differs from the standard regression approach in that the model is for counterfactual outcomes rather than observed outcomes.[124,279] The counterfactual approach allows the estimation of the controlled direct effect by comparing the extent to which an outcome would change if the intermediate variable were controlled at level *m* uniformly in the population but the exposure was changed from the observed level, *x* to a counterfactual level, *x**.[124]

The marginal structural model is not conditioned on any covariates but uses a weighting approach to account for confounding factors.[219,280] Under the assumption of no unmeasured confounding factors, the weighting method creates a pseudo population in

which the association between confounding factors and the exposure is removed. The weighting approach takes into account the confounding effect of *L* to allow a better estimation of the direct effect.[125]

Table 5.3 displays the estimated stabilised inverse probability of treatment weights. The means of the stabilised weights were around 1.00, suggesting no misspecification of the model.[219]

**Table 5.3 Stabilised inverse probability of treatment weights**

| | **Estimated weight** | | **Weight truncated at 99th percentile** | |
|---|---|---|---|---|
| | **Mean (SD)** | **Minimum/maximum** | **Mean (SD)** | **Minimum/maximum** |
| Reactivity | 1.00 (0.55) | 0.03/54.12 | 0.99 (0.38) | 0.25/4.39 |
| Approach | 1.00 (0.46) | 0.04/41.36 | 0.99 (0.32) | 0.27/3.54 |
| Persistence | 1.00 (0.47) | 0.05/33.10 | 0.99 (0.31) | 0.29/3.30 |

## Appendix C Confounders of the association between temperament and outcomes

Indicators of socio-economic position included maternal education (tertiary, diploma and certificate, and schooling only), housing tenure (owned/mortgaged versus rented/other), Aboriginal or Torres Strait Islander (yes/no), financial hardship, and neighbourhood disadvantage. Financial hardship was assessed from 7 items asking whether mothers had experienced the following due to shortage of money: adults or children went without meals; they were unable to heat or cool their home; they pawned or sold something; or they sought assistance from a welfare or community organization. Hardship scores ranged from 0 to 7, with a higher score indicating greater financial difficulties. The Index of Relative Socio-economic Disadvantage (IRSD) was used as an indicator of neighbourhood disadvantage. The IRSD is based on postcode of residence and has a national mean of 1000 and a standard deviation of 100, with lower scores indicating greater disadvantage.

Intrauterine factors included self-reported gestational hypertension (yes/no), gestational diabetes (yes/no), and smoking and alcohol intake during pregnancy (yes/no). Child factors included sex, birth weight for gestational age z-score and duration of breast feeding (never breastfed, < 1 month, <3 months, < 6 months, and ≥6 months). Birth weight percentiles were calculated based on Australian birth weight percentiles data collected from 1998 to 2007 stratified by sex and gestational age.[281]

Maternal and family factors included age, maternal country of birth (Australia versus other countries), psychological distress, mother and partner argumentative relationship, and single-parent household (yes/no). Mother's psychological distress was assessed using the Kessler K6 scale with a total score that ranged from 1 to 5.[282] Higher scores indicated

less distress. Five questions were asked about conflicts the mother had with her partner, including disagreement, argument, stressful conversation, verbal hostility, and physical hostility. Responses for each item ranged from 1 (not at all) to 5 (all the time). Higher scores indicated more argumentative relationships.

## Appendix D Sensitivity analysis for controlled direct effects

Sensitivity analyses were conducted to assess the effect of a binary unmeasured confounder, *U* of the association between *M* and *Y*, on the CDEs of the temperament subscale of reactivity on cognitive and academic outcomes. The bias for the conditional CDE is defined as[221]:

$$Bias\left(CDE_{x,x*}(\mathrm{m})\right) = \delta\gamma$$

Where $\delta$ denotes the difference in prevalence of *U* in the exposed, *x* and counterfactual, *x** group while $\gamma$ denotes the effect size of *U* on cognitive and academic outcomes.

Table 5.4 showed that to invalidate the observed CDE (β=-0.37) of reactivity on the PPVT score, the unmeasured confounder *U* would need to have a difference in prevalence of 80% (for example 90% for exposure level *x*, and 10% for counterfactual exposure level *x**) and would be required to decrease the PPVT score by at least 0.6. *U* might have an effect size of 0.6 on the PPVT score. However, it does not appear plausible for a *U* to have a prevalence difference of 0.8 or more that could eliminate the CDE of reactivity on PPVT.

**Table 5.4 The effect of a potential unmeasured binary confounder *U* of parenting practices to outcomes pathway**

| $P(U=1\|x, m, c)^{(1)}$ | $P(U=1\|x^*, m, c^{(2)}$ | $\delta=^{(1)-(2)}$ | $\gamma$ | $d=\delta\gamma$ |
|---|---|---|---|---|
| **Peabody Picture Vocabulary Test (direct effect of reactivity β=-0.37)** | | | | |
| 0.6 | 0.4 | 0.2 | -0.40 | -0.08 |
| 0.7 | 0.3 | 0.4 | -0.40 | -0.16 |
| 0.8 | 0.2 | 0.6 | -0.40 | -0.24 |
| 0.9 | 0.1 | 0.8 | -0.40 | -0.32 |
| 0.6 | 0.4 | 0.2 | -0.60 | -0.12 |
| 0.7 | 0.3 | 0.4 | -0.60 | -0.24 |
| 0.8 | 0.2 | 0.6 | -0.60 | -0.36 |
| 0.9 | 0.1 | 0.8 | -0.60 | -0.48 |
| 0.6 | 0.4 | 0.2 | -0.80 | -0.16 |
| 0.7 | 0.3 | 0.4 | -0.80 | -0.32 |
| 0.8 | 0.2 | 0.6 | -0.80 | -0.48 |
| 0.9 | 0.1 | 0.8 | -0.80 | -0.64 |
| **Matrix Reasoning Test (direct effect of reactivity β =-0.11)** | | | | |
| 0.6 | 0.4 | 0.2 | -0.10 | -0.02 |
| 0.7 | 0.3 | 0.4 | -0.10 | -0.04 |
| 0.8 | 0.2 | 0.6 | -0.10 | -0.06 |
| 0.9 | 0.1 | 0.8 | -0.10 | -0.08 |
| 0.6 | 0.4 | 0.2 | -0.20 | -0.04 |
| 0.7 | 0.3 | 0.4 | -0.20 | -0.08 |
| 0.8 | 0.2 | 0.6 | -0.20 | -0.12 |
| 0.9 | 0.1 | 0.8 | -0.20 | -0.16 |
| 0.6 | 0.4 | 0.2 | -0.30 | -0.06 |
| 0.7 | 0.3 | 0.4 | -0.30 | -0.12 |
| 0.8 | 0.2 | 0.6 | -0.30 | -0.18 |
| 0.9 | 0.1 | 0.8 | -0.30 | -0.24 |
| **ARS-Literacy (direct effect of reactivity β =-0.08)** | | | | |
| 0.6 | 0.4 | 0.2 | -0.05 | -0.01 |
| 0.7 | 0.3 | 0.4 | -0.05 | -0.02 |
| 0.8 | 0.2 | 0.6 | -0.05 | -0.03 |
| 0.9 | 0.1 | 0.8 | -0.05 | -0.04 |
| 0.6 | 0.4 | 0.2 | -0.10 | -0.02 |
| 0.7 | 0.3 | 0.4 | -0.10 | -0.04 |
| 0.8 | 0.2 | 0.6 | -0.10 | -0.06 |
| 0.9 | 0.1 | 0.8 | -0.10 | -0.08 |
| 0.6 | 0.4 | 0.2 | -0.15 | -0.03 |
| 0.7 | 0.3 | 0.4 | -0.15 | -0.06 |
| 0.8 | 0.2 | 0.6 | -0.15 | -0.09 |
| 0.9 | 0.1 | 0.8 | -0.15 | -0.12 |
| **ARS-Numeracy (direct effect of reactivity β =-0.07)** | | | | |
| 0.6 | 0.4 | 0.2 | -0.05 | -0.01 |
| 0.7 | 0.3 | 0.4 | -0.05 | -0.02 |
| 0.8 | 0.2 | 0.6 | -0.05 | -0.03 |
| 0.9 | 0.1 | 0.8 | -0.05 | -0.04 |
| 0.6 | 0.4 | 0.2 | -0.10 | -0.02 |
| 0.7 | 0.3 | 0.4 | -0.10 | -0.04 |
| 0.8 | 0.2 | 0.6 | -0.10 | -0.06 |
| 0.9 | 0.1 | 0.8 | -0.10 | -0.08 |

$P(U = 1|x, m, c)$ = prevalence of the unmeasured confounder for exposure level *x*; $P(U = 1|x^*, m, c)$ = prevalence of the unmeasured confounder for exposure level *x**. $\delta$ = difference in the prevalence of the unmeasured confounder between exposure level *x* and *x**; $\gamma$= beta coefficient of the direct effect of *U* on *Y*; *d* =magnitude of bias, is the product of $\delta$

# CHAPTER 6. Parenting practices at 24 to 47 months and IQ at age 8: Effect-measure modification by infant temperament

## 6.1 Preface

This chapter contains the third article contributing to this thesis. The article was published in the journal *PLoS One* in March 2016.[283]

In Chapter 5, the direct effect of temperament on cognitive and academic achievement was examined. After taking into account parenting practices and a range of confounders, the effect sizes of temperament were small (≤0.11 SD). Since temperament had such a small influence on children's cognitive and academic outcomes, the focus of this thesis then shifted to parenting as the exposure, and temperament as an effect modifier. As reviewed in Chapter 2, developmental theories such as the differential susceptibility model suggest that children of difficult temperament may be more susceptible to the influence of negative parenting than children with easy temperament. Chapter 6 focused on the impact of parenting on cognitive abilities, and considered whether negative parenting might have a more pronounced impact on cognitive abilities for children with particular temperament profiles.

In Chapter 6, the associations between two important aspects of parenting practices*, i.e*. parental warmth and control, and cognitive ability was explored. The associations between parenting and IQ were stratified to examine whether the associations differed in children with different temperament profiles, *i.e.* the effect-measure modification. This chapter explains the limitations of the conventional approach used to examine

“interaction” used in many previous studies and uses a more advanced epidemiological approach to analyse and present effect-measure modification analyses. This article may be useful to inform whether parenting interventions and policies should be targeted to parents of children with difficult temperament.

## 6.2 Publication: Parenting practices at 24 to 47 months and IQ at age 8: Effect-measure modification by infant

### 6.2.1 Statement of authorship

| | |
|---|---|
| Title of Paper | Parenting practices at 24 to 47 months and IQ at age 8: Effect-measure modification by infant |
| Publication Status | ☑ Published<br>☐ Accepted for Publication<br>☐ Submitted for Publication<br>☐ Unpublished and Unsubmitted w ork w ritten in manuscript style |
| Publication Details | Chong SY, Chittleborough CR, Gregory T, Lynch JW, Mittinty M, Smithers L. Parenting practices at 24 to 47 months and IQ at age 8: effect-measure modification by infant temperament. *PLoS One*. 2015;11:e0152452. |

### Principal Author

| | | | |
|---|---|---|---|
| Name of Principal Author (Candidate) | Shiau Yun Chong | | |
| Contribution to the Paper | Conceived and designed the study, analysed and interpreted the data, drafted the manuscript. | | |
| Overall percentage (%) | 85 | | |
| Certification: | This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper. | | |
| Signature | | Date | 15-08-2016 |

### Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

vii. the candidate's stated contribution to the publication is accurate (as detailed above);

viii. permission is granted for the candidate in include the publication in the thesis; and

ix. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

| Name of Co-Author | Catherine Chittleborough | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, data acquisition, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | Tess Gregory | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | Murthy Mittinty | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | John Lynch | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | Lisa Smithers | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

### 6.2.2 Abstract

Cognitive development might be influenced by parenting practices and child temperament. We examined whether the associations between parental warmth, control and intelligence quotient (IQ) may be heightened among children in difficult temperament. Participants were from the Avon Longitudinal Study of Parents and Children (n=7044). Temperament at 6 months was measured using the Revised Infant Temperament Questionnaire and classified into 'easy' and 'difficult'. Parental warmth and control was measured at 24 to 47 months and both were classified into 2 groups using latent class analyses. IQ was measured at 8 years using the Wechsler Intelligence Scale for Children and dichotomized (<85 and ≥85) for analysing effect-measure modification by temperament. Linear regression adjusted for multiple confounders and temperament showed lower parental warmth was weakly associated with lower IQ score [β=-0.52 (95% CI 1.26, 0.21)], and higher parental control was associated with lower IQ score [β=-2.21 (-2.95, -1.48)]. Stratification by temperament showed no increased risk of having low IQ in temperamentally difficult children [risk ratio (RR) =0.97 95% CI 0.65, 1.45)] but an increased risk among temperamentally easy children (RR=1.12 95% CI 0.95, 1.32) when parental warmth was low. There was also no increased risk of having low IQ in temperamentally difficult children (RR=1.02 95% CI 0.69, 1.53) but there was an increased risk among temperamentally easy children (RR=1.30 95% CI 1.11, 1.53) when parental control was high. For both parental warmth and control, there was some evidence of negative effect-measure modification by temperament on the risk-difference scale and the risk-ratio scale. It may be more appropriate to provide parenting interventions as a universal program rather than targeting children with difficult temperament.

### 6.2.3 Introduction

Cognitive ability is an important aspect of healthy child development. Intelligence quotient (IQ), derived from intelligence tests, is a marker of future health, academic achievement, occupational outcomes, and social development.[18-20] At the population level, increases in average IQ are associated with improvements in economic growth.[284] Children with lower IQ have increased risk of mortality and morbidity as well as lower occupational status and earnings in adulthood.[17] Cognitive abilities develop in early life through social interactions,[285] and a nurturing environment is particularly important to facilitate cognitive development.[96] For instance, children with supportive parents who engage in learning activities are more ready to learn and develop their cognitive abilities.

Parental warmth and control are two aspects of parenting that are important for children's development. There is some evidence that parental warmth is associated with children's cognitive development.[87,95]. Parents who understand the child's needs may also be more likely to provide support to assist the child in developing learning skills such as mastery, security, autonomy, and self-efficacy.[286] On the other hand, low warmth parenting that uses verbal and physical punishments may hinder the child's cognitive development. High parental control is characterized by behaviours that involve the use of pressure, solving problems for children, and making decisions for the child based on a parental perspective.[89,90] High parental control may be associated with lower levels of intrinsic motivation as children with controlling parenting are less likely to engage in activities and attempts that help them to learn.[287] Children of more controlling parents may also have poorer self-regulation,[89] hence affecting their cognitive development.

Besides parenting practices, individual factors such as the child's temperament also play a role in their development. Children who are high in approach to unfamiliar persons or objects, more adaptable to new environments and more positive in mood are linked to have higher IQ.[144,171] There is also some evidence that self-regulation is associated with IQ.[288,289] Self-regulation, which is the ability to consciously control activity, emotion, and attention, is an emerging component of temperament that is observable from the age of two.[6,290] Studies categorize temperament by combining traits including rhythmicity, approach, adaptability, intensity, and mood to create a construct of difficult temperament have found mixed results. One study reported no association between temperament and IQ,[33] whereas another found that IQ was higher in children with difficult temperament than children with easy temperament.[140] Further studies are needed to determine the influence of temperament on children's IQ.

This study aims to examine whether the associations between parenting warmth, control and IQ differ among children with different temperaments, *i.e*. the possibility that temperament is an *effect modifier* of the association between parental warmth and control on IQ. Our hypothesis that temperament would modify the associations between parenting warmth, control and IQ is based on the following rationale. Temperament is the individual differences in styles that are observable from early childhood.[27] In this study we use the Revised Infant Temperament Questionnaire (RITQ) measured at 6 months of age, and dichotomized into difficult versus easy or other categories as has been recommended.[291] It is possible that some temperament characteristics may help buffer children from adverse effects of negative parenting practices. For instance, children who are temperamentally easy are more capable at controlling their own emotions and more

adaptable to their environment. Therefore, they are better in finding ways to fit in into their environmental context and less likely to be affected by parenting practices than temperamentally difficult children who have problems controlling their emotions and are less adaptable to their environment. If this hypothesis is true, we would expect that, given similar circumstances of growing up in a less positive parenting environment, children who are temperamentally easy would have fewer adverse outcomes compared to children with more difficult temperament. Understanding the associations between parenting practices (warmth and control) on children with different temperaments will help to determine whether parenting interventions should be targeted to children with specific temperaments or to all children.

In this current study, we are interested in effect-measure modification by temperament because our interest is to intervene on parenting practices (warmth and control) rather than child temperament. We focus on parenting practices because of the preponderance of interventions currently being used that are designed to improve parenting.[14,292] Interventions on temperament of infants and young children are typically through improving parent-child social relationships, and reinforcement activities delivered by parents or teachers.[293] Although parenting and temperament are measured at different time points (temperament at 6 months, parenting at 24 to 47 months), this study is interested in 'effect-measure modification' (whether the association of warmth, control, and IQ differs in children with different temperament), but not 'mediation' (examining the direct effect of temperament on IQ, and the indirect effect of temperament on IQ that goes through parenting). Therefore, the hypothesis put forward in this paper is an issue of effect-measure modification. In practice, the distinctions between effect-measure

modification and interaction are often ignored. Traditionally, effect-measure modification and interaction are tested in regression analyses by including a product term. However, using regression models, it is not clear whether the coefficient of the interaction term should be interpreted as effect-measure modification or interaction, or both, or neither.[135] Interaction is widely used when interpreting findings that should be interpreted as effect-measure modification.[135] The distinctions between effect-measure modification and interaction are important especially when considering potential intervention and policy recommendations.[135,233] For public health intervention purposes, only one intervention is considered in effect-measure modification, *i.e.* the main exposure, while in interaction, potential intervention on both exposures is considered. Conceptually, effect-measure modification occurs when the effect of the main exposure (parenting) on an outcome (IQ) differs across strata of a second exposure (temperament).[135] Effect-measure modification on the risk-difference scale can be written as:

$$E\left[Y_{P_1}|T=t_1, C=c\right] - E\left[Y_{P_0}|T=t_1, C=c\right] \neq E\left[Y_{P_1}|T=t_0, C=c\right] - E\left[Y_{P_0}|T=t_0, C=c\right] \quad (1)$$

where *Y* denotes the outcome under study, *T* denotes the effect modifier, *P* denotes the exposure of interest, and *C* denotes a set of confounders. Equation 1 is read as the expectation of the difference in outcome (*Y*) between low warmth parenting ($P_1$) and high warmth parenting ($P_0$) in stratum $T_1$ of effect modifier (conditioned on *C*) is not equal to the expectation of the difference in outcome (*Y*) between low warmth parenting ($P_1$) and high warmth parenting ($P_0$) in stratum $T_0$ (conditioned on *C*). In effect-measure modification (Equation 1), the relationship between *P* and *T* are asymmetric, *i.e*. only the effect of *P* on *Y* is of interest, *T* only concerns whether the effect of *P* on *Y* differ across different value of *T*. Interaction is different from effect-measure modification in that it

concerns whether the joint effect of the two exposures differs from the combined independent effects. Interaction on the risk-difference scale can be written as:

$$E\left[Y_{P_1T_1}\middle|C=c\right]-E\left[Y_{P_0T_1}\middle|C=c\right]\neq E\left[Y_{P_1T_0}\middle|C=c\right]-E\left[Y_{P_0T_0}\middle|C=c\right] \quad (2)$$

In interaction (Equation 2), the role of *P* and *T* are symmetric, *i.e*. both *P* and *T* have causal effects on *Y*. Given that our interest is to intervene on *P* (parenting) but not on *T* (temperament), it is essential to understand the effect of *P* on *Y* rather than the joint effect of *P* and *T* on *Y*, therefore effect-measure modification is of interest in this current study rather than interaction.

The presence or absence and the magnitude of effect-measure modification depend on which scale the association is measured - risk-difference or risk-ratio scale. The risk-difference scale estimates the extent to which the effect of the two exposures, *i.e*. parenting and temperament operating together exceeds the effect of each added together.[135] The risk-difference scale helps us to identify target groups as it allows us to see the absolute gain in outcome if an intervention is targeted at certain subgroup, which can help making public policy decisions when resources are finite (for example, see Appendix E). On the other hand, the risk-ratio scale estimates the extent to which the effect of both exposures operating together exceeds the product of the effects of the two exposures. It is unclear how to interpret effect-measure modification on the risk-ratio scale, as it does not allow us to determine which subgroup to treat, but it is thought to be useful for investigation of possible biological pathways.[294] While there is consensus that the risk-difference scale is considered more important for public policy action and interventions, the risk-difference scale is often not reported in many studies.[136] Even though effect-measure modification is widely studied in epidemiological research, most

studies have not provided enough information about the size and statistical significance of the effect-measure modification on both the risk-difference and risk-ratio scale.[135,136]

This study examined the *effect-measure modification* of the association between parental practices (warmth and control) and IQ by child temperament. We reported the effect-measure modification on both the risk-difference and risk-ratio scale for transparency and completeness,[137,138] however, results on the risk-difference scale are more pertinent to this research question because of the implications for public health intervention.

### 6.2.4 Methods

***Study design***

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a population-based prospective study investigating the influence of genetic and environmental characteristics on health and development in parents and children. A total of 14541 pregnant women who resided in the Southwest of England with expected delivery date between 1st April 1991 and 31st December 1992 were recruited, and this includes 72% of the eligible mothers.[198] The ALSPAC sample is broadly representative of the population living in Avon and the whole of Britain at the time although ethnic minorities and unmarried couples were slightly underrepresented.[198] Follow up assessments have been administered frequently through questionnaires and clinical assessments. The length of follow-up and the breadth of data collection provide valuable data that can be used as confounders. The ALSPAC sample consists of 13988 infants who were alive at 1 year (Figure 6.1). The study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/).

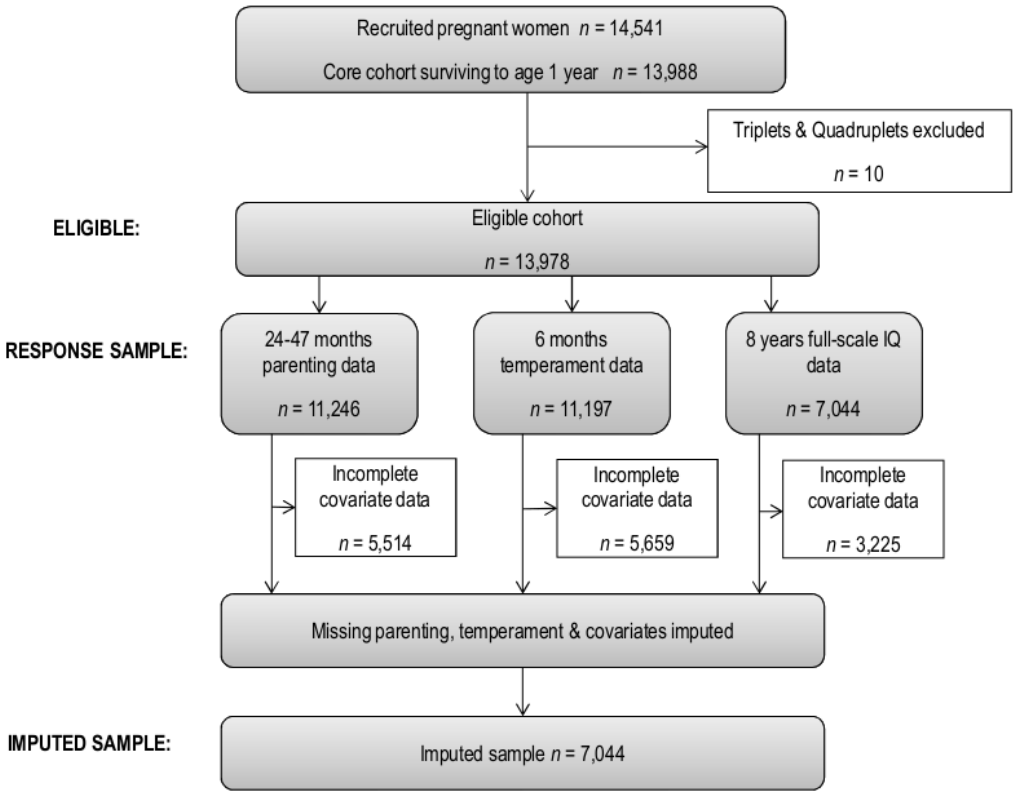


**Figure 6.1 Eligible cohort and numbers included**

### *Ethics statement*

Ethical approvals were obtained from the Human Research Ethics Committee, University of Adelaide, ALSPAC Law and Ethics committee and the four Local Research Ethics committees: Southmead, Frenchay, Bristol and Weston health authorities. Written consent was obtained from the original participants and from the parents, next of kin, caretakers, or guardians on behalf of the children enrolled in ALSPAC.

### *Measures*

### *IQ (the outcome)*

IQ was assessed using the Wechsler Intelligence Scale for Children (WISC-III $^{UK}$) when the children were 8 years of age.[199] Children's scores on five verbal and five performance subtests were summarized into domains of verbal IQ and performance IQ, which were

then combined to yield the full-scale IQ. All tests were administered by the ALSPAC psychology team. To reduce the length of the session, alternate items were used for all subtests, except the coding subtest which was administered in full. Individual items within each subtest were summed and multiplied by 2 for picture completion, information, arithmetic, vocabulary, comprehension, and picture arrangement; multiplied by 5/3 for similarities; and multiplied by 3/2 for object assembly and block design. This made the raw scores comparable to those that would have been achieved had the full test been administered. Raw scores were converted to age-scaled scores according to standard procedures.[199] The IQ scores were standardized on a normal British population in the early 1990s to have a mean of 100 and standard deviation of 15.

IQ was used as a continuous variable in regressions but we used dichotomized outcome when analysing effect-measure modification for ease of interpretation of results.[138] IQ was dichotomized at <85 ('low IQ') and ≥85 ('normal to high IQ'). Although other IQ cut-points were tested and results found to be similar (Appendix H), the cut-point at IQ 85 reflected a balance between having an IQ score at one SD below the mean and an adequate sample in each cell for effect-measure modification analysis. Furthermore, this group of children have lower human capital and greater likelihood of needing supportive resources.[21] Of the participants for whom full-scale IQ score was available, 12.0% (*n*=847) were classified as having a low IQ score.

***Parenting warmth and control (the exposures of interest)***

Parenting warmth and control were reported by mothers using items selected from ALSPAC questionnaires. There were eight items for warmth and nine items for control

across 24 to 47 months (see Table 6.1). The parenting measures correspond to warmth and control which are identified as important aspects of parenting.[295] In order to encourage continued participation of ALSPAC members, parenting questions were phrased in a way that minimized offence. Mothers answered the questions with 3- to 5-point Likert scales that ranged from ‘never’ to ‘frequently’.

**Table 6.1 Parenting measures in the ALSPAC questionnaires**

| Parenting questions | Age | Responses |
|---|---|---|
| ***Parental warmth*** | | |
| Mother smacks child during tantrums | 42 months | Never, rarely, once a month, once a week, daily |
| Mother shouts at child when naughty | 24, 42 months | Every day, several times a week, once a week, rarely, never |
| Child is slapped | 24 months | Every day, several times a week, once a week, rarely, never |
| Child is kissed or cuddled | 24, 38, 42 months | Nearly every day, 3-5 times per week, 1-2 times per week, < once per week, never |
| Child is praised | 24 months | Every day, several times a week, once a week, rarely, never |
| ***Parental control*** | | |
| Mother reasons with child during tantrums | 30, 42 months | Often, sometimes, never |
| Child has some choice at meals | 30, 42 months | Free choice, select choice, no choice |
| Child has some choice with clothes | 30, 42 months | Free choice, select choice, no choice |
| Parent and child have battle of wills | 30, 42, 47 months | Never, rarely, sometimes, frequently |

We used latent class analysis (LCA) in SAS statistical software version 9.2 (SAS Institute, Inc, Cary, North Carolina) to investigate the patterns of parenting in our study population. LCA identifies groups of parents with distinct characteristics based on their responses to the parenting items. LCA is a model-based cluster analysis that provides a set of class-assignment probabilities for each respondent.[226,231,232] The latent classes were created based on two parameters: the latent class membership probabilities and the conditional item-response probabilities.[231] The latent class membership probabilities estimate the proportion of the population that fall into a given class. The conditional item-response probabilities are probabilities of responding to each question, given the class membership. The item-response probabilities are conceptually similar to factor loadings

in factor analysis. However, they are probabilities rather than coefficients.[227] Individual posterior probabilities of membership in each latent class were obtained by applying Bayes' theorem.[231] Models with two-, three-, and four-class solutions were tested and the best model was chosen based on the log-likelihood, Bayesian Information Criterion, and the face validity of the classes.[226] A two-class solution was chosen for both warmth and control. Each respondent was assigned a class based on their highest probability of membership. The classes were given descriptive labels based on consensus of the authors after reviews of each class's characteristics.

***Temperament (the potential effect-modifier)***

Temperament was assessed when the infants were 6 months old using an adaptation of the RITQ.[202] RITQ is a valid and reliable measure of temperament.[6,42] The internal consistency of the adapted RITQ is consistent with the internal consistencies of the original questions included in the RITQ.[244] Scores for each subscale were derived according to the procedure introduced in the test manual.[202]

Infant temperament was classified into two groups based on the scores on five temperament subscales (rhythmicity, approach, adaptability, intensity and mood).[41] Infants were defined as difficult if they scored greater than the mean on at least four of the five subscales and greater than one standard deviation above the mean on at least two subscales. Infants who did not satisfy the difficult definition were classified into a single group called 'easy or other'. The classification of infant temperament utilized ALSPAC-specific norms, rather than norms from the test manual, which we have previously demonstrated may reduce misclassification of temperament.[244] There were

1655 (15%) of the infants with temperament data available who were classified as difficult.

***Confounding***

Confounders were determined *a priori*[213] based on factors that might confound the association between parenting and IQ. Confounders included parent level factors such as indicators of socio-economic position, parents' physical and mental health and child level factors such as intrauterine growth (see Table 6.2 for full list).

Measures of socio-economic position were obtained from mothers' self-reported questionnaires from pregnancy until 8 weeks postpartum. Mother's and partner's education were categorised into 4 levels consistent with the UK education system: none or certificate of secondary education (CSE) or vocational; O level; A level; degree. Parents' social class was based on the highest occupation of either parent, and categorised using standard UK classifications of occupation, ranging from class I (highest), II, III-non-manual, III-manual, IV, and V (lowest). Mothers' financial difficulties in affording food, clothing, heating, rent or mortgage was assessed at 32 weeks gestation with possible responses ranging from 1 (very difficult) to 4 (not difficult). The sum of the scores of each of the 5 items was subtracted from 20 to derive a total financial difficulties score. A total score of 0 represented no financial difficulties and 15 represented maximum financial difficulties. Mothers with a score of 9 and above were defined as experiencing financial difficulties and this included approximately 10% of the cohort.[296] Household crowding was categorised according to whether there were ≤ 1 or > 1 person per room. Home ownership was categorised into owned or mortgaged and rented or other. Mother's

social support was measured using a set of 10 items specifically designed for the cohort. These items represented statements in relation to financial, emotional and instrumental support the mothers received from their partners, friends, families and official agencies. Scores were summed from 0 (lowest) to 30 (highest). The social support score was separated into three groups of equal size ('low', 'medium', and 'high'). Mother's marital status was classified as 'married or cohabiting' and 'not married or not living with partner'.

At 8 months postpartum, mothers self-rated their own health as always well, mostly well, often feel unwell, or hardly ever well. Mothers' and partners' depression was assessed using the ten items from the Edinburgh Postnatal Depression Scale (EPDS) at 18 weeks antenatal and 8 weeks postpartum. The EPDS is validated for use among parent groups during the postnatal and pregnancy period.[297,298] Mothers and partners with a score of ≥12 in either measurement time were considered as displaying depressive symptoms.[297]

Mothers were asked whether or not they smoked during the first three months of pregnancy. Mothers' alcohol consumption during the first three months of pregnancy was classified as never, < 1 glass per week, and ≥ 1 glasses per week or ≥ 1 glasses per day. Child's gestational age and birth weight were collected by ALSPAC staff from obstetric data.

***Analysis***

Correlations between parenting, temperament, IQ and all confounders are included in Appendix F. A series of analyses were undertaken to examine the effect of the parenting

dimensions (warmth, control) on children's full-scale IQ (continuous score) using linear regression modelling. In the first step, univariable associations between parenting warmth on IQ and parenting control on IQ were examined separately (Model 1). We then adjusted for temperament (Model 2). In Model 3, parenting warmth model was adjusted for parenting control and all other confounding variables described above, parenting control model was adjusted for parenting warmth and all other confounding variables.

We estimated effect-measure modification on both risk-difference and risk-ratio scales as outlined by Knol and VanderWeele.[138] Using dichotomized IQ, Poisson regressions were used to estimate risk ratio (RR) estimates for each stratum of parenting (P) and temperament (T): i) high warmth parenting and easy temperament ($\mathrm{RR}_{\mathrm{P_0T_0}}$)(reference category); ii) high warmth parenting and difficult temperament ($\mathrm{RR}_{\mathrm{P_0T_1}}$); iii) low warmth parenting and easy temperament ($\mathrm{RR}_{\mathrm{P_1T_0}}$); and iv) low warmth parenting and difficult temperament ($\mathrm{RR}_{\mathrm{P_1T_1}}$). Next, the RR for parenting within strata of temperament was estimated. Poisson models with robust errors were used to estimate RR due to convergence problems with log-binomial models.

A relative excess risk due to interaction (RERI) was calculated to give the measure of effect-measure modification on the risk-difference scale, and 95% CIs were obtained by the delta method.[299]

$$RERI = RR_{P_1T_1} - RR_{P_0T_1} - RR_{P_1T_0} + RR_{P_0T_0} \qquad (3)$$

RERI > 0 indicates the effect-measure modification is positive (the effect of the exposure and the effect modifier operating together is greater than the effect of each added together), RERI < 0 indicates the effect-measure modification is negative, RERI of 0

indicates there is no effect-measure modification on the risk-difference scale. Effect-measure modification on the risk-ratio scale is taken as:

$$Ratio\ of\ RRs = \frac{RR_{P_1T_1}\ X\ RR_{P_0T_0}}{RR_{P_0T_1} X\ RR_{P_1T_0}} \qquad (4)$$

If the ratio of RRs >1, the effect-measure modification is positive (the effect of the exposure and the effect modifier operating together is greater than the product of the effect of the exposure and the effect modifier). A ratio of RRs < 1 indicates the effect-measure modification is negative. A ratio of RRs = 1 means the effect of both exposures together is equal to the product of the effect of the two exposures considered separately.

Analyses were performed using STATA version 13.0 (Stata Corp, College Station, Texas).

***Multiple imputation for missing data***

We used multiple imputation by chained equation to impute missing data. Imputed datasets were generated under the missing at random assumption that the probability of data being missing is dependent on the observed data.[235] Variables included in the imputation model were parenting warmth, control, temperament, all confounders, breastfeeding, HOME inventory, all three measures of IQ (full-scale IQ, verbal IQ, and performance IQ) and interaction terms between parenting and temperament. Breastfeeding and HOME inventory variables were two auxiliary variables that were added to the imputation model because they are related to the outcome (IQ) and may enhance the prediction of missing values. Fifty cycles of regression switching were undertaken and 20 imputed datasets were generated. We used the multiple imputation then deletion technique[242] where analyses were conducted on respondents only with non-imputed outcome data. All analyses were performed on imputed data (n=7044).

### 6.2.5 Results

Table 6.2 shows the socio-demographic characteristics of ALSPAC response sample, respondents who had complete data on IQ, parenting, temperament, and all covariables (complete case, *n*=3665), and the imputed sample (*n*=7044). Participants with complete data had a higher proportion of mothers with higher warmth and lower control parenting, higher education, and higher social class, and lower proportions of mothers with financial difficulties, who smoked during pregnancy and were unmarried or not living with partner. Lower proportions of children in the complete case sample were non-white, and had low IQ.

**Table 6.2 IQ, parenting, temperament, and demographic characteristics of response, complete case and imputed sample**

| | Response sample | | Complete data sample (n=3665) | Imputed sample (n=7044) |
|---|---|---|---|---|
| | n | Mean (SD) or % | Mean (SD) or % | Mean (SD) or % |
| ***Total IQ, 8 years*** | 7044 | 104.2 (16.5) | 106.3 (16.4) | 104.2 (16.5) |
| Low IQ (<85) | 847 | 12.0 | 9.7 | 12.0 |
| Normal to high IQ (≥85) | 6197 | 88.0 | 90.3 | 88.0 |
| ***Parenting, 24-47 months*** | | | | |
| Parental warmth | | | | |
| High warmth | 5205 | 45.7 | 48.6 | 46.9 |
| Low warmth | 6184 | 54.3 | 51.4 | 53.1 |
| Parental control | | | | |
| Less controlling | 6522 | 58.0 | 59.1 | 59.4 |
| High controlling | 4724 | 42.0 | 40.9 | 40.6 |
| ***Temperament, 6 months*** | | | | |
| Easy or other | 4169 | 85.1 | 86.1 | 84.6 |
| Difficult | 1655 | 14.9 | 13.9 | 15.4 |
| ***Child covariables*** | | | | |
| Sex, female | 13976 | 48.3 | 49.3 | 50.0 |
| Birth weight (grams) | 13798 | 3392.0 (559.3) | 3446.2 (516.7) | 3414.6 (554.2) |
| Gestational age (weeks) | 13976 | 39.4 (1.9) | 39.5 (1.6) | 39.4 (1.9) |
| Ethnicity, non-white | 12083 | 5.0 | 2.1 | 4.1 |
| ***Parent covariates*** | | | | |
| Maternal age (years) | 13978 | 28.0 (5.0) | 29.2 (4.3) | 28.0 (5.0) |
| Maternal smoking in first 3 months pregnancy | 13158 | 25.0 | 14.4 | 18.3 |
| Maternal alcohol consumption in first 3 months pregnancy | | | | |
| Never | 5917 | 45.5 | 44.3 | 44.1 |
| Less than 1 glass per week | 5034 | 38.7 | 41.3 | 40.7 |
| One or more glass per week | 1804 | 13.9 | 13.1 | 13.3 |
| One or more glass per day | 250 | 1.9 | 1.3 | 1.8 |
| No partner / not living with partner | 13179 | 8.7 | 1.9 | 5.6 |
| Home ownership, rented/other | 13027 | 26.6 | 11.7 | 16.9 |
| Household crowding, >1 person/room | 12084 | 6.9 | 2.6 | 4.1 |
| Maternal highest education | | | | |
| None / CSE / vocational | 3728 | 26.7 | 16.8 | 22.2 |
| O level | 4296 | 30.7 | 35.1 | 34.9 |
| A level | 2794 | 20.0 | 28.6 | 26.7 |
| Degree or higher | 1600 | 11.5 | 19.6 | 16.2 |
| Partner's highest education | | | | |
| None / CSE / vocational | 4124 | 34.5 | 21.2 | 28.9 |
| O level | 2540 | 21.3 | 22.1 | 21.8 |
| A level | 3105 | 26.0 | 30.1 | 27.8 |
| Degree or higher | 2171 | 18.2 | 26.6 | 21.5 |
| Parental highest social class | | | | |
| Professional / managerial (I/II) | 6342 | 55.1 | 66.3 | 60.0 |
| Skilled manual / non-manual (III) | 4481 | 38.9 | 30.9 | 35.5 |

| | Response sample | | Complete data sample (n=3665) | Imputed sample (n=7044) |
|---|---|---|---|---|
| | n | Mean (SD) or % | Mean (SD) or % | Mean (SD) or % |
| Unskilled / semiskilled manual (IV/V) | 682 | 5.9 | 2.8 | 4.5 |
| Financial difficulties | 12086 | 10.0 | 5.7 | 8.1 |
| Social support | | | | |
| Low | 4142 | 38.1 | 32.3 | 36.1 |
| Medium | 3736 | 34.4 | 36.8 | 35.6 |
| High | 2999 | 27.6 | 30.9 | 28.3 |
| Maternal depression | 10929 | 21.9 | 14.8 | 17.9 |
| Partner's depression | 7605 | 7.9 | 5.5 | 7.7 |
| Maternal health, often unwell / hardly ever well | 11317 | 5.5 | 4.9 | 5.1 |

*CSE* Certificate of Secondary Education, *IQ* Intelligence Quotient

Table 6.3 shows the associations between parenting and IQ at 8 years (*n*=7044) using linear regressions. Models 1 and 2 provide some evidence that low parental warmth and high parental control were associated with lower IQ at 8 years. In the fully-adjusted model (Model 3), children experiencing low warmth parenting had 0.52 (95% CI -1.26, 0.21) point lower IQ than children experiencing high warmth parenting. Children whose parents demonstrated high control had 2.21 (95% CI: -2.95,-1.48) point lower IQ than children whose parents demonstrated low control. The association between difficult temperament and IQ was negligible (β = -0.12, 95% CI -1.13, 0.90)

**Table 6.3 Association between parenting warmth and control, and child temperament on children's IQ (Imputed sample, n=7044)**

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | **B** | **95% CI** | ***p*** | **β** | **95% CI** | ***p*** | **β** | **95% CI** | ***p*** |
| Warmth | | | | | | | | | |
| High | Ref | | | Ref | | | Ref | | |
| Low | -3.21 | -3.99, -2.42 | <0.001 | -3.04 | -3.84, -2.24 | <0.001 | -0.52 | -1.26, 0.21 | 0.166 |
| Temperament | | | | | | | | | |
| Easy or Other | | | | Ref | | | Ref | | |
| Difficult | | | | -0.14 | -1.27,1.00 | 0.976 | -0.12 | -1.13, 0.90 | 0.824 |
| Control | | | | | | | | | |
| Less controlling | Ref | | | Ref | | | Ref | | |
| High controlling | -3.33 | -4.13, -2.53 | <0.001 | -3.29 | -4.11, -2.47 | <0.001 | -2.21 | -2.95, -1.48 | <0.001 |
| Temperament | | | | | | | | | |
| Easy or Other | | | | Ref | | | Ref | | |
| Difficult | | | | -0.25 | -1.49, 0.97 | 0.683 | -0.12 | -1.13, 0.90 | 0.824 |

Model 1 is unadjusted. Model 2 is adjusted for temperament. In Model 3, parenting warmth is adjusted for temperament and all the covariables (maternal smoking, alcohol consumption, birth weight, gestation at birth, sex, ethnicity, maternal age, partner status, financial difficulties, maternal and partner's education, parental social class, home ownership, household crowding, maternal health, social support, maternal and partner's depression), parenting control is adjusted for parenting warmth and all the covariables.

Table 6.4 shows effect-measure modification of the association between parenting warmth and IQ by temperament. Among temperamentally easy children, low warmth parenting was associated with 12% higher risk of low IQ (95% CI 0.95, 1.32), whereas in the stratum of temperamentally difficult children, low warmth parenting did not increase the risk of low IQ (RR 0.97 95% CI 0.65, 1.45). Compared with the reference category of children who had high warmth parenting and easy or other temperament, children with low warmth parenting or difficult temperament, or both, had about 12 to 17% increased risk of having low IQ. The RERI of -0.19 (95% CI -0.65, 0.27) indicated small negative effect-measure modification by temperament on the risk-difference scale, *i.e.* the combined risks of both low warmth parenting and difficult temperament (RR 1.12) was lower than expected (RR 1.31) when based on summing the individual risks of low warmth parenting (RR 1.14) and difficult temperament (RR 1.17). Similarly, for effect-measure modification on the risk-ratio scale, the ratio of RRs was 0.84 (95% CI 0.56, 1.25) indicating that the combined risks of both low warmth parenting and difficult temperament (RR 1.12) was lower than expected (RR 1.33) when based on multiplying the individual risks of low warmth parenting and difficult temperament.

**Table 6.4 Effect-measure modification of the effect of parenting warmth on IQ (<85) by child temperament (Imputed sample, n=7044)**

| | High warmth parenting | | Low warmth parenting | | RR (95% CI) for low warmth parenting within strata of temperament type |
|---|---|---|---|---|---|
| | N Low IQ/High IQ | RR (95% CI) | N Low IQ/High IQ | RR (95% CI) | |
| Easy or other temperament | 283/2565 | 1.00 (Ref) | 433/2749 | 1.14 (0.97, 1.34), p=0.106 | 1.12 (0.95, 1.32), p=0.182 |
| Difficult temperament | 47/391 | 1.17 (0.85, 1.61), p=0.327 | 84/492 | 1.12 (0.86, 1.45), p=0.393 | 0.97 (0.65, 1.45), p=0.872 |

Effect-measure modification on the risk-difference scale: RERI= -0.19 (-0.65, 0.27), p=0.413
[RERI=1.12-1.14-1.17+1.00 = -0.19 when estimated from the table]

Effect-measure modification on the risk-ratio scale: Ratio of RRs= 0.84 (0.56, 1.25), p=0.385
[Ratio of RRs=1.12/(1.14 x 1.17) = 0.84 when estimated from the table]

RRs are adjusted for parenting control, maternal smoking, alcohol consumption, birth weight, gestation at birth, sex, ethnicity, maternal age, partner status, financial difficulties, maternal and partner's education, parental social class, home ownership, household crowding, maternal health, social support, maternal and partner's depression

Table 6.5 shows the effect-measure modification of the association between parental control and IQ by child temperament. The increased risk of low IQ associated with high control parenting was 30% in easy or other temperament. There was no increased risk of having low IQ in difficult temperament children (RR 1.02 95% CI 0.69, 1.53) but the confidence intervals were wide in the stratum of temperamentally difficult children due to smaller numbers. Compared with children who had easy or other temperament and less controlling parenting, children with easy or other temperament and high control parenting had a 31% increased risk of having low IQ, children with difficult temperament and less controlling parenting and children with both difficult temperament and high control parenting had a 18% increased risk of having low IQ.

For parental control, the RERI of -0.30 (95% CI -0.78, 0.18) suggested a small negative effect modification, although confidence intervals were wide. The measure of effect-measure modification on the risk-ratio scale was 0.77 (95% CI 0.52, 1.15) also indicated a small negative effect-measure modification.

**Table 6.5 Effect-measure modification of the effect of parenting control on IQ (< 85) by child temperament (Imputed sample, n=7044)**

| | Less controlling parenting | | High controlling parenting | | RR (95% CI) for high control parenting within strata of temperament type |
|---|---|---|---|---|---|
| | N Low IQ/High IQ | RR (95% CI) | N Low IQ/High IQ | RR (95% CI) | |
| Easy or other temperament | 349/3195 | 1.00 (Ref) | 367/2119 | 1.31 (1.12, 1.53), p=0.001 | 1.30 (1.11, 1.53), p=0.001 |
| Difficult temperament | 77/551 | 1.18 (0.90, 1.53), p=0.228 | 54/332 | 1.18 (0.89, 1.59), p=0.261 | 1.02 (0.69, 1.53), p=0.907 |

Effect-measure modification on the risk-difference scale: RERI=-0.30 (-0.78, 0.18), p=0.226
[Due to rounding, RERI=1.18-1.31-1.18+1.00 = -0.31 when estimated from the table]

Effect-measure modification on the risk-ratio scale: Ratio of RRs= 0.77 (0.52, 1.15), p=0.204
[Due to rounding, ratio of RRs=1.18/(1.31 x 1.18) = 0.76 when estimated from the table]

RRs are adjusted for parenting warmth, maternal smoking, alcohol consumption, birth weight, gestation at birth, sex, ethnicity, maternal age, partner status, financial difficulties, maternal and partner's education, parental social class, home ownership, household crowding, maternal health, social support, maternal and partner's depression

Graphical illustrations of the effects of parenting warmth and control across stratum of temperament are included in Appendix G. Other IQ cut points (80 and 90) were also tested and the results were similar (Appendix H).

### 6.2.6 Discussion

This study found some evidence that parental warmth and control are associated with children's cognitive development. Children whose parent's demonstrated low warmth at age 24 to 47 months had a 0.52-point lower IQ at age 8 than children whose parents demonstrated high warmth. The study also found that high controlling parenting was associated with 2.21-point lower IQ than less controlling parenting. Effect sizes of parenting practices were small (warmth: 0.03 standard deviation; control: 0.15 standard deviation) but may have an important impact at a population level.[21] It has been suggested that parental warmth influences children's IQ by providing more support in problem solving, more engagement in positive parent-child interaction, and increased likelihood to encourage exploration and task persistence.[300] Meanwhile, high controlling parenting may have restricted the child's ability to make autonomous choices, and impeded the child's free expression of feeling and thinking, which in turn hinders their cognitive development.

The results of the effect-measure modification analyses provide some evidence to suggest that the associations between parenting practices (warmth and control) on childhood IQ differ according to temperament, although results need to be interpreted cautiously because the confidence intervals were wide in the strata of difficult temperament due to smaller numbers of children. We hypothesized that the associations between low

parental warmth or high control and IQ would be more prominent among children with a difficult temperament, but there was almost no evidence of an exacerbated risk of lower IQ in this stratum. Instead, low warmth and high control parenting was associated with higher risks of low IQ among children with easy and other temperaments. This was surprising as previous research has suggested that children with easy temperament might be more adaptable or less susceptible to parenting practices.[168,301] Compared with children who have easy temperaments, there was a small 12-18% increased risk of lower IQ among children with a difficult temperament and therefore it is important that children with difficult temperaments are supported to realize their full cognitive potential. This might require different types of parenting support and this may be the subject of further research.

For both parenting warmth and control, results on effect-measure modification showed that there was no evidence that parenting interventions should be targeted to children with difficult temperament. Although we found that children with easy or other temperament may be more susceptible to low warmth or high control parenting than children with difficult temperament, children with easy temperament comprise a much larger proportion of the population (85%). As such, rather than targeting children with specific temperament, it may be more appropriate to provide parenting interventions as a universal program. While there is some evidence of effect-measure modification by temperament on the risk-ratio scale, it is difficult to determine the applicability of effect-measure modification on the risk-ratio scale for this research question.

This study has several advantages over previously published research on the association between parenting, temperament and IQ. First, we were able to make better causal inferences by adjusting for a wider range of potential confounders than have been used in many previous studies.[134,302] especially when studying the effect-measure modification. However, it is possible that the results of this longitudinal cohort study remain open to residual and unmeasured confounding, as with all cohort studies. Second, the differential effect of parenting on IQ by child temperament was examined based on a strict definition of effect-measure modification. Other studies[302,303] have not differentiated effect-measure modification from interaction, and results are often not interpreted correctly. Third, assessing effect-measure modification on both the risk-difference and risk-ratio scale is transparent and provides information for readers to draw conclusions about effect-measure modification and the implications of the results. To our knowledge, this is the first study that has investigated the effect-measure modification by temperament on the association between parenting and IQ on both risk-difference and risk-ratio scales. Fourth, data were from a population-based prospective longitudinal study with a large sample, and we used multiple imputation to address potential bias due to missing data. However, several study limitations should also be noted. While the sample was representative of the population in the United Kingdom, the study sample was not very culturally diverse and since parenting styles may vary across cultures, we cannot generalize the results of this study to other ethnic groups or cultures. Future studies with more diverse cultural groups are required. In addition, although we have a large sample size (n=7044), the wide confidence intervals from the effect-measure modification may be influenced by the small number of children with difficult temperament which reduces the power to detect effect-measure modification. It is also possible that children's IQ

might affect parenting, however we are unable to examine this due to the temporal order in which data were collected.

In summary, this study showed small effect sizes of parenting warmth and control at age 24 to 47 months on children's IQ at age 8. We found no increased risk of low IQ as a result of parental warmth or control in temperamentally difficult children. As such, to improve children's IQ, it may be more appropriate to offer interventions to improve parental warmth and decrease parental control as a universal program, rather than targeting to parents who have children with difficult temperament.

### 6.2.7 Online appendices

**Appendix E Effect-measure modification**

The presence or absence and the magnitude of effect-measure modification depend on which scale the association is measured - risk-difference or risk-ratio scale.[304] The effect-measure modification on the risk-difference scale, the relative excess risk due to interaction (RERI) is calculated as:[138,233]

$$RERI = RR_{P_1T_1} - RR_{P_0T_1} - RR_{P_1T_0} + RR_{P_0T_0}$$

where *P* denotes parenting (0=high warmth; 1=low warmth) and *T* denotes temperament (0=easy/other; 1=difficult). RERI = 0 means no effect-measure modification; RERI > 0 means positive effect-measure modification; RERI< 0 means negative effect-measure modification.

The effect-measure modification on the risk-ratio scale, ratio of RRs is taken as:

$$\frac{RR_{P_1T_1}RR_{P_0T_0}}{RR_{P_0T_1}RR_{P_1T_0}}$$

Ratio of RRs = 1 means no effect-measure modification; ratio of RRs > 1 means positive effect-measure modification; ratio of RRs < 1 means negative effect-measure modification.

Examples 1 to 3 in the following were taken from Vanderweele[233] to illustrate the scale dependence of effect-measure modification and the public health significance of the risk-difference scale.

1. An example of when effect-measure modification occurs on the risk-difference but not risk-ratio scale

| | No alcohol | Alcohol |
|---|---|---|
| Non-smokers | 0.02 | 0.05 |
| Smokers | 0.04 | 0.10 |

**Risk-difference scale:** 0.10-0.05-0.04+0.02 = 0.03

**Risk-ratio scale:** (0.10 x 0.02) / (0.05 x 0.04) = 1

2. An example of when effect-measure modification occurs on the risk-ratio but not risk-difference scale

| | No alcohol | Alcohol |
|---|---|---|
| Non-smokers | 0.02 | 0.05 |
| Smokers | 0.07 | 0.10 |

**Risk-difference scale:** 0.10-0.05-0.07+0.02 = 0

**Risk-ratio scale:** (0.10 x 0.02) / (0.05 x 0.07) = 0.57

3. An example of effect-measure modification on the risk-difference scale for public health interest

| | No drug | Drug | Risk-difference | Risk-ratio |
|---|---|---|---|---|
| Genotype A | 20 | 10 | 20-10=10 | 20/10=2 |
| Genotype B | 3 | 1 | 3-1=2 | 3/1=3 |

The risk-difference allows us to see the absolute gain in outcome if an intervention is targeted at certain subgroup. From the data above, the effect of drug on the risk-difference scale amongst individuals with genotype A is greater than the risk-difference amongst individuals with genotype B (10 vs 2). If we had 100 doses of the drug, we can improve the outcome in 10 additional individuals if the drug is given to people with genotype A. If the drug is given to all people with genotype B, we can improve outcome in 2 additional individuals. We would want to give the drugs to people with genotype A. The risk-ratio scale, however, may indicate the wrong subgroup to be targeted for intervention. As we can see from the data above, the risk-ratio is larger for subgroup with genotype B than for subgroup with genotype A (2 vs 3).

4. An example of effect-measure modification on the risk-ratio scale for investigation of disease aetiology

| | No asbestos | Asbestos |
|---|---|---|
| Non-smokers | 1 | 3 |
| Smokers | 10 | 80 |

From the data above, there is evidence of effect-measure modification on the risk ratio scale [80 / (10 x 3) = 2.67]. This suggests that there may be a multiplicative relation between smoking and asbestos exposure in relation to lung cancer risk.[305] One possible explanation may be that that smoking damages the cilia in the airways and thereby opens up for exposure from contaminated particles. However, it has been pointed out there is potential danger to use statistical interaction to draw conclusion about biological interaction.[137,306]

## Appendix F Pearson's correlation coefficients[a]

| | Warmth | Control | Temperament | IQ | Birth weight | Gestational age | Sex | Ethnicity | Maternal age | Maternal smoking | Maternal alcohol consumption |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Warmth** | 1.00 | | | | | | | | | | |
| **Control** | 0.08 | 1.00 | | | | | | | | | |
| **temperament** | 0.03 | -0.03 | 1.00 | | | | | | | | |
| **IQ** | -0.09 | -0.10 | -0.02 | 1.00 | | | | | | | |
| **Birth weight** | -0.03 | 0.01 | -0.04 | 0.11 | 1.00 | | | | | | |
| **Gestational age** | -0.02 | -0.04 | -0.02 | 0.04 | 0.57 | 1.00 | | | | | |
| **sex** | -0.08 | -0.19 | 0.04 | -0.00 | -0.09 | 0.05 | 1.00 | | | | |
| **Ethnicity** | 0.00 | -0.01 | 0.06 | -0.04 | -0.08 | -0.04 | -0.00 | 1.00 | | | |
| **Maternal age** | -0.18 | 0.02 | -0.02 | 0.21 | 0.08 | -0.01 | -0.02 | -0.05 | 1.00 | | |
| **Maternal smoking** | 0.07 | -0.01 | 0.01 | -0.14 | -0.14 | -0.01 | -0.02 | 0.03 | -0.22 | 1.00 | |
| **Maternal alcohol consumption** | 0.03 | -0.00 | 0.01 | 0.01 | -0.00 | 0.01 | -0.01 | -0.03 | 0.11 | 0.12 | 1.00 |
| **Partner status** | 0.03 | -0.01 | 0.04 | -0.12 | -0.06 | -0.00 | -0.00 | 0.17 | -0.24 | 0.20 | 0.07 |
| **Home ownership** | 0.05 | 0.00 | 0.05 | -0.21 | -0.08 | -0.02 | -0.00 | 0.13 | -0.34 | 0.30 | -0.00 |
| **Household crowding** | 0.01 | 0.02 | 0.04 | -0.14 | -0.02 | 0.01 | -0.01 | 0.08 | -0.15 | 0.15 | 0.00 |
| **Maternal education** | -0.14 | -0.09 | 0.01 | 0.38 | 0.06 | 0.02 | 0.01 | -0.01 | 0.31 | -0.24 | 0.04 |
| **Partner's education** | -0.11 | -0.05 | -0.02 | 0.37 | 0.07 | 0.02 | -0.01 | -0.04 | 0.32 | -0.24 | 0.03 |
| **Parental social class** | 0.09 | 0.05 | 0.02 | -0.27 | 0.05 | -0.00 | 0.00 | 0.03 | -0.29 | 0.20 | -0.03 |
| **Financial difficulties** | 0.06 | -0.01 | 0.08 | -0.13 | -0.04 | 0.00 | -0.00 | 0.07 | -0.08 | 0.15 | 0.01 |
| **Social support** | -0.05 | -0.02 | -0.12 | 0.11 | 0.01 | 0.02 | 0.01 | -0.08 | 0.06 | -0.08 | -0.07 |
| **Maternal depression** | 0.08 | -0.01 | 0.11 | -0.12 | -0.07 | -0.05 | -0.00 | 0.08 | -0.14 | 0.18 | 0.02 |
| **Partner's depression** | 0.05 | -0.01 | 0.06 | -0.09 | -0.04 | -0.04 | -0.00 | 0.10 | -0.10 | 0.13 | 0.04 |
| **Maternal health** | 0.03 | -0.01 | 0.06 | 0.00 | -0.02 | -0.02 | -0.00 | 0.01 | 0.01 | 0.05 | 0.02 |

| | Partner status | Home ownership | Household crowding | Maternal education | Partner's education | Parental social class | Financial difficulties | Social support | Maternal depression | Partner's depression | Maternal health |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Partner status** | 1.00 | | | | | | | | | | |
| **Home ownership** | 0.29 | 1.00 | | | | | | | | | |
| **Household crowding** | 0.12 | 0.30 | 1.00 | | | | | | | | |
| **Maternal education** | -0.15 | -0.26 | -0.16 | 1.00 | | | | | | | |
| **Partner's education** | -0.19 | -0.29 | -0.16 | 0.56 | 1.00 | | | | | | |
| **Parental social class** | 0.19 | 0.31 | 0.18 | -0.45 | -0.44 | 1.00 | | | | | |
| **Financial difficulties** | 0.12 | 0.16 | 0.10 | -0.12 | -0.16 | 0.14 | 1.00 | | | | |
| **Social support** | -0.12 | -0.15 | -0.09 | 0.08 | 0.12 | -0.09 | -0.14 | 1.00 | | | |
| **Maternal depression** | 0.15 | 0.19 | 0.14 | -0.11 | -0.11 | 0.11 | 0.21 | -0.29 | 1.00 | | |
| **Partner's depression** | 0.19 | 0.18 | 0.12 | -0.09 | -0.12 | 0.10 | 0.13 | -0.17 | 0.22 | 1.00 | |
| **Maternal health** | 0.02 | 0.01 | 0.03 | 0.00 | -0.00 | -0.01 | 0.08 | -0.09 | 0.18 | 0.07 | 1.00 |

[a] Correlation coefficients need to be interpreted carefully because all variables are categorical, except IQ, birth weight, gestational age, and maternal age.

**Appendix G Graphical presentation of results on effect-measure modification**

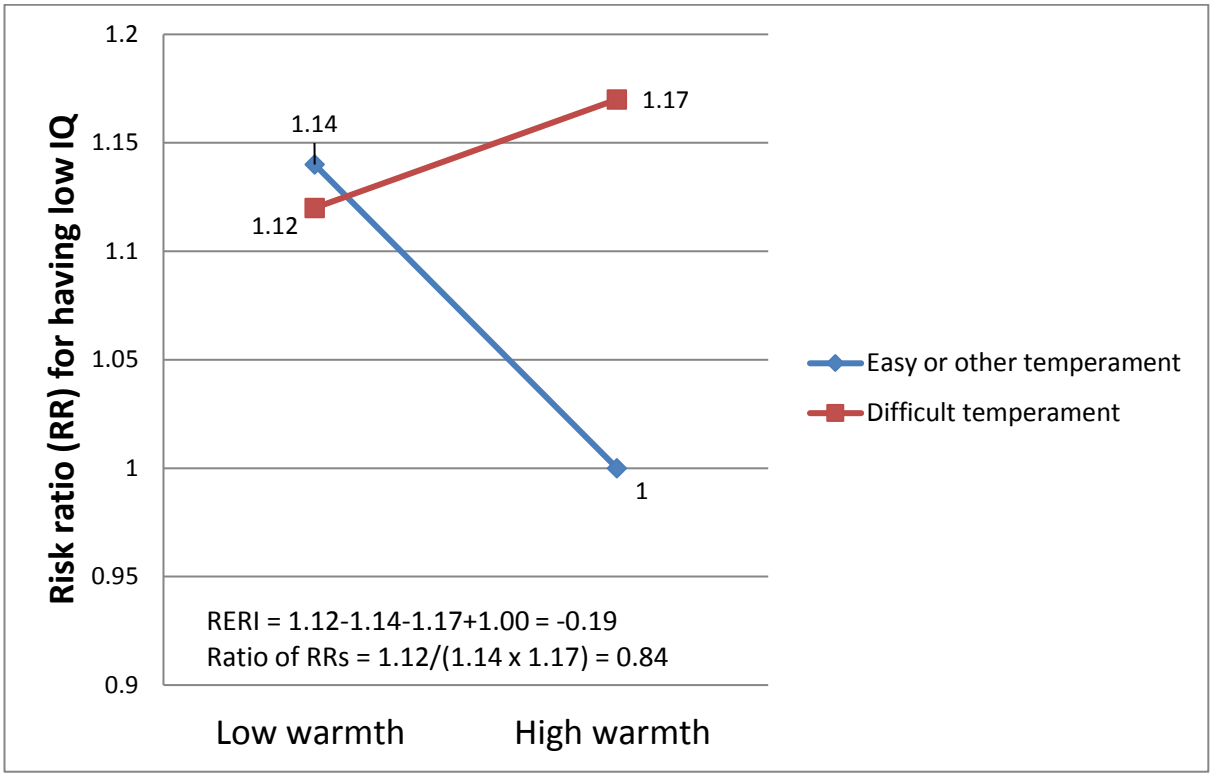


**Figure 6.2 Effect-measure modification of the effect of parenting warmth on IQ by child temperament**

We used the same information from Table 6.4 in the manuscript to present effect-measure modification in a graph (Figure 6.2). Compared with children with high warmth and easy temperament (reference category, RR=1.00), children with low warmth and easy temperament had 14% increased risk of having low IQ; children with high warmth and difficult temperament had 17% increased risk of having low IQ; and children with low warmth and difficult temperament had 12% increased risk of having low IQ.

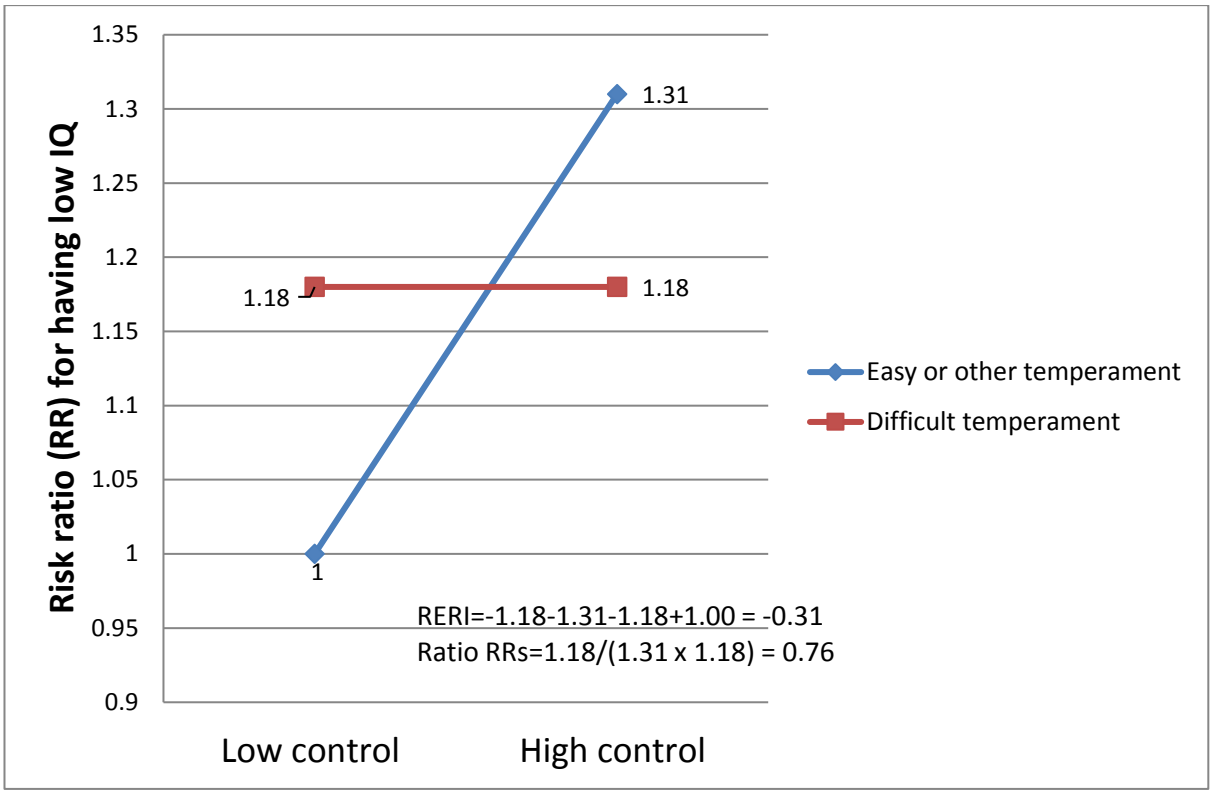


**Figure 6.3 Effect-measure modification of the effect of parenting control on IQ by child temperament**

Estimates in Figure 6.3 were taken from Table 6.5 Effect-measure modification of the effect of parenting control on IQ (< 85) by child temperament (Imputed sample, n=7044) in our manuscript. Compared with children with low control and easy temperament (reference category, RR=1.00), children with high control and easy temperament had 31% increased risk of having low IQ; children with low control and difficult temperament had 18% increased risk of having low IQ; and children with high control and difficult temperament had 18% increased risk of having low IQ.

The RERI was calculated as the difference in estimates between these four groups. The ratio of RRs was calculated as the ratio of estimates between these four groups.

## Appendix H Supporting information tables for effect-measure modification using different IQ cut-offs

### Effect-measure modification of the effect of parenting warmth on IQ by child temperament (Imputed sample, n=7044)

**i) IQ (<80)**

| | High warmth parenting | | Low warmth parenting | | RR (95% CI) for low warmth parenting within strata of temperament type |
|---|---|---|---|---|---|
| | N Low IQ/High IQ | RR (95% CI) | N Low IQ/High IQ | RR (95% CI) | |
| Easy or other temperament | 151/2697 | 1.00 (Ref) | 259/2923 | 1.26 (1.02, 1.57), p=0.032 | 1.23 (0.98, 1.53), p=0.073 |
| Difficult temperament | 28/410 | 1.33 (0.89, 1.99), p=0.158 | 57/519 | 1.43 (1.04, 1.97), p=0.030 | 1.07 (0.65, 1.75), p=0.788 |

RERI= -0.17 (-0.83, 0.50), p=0.619
Ratio of RRs= 0.85 (0.52, 1.39), p=0.514

**ii) IQ (<90)**

| | High warmth parenting | | Low warmth parenting | | RR (95% CI) for low warmth parenting within strata of temperament type |
|---|---|---|---|---|---|
| | N Low IQ/High IQ | RR (95% CI) | N Low IQ/High IQ | RR (95% CI) | |
| Easy or other temperament | 469/2379 | 1.00 (Ref) | 663/2519 | 1.07 (0.95, 1.22), p=0.263 | 1.05 (0.93, 1.19), p=0.473 |
| Difficult temperament | 71/367 | 1.09 (0.85, 1.40), p=0.504 | 133/443 | 1.08 (0.88, 1.33), p=0.465 | 1.00 (0.73, 1.39), p=0.973 |

RERI= -0.08 (-0.43, 0.27), p=0.642
Ratio of RRs= 0.92 (0.67, 1.28), p=0.628
RRs are adjusted for parental control, maternal smoking, alcohol consumption, birth weight, gestation at birth, sex, ethnicity, maternal age, partner status, financial difficulties, maternal and partner's education, parental social class, home ownership, household crowding, maternal health, social support, maternal and partner's depression

**Effect-measure modification of the effect of <u>parenting control</u> on IQ by child temperament (Imputed sample, n = 7044)**

**i) IQ (<80)**

| | Less controlling parenting | | High controlling parenting | | RR (95% CI) for high control parenting within strata of temperament type |
|---|---|---|---|---|---|
| | N Low IQ/High IQ | RR (95% CI) | N Low IQ/High IQ | RR (95% CI) | |
| Easy or other temperament | 194/3350 | 1.00 (Ref) | 216/2270 | 1.34 (1.09, 1.65), p=0.005 | 1.32 (1.06, 1.64), p=0.012 |
| Difficult temperament | 47/581 | 1.31 (0.94, 1.83), p=0.112 | 38/348 | 1.48 (1.01, 2.08), p=0.043 | 1.13 (0.69, 1.84), p=0.623 |

RERI=-0.19 (-0.87, 0.47), p=0.560
Ratio of RRs= 0.83 (0.51, 1.35), p=0.446

**ii) IQ (<90)**

| | Less controlling parenting | | High controlling parenting | | RR (95% CI) for high control parenting within strata of temperament type |
|---|---|---|---|---|---|
| | N Low IQ/High IQ | RR (95% CI) | N Low IQ/High IQ | RR (95% CI) | |
| Easy or other temperament | 580/2964 | 1.00 (Ref) | 552/1934 | 1.21 (1.08, 1.38), p=0.002 | 1.22 (1.08, 1.39), p=0.002 |
| Difficult temperament | 116/512 | 1.12 (0.91, 1.38), p=0.300 | 88/298 | 1.14 (0.90, 1.48), p=0.264 | 1.02 (0.74, 1.41), p=0.883 |

RERI= -0.19 (-0.55, 0.17), p=0.311
Ratio of RRs= 0.84 (0.61, 1.16), p=0.293

RRs are adjusted for parental warmth, maternal smoking, alcohol consumption, birth weight, gestation at birth, sex, ethnicity, maternal age, partner status, financial difficulties, maternal and partner's education, parental social class, home ownership, household crowding, maternal health, social support, maternal and partner's depression

# CHAPTER 7. Child temperament, parental feeding practices and adiposity

## 7.1 Preface

Chapters 5 and 6 of this thesis examined the influence of temperament and parenting on cognitive and academic outcomes in children. Chapter 7 expands the range of outcomes under investigation by focusing on children's adiposity as another important public health outcome and an area of increasing research interest.

This chapter consists of three parts. In Part 1, the influences of temperament at age 0 to 5 years on adiposity at age 7 years were examined using ALSPAC and LSAC data. A child's temperament may influence his adiposity. For instance, children with negative emotion may have higher BMI because mothers are more likely to use food a calm a distressed child.[307]

Part 2 of this chapter is written in manuscript format and is currently under review for publication at a journal. As reviewed in Section 2.5.2, the associations between general parenting (warmth and control) and children's adiposity have been examined in a previous publication by Wake *et al.*[184] using data from LSAC (*n*=3040), and no association was found. Hence this thesis moves beyond general parenting to examine the influence of food-specific parenting (*e.g.* feeding control) on children's adiposity. The second part examines the associations between parental feeding practices (control and use of food to soothe) on adiposity outcomes in childhood and adolescence in ALSPAC data because no feeding practices variables exist in LSAC.

Part 3 examines whether the differential susceptibility theory was supported with adiposity outcomes. The association between parental feeding control and adiposity may be influenced by the child's temperament, in that temperamentally difficult children are more likely to have higher BMI when parents are highly controlling of food intake than temperamentally easy children. This section used an analysis of effect-measure modification to examine whether the effect of parental feeding control observed in Part 2 differs in children with easy and difficult temperaments.

## 7.2 PART 1: Temperament and BMI

### 7.2.1 Introduction

The association between temperament and adiposity has been of interest to temperament research in recent years.[173,308] A 2014 systematic review of 18 papers (5 cross-sectional, 13 longitudinal studies) found some evidence of an impact of temperament on Body Mass Index (BMI) and weight gain in infants and preschool-aged children.[173] However, effect sizes (Cohen's d) of these studies varied markedly from very small (d=0-0.1) to large (d>1.0), possibly due to variation in study design, sample size, measures of temperament, age when outcomes were measured and analytical approach. Part 1 of Chapter 7 examines the associations between temperament (0 to 5 years) and BMI z-score at 7 years in two different cohorts (ALSPAC and LSAC). These two cohorts used different temperament instruments and measured different temperament dimensions at different ages, which may be useful for triangulation of evidence.

### 7.2.2 Methods

***ALSPAC***

In ALSPAC, temperament was reported by mothers when their children were aged 6, 24, and 38 months. Temperament at age 6 months was measured using the adapted Revised Infant Temperament Questionnaire (RITQ) and temperament at 24 months was measured using the Toddler Temperament Scale (TTS). The RITQ and TTS consisted of 9 subscales (activity, rhythmicity, approach, adaptability, intensity, mood, persistence, distractibility, and threshold) which were used to categorise children into easy, intermediate low, intermediate high, and difficult (Chapter 4). Temperament at 38 months was measured using the EAS Temperament Scale, which consisted of four subscales – activity,

emotionality, shyness, and sociability. An average score was created for each dimension with higher scores indicating higher levels of activity, emotionality, shyness, and sociability. BMI was calculated as weight/height$^2$ (kg/cm$^2$). BMI z-scores at age 7 years were calculated based on the 1990 British Growth reference.[309] Confounders of the association between temperament and BMI included ethnicity, financial hardship, partner status, maternal education, social class, household crowding index, home ownership, maternal age, smoking and alcohol during pregnancy, birth weight, gestational age, and maternal postnatal depression. These confounders were measured from 8 weeks gestation to 8 weeks postnatal.

***LSAC***

In LSAC, temperament was reported by the primary caregiver (98.2% mothers) at ages 0 to 1, 2 to 3, and 4 to 5 years. Temperament at age 0 to 1 year was measured using the Short Temperament Scale for Infants (STSI), with 3 subscales - irritability, approach, and cooperation (four items in each subscale). Temperament at ages 2 to 3 years and 4 to 5 years were measured using the Short Temperament Scale for Toddlers (STST), and the Short Temperament Scale for Children, respectively. Both STST and STSC consisted of 3 subscales: reactivity, approach, and persistence (four items in each subscale). BMI z-scores at age 7 years in LSAC sample were created based on the CDC 2000 Growth Charts.[210] Confounders of the association between temperament and BMI z-score included birth weight for gestational age, the Index of Relative Socio-economic Disadvantage (IRSD), Aboriginal and Torres Strait Islander status, financial hardship, language other than English, country of birth, maternal age, highest education, married or living with partner, parental warmth and hostility, gestational diabetes, gestational

hypertension, maternal psychological distress, alcohol and smoking. These confounders were measured at age 0 to 1 year during home interviews.

***Analysis***

Demographic characteristics and BMI z-scores of the ALSPAC and LSAC samples were compared. Multiple linear regressions were used to examine the associations between temperament (0 to 5 years) and BMI z-score at age 7 years in both ALSPAC and LSAC cohorts. Univariable linear regressions were performed to obtain the crude association between temperament and BMI z-score (Model 1), and then adjusted for all confounders listed above (Model 2). Analyses were performed on participants with complete data on BMI z-scores, temperament, and all the covariables.

### 7.2.3 Results

The BMI z-scores at 7 years and characteristics of ALSPAC and LSAC participants are displayed in Table 7.1, showing a greater BMI z-score in the LSAC sample compared with the ALSPAC sample. A higher proportion of mothers in LSAC had a degree level education than in ALSPAC. The mean age of mother in LSAC was older than in ALSPAC while ALSPAC mothers were more likely to smoke during pregnancy than LSAC mothers.

**Table 7.1 Characteristics of ALSPAC and LSAC samples**

| | ALSPAC | LSAC |
|---|---|---|
| | **Mean (SD) or %** | **Mean (SD) or %** |
| BMI z-scores at 7 years | 0.13 (1.04) | 0.38 (1.03) |
| Maternal education, degree | 12.9 | 32.9 |
| Housing tenure, rented or other | 26.6 | 35.8 |
| Maternal age (years) | 28.36 (4.82) | 31.01 (5.51) |
| Child sex, male | 51.7 | 51.1 |
| Birth weight (kg) | 3.39 (0.56) | 3.41 (0.57) |
| Gestational age (weeks) | 39.38 (1.9) | 39.11 (2.03) |
| Smoked during pregnancy | 25.0 | 16.7 |

Table 7.2 shows the association between temperament (as individual dimensions and as categories) and BMI z-score at age 7 years in the ALSPAC sample. There was little evidence of any association between temperament dimensions measured at 6 and 24 months measured using the RITQ or the TTS. When temperament was categorised into easy/difficult, there was some suggestion that children with a difficult temperament at 6 months had a 0.14-point (95% CI -0.22, -0.06) lower BMI z-score at 7 years than children with easy temperament. However, the association between temperament categories and BMI z-score was attenuated at 24 months [β=-0.08, 95% CI (-0.16, 0.00)]. When temperament dimensions were measured using the EAS Temperament Scale at age 38 months, higher emotionality and shyness were associated with lower BMI z-scores, while higher activity and sociability were associated with higher BMI z-score. However, these effect sizes were very small. For instance, one unit increase in emotionality (ranged from 1 to 5) was associated with 0.09 lower BMI z-score.

**Table 7.2 Association between temperament (6, 24 months, and 38 months) and BMI z-score at 7 years in ALSPAC**

| | | Model 1: Unadjusted | | Model 2: Adjusted[a] | |
|---|---|---|---|---|---|
| | Mean (SD) or % | β | 95% CI | β | 95% CI |
| **Temperament 6 months subscales (n=5896)[b]** | | | | | |
| Activity (ranged 1-6) | 4.35 (0.53) | -0.03 | -0.07, 0.02 | -0.04 | -0.09, 0.08 |
| Rhythmicity | 2.60 (0.69) | -0.07 | -0.11, -0.04 | -0.06 | -0.10, -0.03 |
| Approach | 2.53 (0.64) | -0.10 | -0.14, -0.06 | -0.07 | -0.12, -0.04 |
| Adaptability | 2.42 (0.57) | -0.06 | -0.10, -0.01 | -0.04 | -0.09, 0.00 |
| Intensity | 3.51 (0.56) | -0.01 | -0.05, 0.03 | -0.03 | -0.07, 0.02 |
| Mood | 2.78 (0.66) | -0.07 | -0.10, -0.03 | -0.06 | -0.10, -0.02 |
| Persistence | 2.95 (0.73) | -0.04 | -0.07, -0.01 | -0.04 | -0.07, -0.00 |
| Distractibility | 2.42 (0.56) | -0.09 | -0.13, 0.04 | -0.07 | -0.12, -0.02 |
| Threshold | 3.75 (0.61) | -0.00 | -0.05, 0.02 | -0.00 | -0.01, 0.00 |
| **Temperament 6 months easy-difficult construct (n=5896)[c]** | | | | | |
| Easy | 36.94 | Ref | | | |
| Intermediate low | 33.88 | -0.07 | -0.12, -0.01 | -0.05 | -0.11, 0.01 |
| Intermediate high | 14.41 | -0.01 | -0.08, 0.07 | 0.00 | -0.08, 0.08 |
| Difficult | 14.76 | -0.16 | -0.24, -0.09 | -0.14 | -0.22, -0.06 |
| **Temperament 24 months subscales (n=5738)** | | | | | |
| Activity (ranged 1-6) | 3.57 (0.50) | 0.10 | 0.06, 0.15 | 0.08 | 0.03, 0.13 |
| Rhythmicity | 2.54 (0.52) | -0.00 | -0.07, 0.03 | 0.00 | -0.04, 0.05 |
| Approach | 2.83 (0.70) | -0.05 | -0.08, -0.01 | -0.04 | -0.08, -0.01 |
| Adaptability | 2.83 (0.60) | 0.03 | -0.01, 0.07 | 0.02 | -0.00, 0.06 |
| Intensity | 3.38 (0.50) | -0.00 | -0.03, 0.02 | -0.03 | -0.08, 0.02 |
| Mood | 2.51 (0.47) | -0.08 | -0.13, 0.03 | -0.06 | -0.11, 0.00 |
| Persistence | 2.81 (0.54) | -0.04 | -0.09, 0.01 | -0.04 | -0.09, 0.01 |
| Distractibility | 3.45 (0.47) | -0.01 | -0.06, 0.05 | 0.00 | -0.00, 0.01 |
| Threshold | 3.38 (0.55) | -0.06 | -0.10, -0.01 | 0.00 | -0.05, 0.06 |
| **Temperament 24 months easy-difficult construct (n=5738)** | | | | | |
| Easy | 37.48 | Ref | | | |
| Intermediate low | 32.63 | 0.02 | -0.04, 0.08 | 0.02 | -0.04, 0.08 |
| Intermediate high | 14.67 | -0.03 | -0.01, 0.04 | -0.04 | -0.12, 0.04 |
| Difficult | 15.23 | -0.07 | -0.15, -0.01 | -0.08 | -0.16, 0.00 |
| **EAS Temperament Scale 38 months (n=5896)** | | | | | |
| Emotionality (ranged 1-5) | 2.48 (0.85) | -0.09 | -0.12, -0.07 | -0.09 | -0.12, -0.05 |
| Activity | 4.32 (0.63) | 0.08 | 0.04, 0.12 | 0.09 | 0.05, 0.13 |
| Shyness | 2.49 (0.82) | -0.08 | -0.11, -0.05 | -0.07 | -0.10, -0.04 |
| Sociability | 3.64 (0.62) | 0.08 | 0.04, 0.12 | 0.09 | 0.04, 0.13 |

[a] Adjusted for ethnicity, financial hardship, partner status, and maternal education, social class, crowding index, home ownership, maternal age, smoking and alcohol during pregnancy, birth weight, gestational age, and maternal depression at 8 weeks gestation

[b] Scores were summed and divided by the number of items in each dimension. Higher score indicating higher difficulty

[c] Temperament categories were created using Carey's algorithm.

Table 7.3 shows the associations between temperament dimensions measured in LSAC and BMI z-score at age 6 to 7 years. At age 0 to 1 year, there was no evidence of an association between temperament dimensions of irritability, approach, and cooperation on BMI z-score at age 6 to 7 years. This was also the case for temperamental reactivity measured at 2 to 3 and 4 to 5 years. However, the association between approach at 2 to 3 years and persistence at 2 to 3 and 4 to 5 years with higher BMI z-score was slightly increased, albeit, these effect sizes were very small. For instance, one unit increase in persistence (ranged from 1 to 6) at age 2 to 3 years was associated with 0.08 lower BMI z-score at age 6 to 7 years.

**Table 7.3 Association between temperament (0 to 1, 2 to 3 and 4 to 5 years) and BMI z-score at 6 to 7 years in LSAC**

| | | Model 1: Unadjusted | | Model 2: Adjusted[a] | |
|---|---|---|---|---|---|
| | **Mean (SD)** | **β** | **95% CI** | **β** | **95% CI** |
| **Temperament 0-1y (n=3254)[b]** | | | | | |
| Irritability (ranged 1-6) | 2.50 (0.82) | 0.01 | -0.03, 0.05 | 0.03 | -0.01, 0.08 |
| Approach | 4.72 (0.85) | 0.02 | -0.02, 0.06 | 0.01 | -0.03, 0.06 |
| Cooperation | 4.17 (0.89) | 0.03 | -0.01, 0.07 | 0.03 | -0.01, 0.07 |
| **Temperament 2-3y (n=2710)** | | | | | |
| Reactivity (ranged 1-6) | 2.98 (0.96) | 0.01 | -0.03, 0.04 | -0.02 | -0.06, 0.02 |
| Approach | 3.93 (0.97) | 0.04 | 0.01, 0.07 | 0.04 | 0.01, 0.08 |
| Persistence | 4.27 (0.74) | 0.08 | 0.03, 0.12 | 0.08 | 0.02, 0.13 |
| **Temperament 4-5y (n=2920)** | | | | | |
| Reactivity (ranged 1-6) | 2.58 (0.86) | -0.01 | -0.05, 0.03 | -0.02 | -0.07, 0.02 |
| Sociability | 3.82 (1.12) | 0.07 | 0.04, 0.10 | 0.06 | 0.03, 0.10 |
| Persistence | 3.88 (0.89) | 0.02 | -0.02, 0.06 | 0.03 | -0.01, 0.08 |

[a] Adjusted for birth weight for gestational age, IRSD, Aboriginal and Torres Strait Islander status, financial hardship, language other than English, country of birth, maternal age, highest education, married or living with partner, parenting warmth and hostility, gestational diabetes & hypertension, maternal psychological distress, alcohol and smoking (all measured at age 0 to 1 year)
[b] Mean score in each dimension.

### 7.2.4 Discussion

The results of these population-based cohort studies suggested very weak or null associations between temperament and BMI z-scores at 7 years of age. In ALSPAC the associations were strongest when individual dimensions were categorised into the

easy/difficult dichotomy, which summarises multiple aspects of temperament. Nevertheless, the effect sizes from both studies were extremely small and raise concern that the associations could be due to residual or unmeasured confounding. Furthermore, the fact that these associations were of a similar magnitude in two different cohorts using different measures of temperament, and collected at multiple times from ages 0 to 1 through to 4 to 5 years, tends to strengthen the findings that temperament measured by generalised tools has negligible effects on BMI in childhood.

Although different temperament tools are applied in ALSPAC and LSAC, there are similarities across some of the underlying dimensions, making it possible to compare the performance of individual dimensions between these and other studies. Most literature suggests children with negative emotionality/mood/reactivity are more likely to have higher weight status.[173,308] because they were more were more likely to be fed with obesogenic food.[180,310] However this is not supported by the current study with mood (ALSPAC) and irritability/reactivity dimensions (LSAC) not associated with BMI, and the reverse was seen for negative emotionality (LSAC) which was associated with lower BMI. While it is possible that children who are temperamentally difficult or have negative emotionality may have problems with emotional eating, other studies also suggest that children with negative emotionality were more likely to have fussy and picky eating behaviours and refuse to try new foods.[311] In addition, the lower BMI in temperamentally difficult children may be because parents have greater difficulties feeding children who are temperamentally difficult.

While little is known about the approach/sociability dimensions of temperament and BMI, others studies have reported associations between persistence and BMI. The positive association between persistence (LSAC) and BMI z-scores is contradicted by findings from Faith and Hittner[312] who found a 54% reduced odds of being overweight in 6-year-old boys who have higher persistence level. However, it is difficult to explain why there was no association between persistence and BMI in the ALSPAC study. Persistence is sometimes conceptualised as a component of self-regulation and inhibitory control. A number of studies have suggested that poor self-regulation (inhibitory control) is associated with higher body weight because children with poor self-regulation are less able to control their intake of palatable foods.[179,313] This could not be examined here as inhibitory control was not measured in both ALSPAC and LSAC, but the literature suggests parents may use feeding behaviours such as restriction on children with poor self-regulation.

In summary, evidence from ALSPAC and LSAC suggested that the association between temperament and BMI is, at best, very small. Contrary to other literature, there is no evidence from both ALSPAC and LSAC that children who are temperamentally difficult or have higher levels of negative mood and reactivity have higher BMI z-scores. The association between temperament dimensions of approach/sociability and persistence and BMI z-score are less consistent between ALSPAC and LSAC and these disparities could be exploited by comparing the content of the temperament tools to improve our understanding of which aspects of temperament are contributing to such differences. However, this research is beyond the scope of this doctoral project and the small associations tend to suggest that this is not likely to be of major importance to public

health researchers. It is for this reason, the attention shifts to parenting and adiposity in the next section.

## 7.3 PART 2 (Publication): Associations of parental feeding control of food intake and use of food to soothe with adiposity in childhood and adolescence

### 7.3.1 Statement of authorship

| | |
|---|---|
| Title of Paper | Associations of parental feeding control of food intake and use of food to soothe with adiposity in childhood and adolescence |
| Publication Status | ☐ Published<br>☐ Accepted for Publication<br>☑ Submitted for Publication<br>☐ Unpublished and Unsubmitted w ork w ritten in manuscript style |
| Publication Details | This manuscript is currently under review. |

### Principal Author

| | | | |
|---|---|---|---|
| Name of Principal Author (Candidate) | Shiau Yun Chong | | |
| Contribution to the Paper | Conceived and designed the study, analysed and interpreted the data, drafted the manuscript. | | |
| Overall percentage (%) | 80 | | |
| Certification: | This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper. | | |
| Signature | | Date | 15-08-2016 |

### Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

x. the candidate's stated contribution to the publication is accurate (as detailed above);

xi. permission is granted for the candidate in include the publication in the thesis; and

xii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

| Name of Co-Author | Catherine Chittleborough | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | Tess Gregory | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | John Lynch | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | Murthy Mittinty | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| Name of Co-Author | Lisa Smithers | | |
|---|---|---|---|
| Contribution to the Paper | Contributed to the design of the study, data acquisition, interpretation of the results, and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

## What is already known about this subject?

- Parental feeding practices may contribute to children's adiposity.
- Evidence on this association is mixed and most often involves small, cross-sectional studies.

## What this study adds?

- High parental feeding control at 3.5 years is associated with lower BMI z-score and fat mass at 15 years, but not BMI z-score at 7 years.
- The use of food to soothe at 3.5 years is not associated with children's BMI z-scores or fat mass at 7 and 15 years.

### 7.3.2 Abstract

**Background:** Associations of parental feeding techniques with adiposity are mixed and largely rely on cross-sectional studies. We examined associations between parental feeding control and using food to soothe at 3.5 years on adiposity at 7 and 15 years.

**Methods:** Participants were from the Avon Longitudinal Study of Parents and Children (n=7674). Feeding control was assessed using the item 'how much choice do you allow him/her in deciding what foods he eats at meals?'. Use of food to soothe was reported by mothers on the item 'how often do you use sweets or other foods to stop his/her crying or fussing?'. BMI at 7 and 15 years was converted to sex- and age-adjusted z-scores. Fat mass was assessed at 15 years using dual energy X-ray absorptiometry.

**Results:** In fully-adjusted models, children given the least choice had 0.12 point lower BMI z-score and 1.54 kg lower fat mass at 15 years than children with the most choices. There was no evidence of an association between using food to soothe and children's adiposity.

**Conclusions:** Contrary to some studies, higher parental control over food choice was associated with lower adiposity, but use of food to soothe was not associated with adiposity at ages 7 and 15.

### 7.3.3 Background

Overweight and obesity is a global public health challenge with 42 million children under 5 years overweight or obese.[22] Due to the high prevalence of overweight and obesity, there is interest in examining how early life feeding experiences influence children's later food preferences, diets and adiposity.[314-317]

Parental feeding practices that may influence children's adiposity include the use of feeding control and food to soothe.[316] There is a lack of a universal definition of feeding control.[105] Most studies conceptualize feeding control as pressuring children to eat certain foods or deliberately restricting children's intake of energy-dense food. Systematic reviews have shown that higher parental pressure to eat is associated with lower children's body mass index (BMI).[318,319] It has been suggested that parental use of pressure increases children's dislike for foods and likelihood of picky eating. Findings on the association between feeding control and children's eating and adiposity are inconsistent. Birch and Fisher reported that restrictive feeding is associated with more snacking and adiposity in 5-year-old girls.[101,320] The authors suggested that restrictive feeding increases children's preference for the restricted foods and children are more likely to eat in the absence of hunger.[100,321] Several cross-sectional studies found no association or a positive association between restrictive feeding and adiposity outcomes particularly in toddlers and preschoolers,[318,319] while two longitudinal studies reported a negative association.[106,322]

Food to soothe is conceptualized as using food in response to the child's crying or fussing due to reasons other than hunger,[103,112,190,323] for instance, when a child is upset or has

temper tantrums.[103,310] Most studies that examined food to soothe and child's eating and adiposity are cross-sectional,[103,112] with findings varying from no association[111,188,190] to positive associations.[102,103,324]

A review of 32 articles on parental feeding practice on children's eating behaviours and weight found that 22 studies were cross-sectional and therefore unable to establish temporal order between parenting and adiposity.[316] For instance, it is unclear whether parental restriction influences children's adiposity, or parents who perceived their children as overweight are more controlling.[322] Longitudinal studies with short-term follow up may be difficult to detect the effects, especially when sample sizes are small. The current study used data from a large, population-based prospective study to investigate the associations between parental feeding control and food to soothe at 3.5 years on children's BMI z-scores at 7 and 15 years and fat mass at age 15 years. Our primary analyses used BMI as an outcome at ages 7 and 15 to examine consistency of associations between feeding practices and adiposity in childhood and adolescence, and fat mass at age 15 when adiposity is strongly predictive of adult outcomes.[24] In secondary analyses, we explored the maternal and child characteristics that may influence parental use of feeding control and food to soothe. For instance, child temperament may influence maternal feeding and adiposity.[173,308].[173,308]

### 7.3.4 Methods

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a population-based, prospective study investigating the determinants of children's health and development. A total of 14541 pregnant women resided in the Southwest of England with delivery date

between 1st April 1991 and 31st December 1992 were recruited into the study. The core eligible ALSPAC sample consists of 13978 infants who were alive at 1 year of age (Figure 7.1). The cohort is considered broadly representative of the population living in Avon at the time although ethnic minorities, single parents and unmarried couples were slightly underrepresented.[197] The study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/). Ethical approvals were obtained from the ALSPAC Law and Ethics committee and local Research Ethics committees. Written informed consent was obtained from all participants.

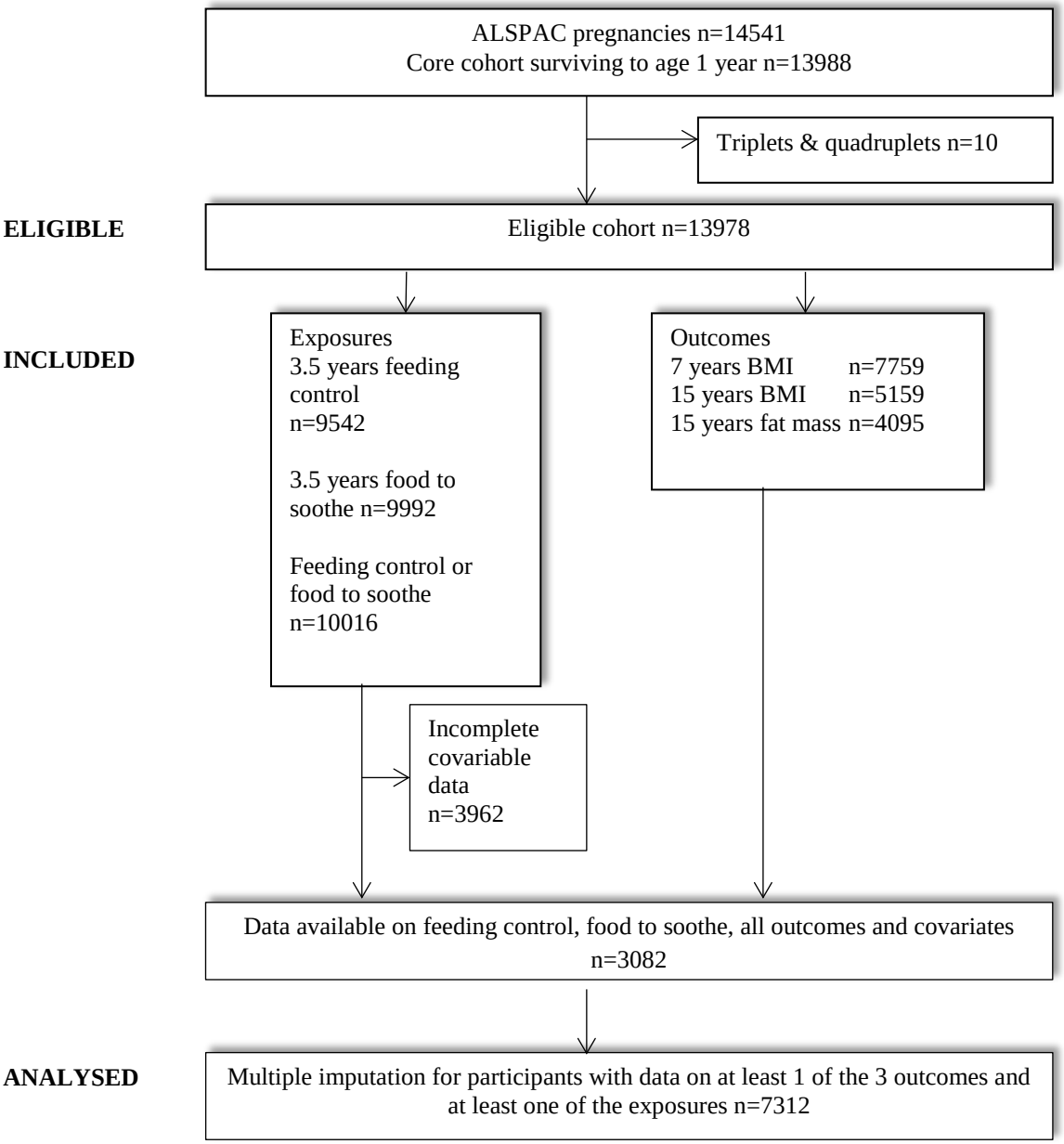


**Figure 7.1 Flow chart of participants**

***BMI and fat mass***

At age 7 and 15 years, participants were invited to attend the annual assessment clinics where anthropometric measurements were taken by trained researchers. Height was measured without shoes to the nearest 0.1 cm using the Harpenden Stadiometer (Holtain Croswell, Dyfed, UK). Weight was measured in light clothing (Tanita, Arlington Heights, IL) to an accuracy of 0.1 kg. BMI was calculated as weight (kg)/ height ($m^2$) and converted to

z-scores based on the British 1990 growth reference as these are temporally and culturally appropriate to the ALSPAC cohort.[309] Total fat mass was assessed at a mean age of 15.5 years using dual energy X-ray absorptiometry (DXA) (Prodigy scanner, Lunar Radiation Corp, Madison, Wisconsin, US).

***Parental feeding control and food to soothe***

Parental feeding control was measured at age 3.5 years using an item 'how much choice do you allow him/her in deciding what foods he eats at meals?'. Mothers responded to this item as 'he/she can choose from any food available', 'he/she is given a choice from a few alternatives that I select', and 'I decide what he/she will eat'. Mothers who decided what the child will eat were considered as endorsing higher control than mothers who allowed their child to have food choices.[325] Mothers' use of food to soothe at age 3.5 years was measured using one item 'how often do you use sweets or other foods to stop his/her crying or fussing?'. Response to this item was 'frequently' (once a day or many times a week), 'infrequently' and 'never'.

***Confounding***

Confounding of the association between parental feeding control, use of food to soothe and adiposity were decided *a priori* based on current literature. Confounding factors included maternal education, home ownership, household crowding, parental social class, and financial difficulties), child's dietary patterns at 3 years of age, and child and parent factors (child temperament, birth weight z-score, ethnicity, eating difficulties, maternal age, maternal pre-pregnancy BMI, sole parenting, postnatal depression, smoking, alcohol

intake during pregnancy, and number of other children). Full details of how these confounding factors were measured are provided in supplementary material.

***Analysis***

Multiple linear regression models were used to estimate the associations between feeding control, use of food to soothe and BMI z-scores at age 7 and 15 years and fat mass at 15 years. First, we estimated the unadjusted effect of feeding control and use of food to soothe on outcomes separately (Model 1). We then adjusted for children's temperament (Model 2) and dietary patterns (Model 3). In Model 4, the feeding control model was adjusted for all confounding factors and parental use of food to soothe (and age, sex, and height for fat mass); the food to soothe model was adjusted for all confounding factors and feeding control (and age, sex, and height for fat mass). All analyses were performed using STATA version 13.0 (Stata Corp, College Station, Texas).

***Multiple imputation***

Twenty imputed datasets were generated using the 'mi impute chained' command in STATA under the missing at random assumption.[235] Multiple imputation was performed for participants with data available on at least one of the three outcomes and at least one of the exposures (n=7312). The imputation model included all the exposures, outcomes, confounding variables, interactions between feeding control, use of food to soothe and temperament, and auxiliary variables (energy intake at age 13, temperament at age 2 (easy/difficult), and time spent watching television). We performed analyses on the complete case sample and the imputed sample and the results were similar (data not shown). Results from the imputed sample are reported.

### 7.3.5 Results

Table 7.7 in Appendix J shows the characteristics of the response, complete case (n=3082) and imputed sample (n=7312). The complete case sample was more socio-economically advantaged and had lower BMI z-scores and fat mass than the response sample. The imputed sample was more comparable to the response sample in terms of the socio-economic position, BMI z-scores and fat mass.

Table 7.4 shows the maternal and child characteristics of ALSPAC participants according to their responses to the parental feeding control and use of food to soothe items. When asked about the extent the child was allowed to have food choices, 9% of the participants responded that the child was allowed to choose from any available foods, 61% responded that the child was allowed to have some food choices, and 30% responded that the parent decided what the child will eat. Mothers who allowed their child to choose from any foods available were more socio-economically disadvantaged, had higher BMI and were more likely to smoke during pregnancy than mothers who allowed their child to have a few choices and mothers who made food decisions for their child. Children who were allowed to choose from any available foods were more likely to have 'junk/processed' dietary pattern, more likely to be choosy about food, and had higher BMI z-scores and fat mass than children who were given a few alternatives and children whose food intake was decided by their parents. The proportion of participants who frequently, infrequently, and never used food to soothe a child's crying or fussing was 12%, 51% and 37% respectively. Mothers who used food to soothe frequently were more socio-economically disadvantaged, had higher BMI and were more likely to smoke during pregnancy than mothers who never or infrequently used food to soothe. Children who

were temperamentally difficult or choosy about food were more likely to have parents who used food to soothe.

**Table 7.4 Maternal and child characteristics of those participants according to their responses to the feeding control and use of food to soothe items at 3.5 years (n=7312)**

| | Parental feeding control | | | Parental use of food to soothe | | |
|---|---|---|---|---|---|---|
| | **Can choose any (8.9%)** | **Choice of few (61.2%)** | **Parent decides (29.9%)** | **Frequent (11.4%)** | **Infrequent (51.2%)** | **Never (37.4%)** |
| ***Maternal characteristics*** | | | | | | |
| Maternal education (%) | | | | | | |
| CSE/vocational | 34.2 | 18.9 | 23.6 | 28.6 | 20.4 | 21.4 |
| O level | 35.2 | 35.5 | 36.3 | 36.6 | 36.5 | 34.3 |
| A level | 19.9 | 28.2 | 25.2 | 23.0 | 27.3 | 26.6 |
| Degree | 10.8 | 17.4 | 15.0 | 11.8 | 15.8 | 17.8 |
| Parental highest social class (%) | | | | | | |
| I/II (highest) | 48.4 | 62.6 | 58.6 | 55.3 | 60.6 | 61.0 |
| III | 45.5 | 33.6 | 36.3 | 39.4 | 35.1 | 34.8 |
| IV/V (lowest) | 5.1 | 3.8 | 5.1 | 5.4 | 4.3 | 4.2 |
| Household crowding (%) | 6.2 | 3.7 | 4.0 | 5.1 | 3.7 | 4.1 |
| Home ownership, rented/other (%) | 25.2 | 15.1 | 17.1 | 17.9 | 15.8 | 17.3 |
| Had financial difficulty (%) | 11.9 | 7.6 | 8.0 | 10.2 | 8.1 | 7.6 |
| Maternal BMI, mean (SD) | 23.49 (4.36) | 22.92 (3.67) | 22.76 (3.64) | 23.19 (3.95) | 22.91 (3.64) | 22.84 (3.79) |
| Maternal age, mean (SD) | 28.59 (5.02) | 29.16 (4.52) | 29.32 (4.31) | 28.57 (4.35) | 29.19 (4.46) | 29.29 (4.62) |
| Smoking during pregnancy (%) | 23.4 | 18.0 | 17.5 | 20.7 | 18.5 | 17.2 |
| ***Child characteristics*** | | | | | | |
| Diet patterns at 3 years, mean (SD) | | | | | | |
| Junk/Processed | 0.25 (1.08) | -0.12 (0.92) | -0.16 (0.95) | 0.36 (1.06) | -0.08 (0.91) | -0.27 (0.91) |
| Health conscious | -0.10 (0.97) | 0.05 (1.00) | -0.01 (1.00) | -0.11 (0.94) | -0.01 (0.98) | 0.10 (1.04) |
| Traditional | -0.20 (1.04) | -0.03 (1.00) | 0.10 (0.94) | -0.16 (0.94) | -0.02 (0.97) | 0.06 (1.03) |
| Snacks | -0.08 (1.10) | 0.10 (0.97) | 0.02 (0.96) | 0.16 (1.03) | 0.11 (0.97) | -0.04 (0.97) |
| Difficult temperament at 6 months (%) | 15.0 | 15.4 | 13.3 | 18.8 | 15.3 | 12.7 |
| Choosy about food at 15 months (%) | 72.7 | 70.7 | 64.5 | 75.9 | 71.1 | 63.7 |
| Overeaten at 15 months (%) | 16.1 | 15.6 | 15.8 | 18.0 | 17.1 | 13.2 |
| Birth weight z-score, mean (SD) | 0.04 (1.00) | 0.11 (0.99) | 0.18 (0.97) | 0.15 (1.00) | 0.12 (0.99) | 0.12 (0.97) |
| BMI z-score at 7 years, mean (SD) | 0.21 (1.13) | 0.12 (1.03) | 0.10 (1.00) | 0.15 (1.05) | 0.12 (1.02) | 0.12 (1.05) |

| | Parental feeding control | | | Parental use of food to soothe | | |
|---|---|---|---|---|---|---|
| | **Can choose any (8.9%)** | **Choice of few (61.2%)** | **Parent decides (29.9%)** | **Frequent (11.4%)** | **Infrequent (51.2%)** | **Never (37.4%)** |
| BMI z-score at 15 years, mean (SD) | 0.55 (1.16) | 0.37 (1.09) | 0.36 (1.05) | 0.41 (1.13) | 0.39 (1.07) | 0.36 (1.11) |
| Fat mass (kg) at 15 years, mean (SD) | 16.71 (10.59) | 15.43 (9.14) | 14.65 (8.89) | 15.92 (9.69) | 15.35 (9.13) | 15.07 (9.20) |

[a] CSE= Certificate of Secondary Education. CSE, O, and A levels were high school qualifications. O-levels were usually studied at age 16 and A-levels at age 18.

[b] 'Junk/Processed' pattern was described by foods with high fat and sugar content, and processed and convenience foods. 'Health conscious' pattern was associated with consumption of salads, fruit, vegetables, fish, pasta and rice. 'Traditional' pattern was highly associated with meat, poultry, potato and vegetable consumption. 'Snack' pattern was described as diet consisting primarily of snack and finger foods as opposed to foods where cooking is required.[326]

Table 7.5 shows the associations between the parental feeding control item at 3.5 years and children's adiposity outcomes at 7 and 15 years. In the unadjusted model, there was evidence of lower BMI z-scores and fat mass among children whose parents used higher feeding control. However, effects were attenuated after adjusting for both temperament and dietary patterns at 3 years. In the fully-adjusted model, there was weak evidence of an association of parental control with BMI z-score at age 7 years (β=-0.07, 95% CI -0.16, 0.01). At 15 years, children who were given the least choices (parent decided food intake) had 0.12 (95% CI -0.23, -0.02) lower BMI z-score and 1.54 kg (95% CI -2.35, -0.74) lower fat mass compared to children who were given free choices. Regression models in Table 7.6 showed no evidence of parental use of food to soothe at age 3.5 years and children's adiposity outcomes at either age.

**Table 7.5 Associations between parental feeding control (3.5 years) on BMI z-score at 7 and 15 years, and fat mass at 15 years in imputed sample (n=7312)**

| | Model 1[a] | | Model 2[b] | | Model 3[c] | | Model 4[d] | |
|---|---|---|---|---|---|---|---|---|
| | β | 95% CI | β | 95% CI | β | 95% CI | β | 95% CI |
| **BMI z-score at 7 years** | | | | | | | | |
| **Feeding control** | | | | | | | | |
| Can choose any | Ref | | Ref | | Ref | | Ref | |
| Choice of few | -0.09 | (-0.18, -0.01) | -0.09 | (-0.18, -0.01) | -0.08 | (-0.16, 0.01) | -0.07 | (-0.16, 0.01) |
| Parent decides | -0.10 | (-0.20, -0.01) | -0.11 | (-0.20, -0.02) | -0.10 | (-0.19, -0.00) | -0.08 | (-0.17, 0.01) |
| **BMI z-score at 15 years** | | | | | | | | |
| **Feeding control** | | | | | | | | |
| Can choose any | Ref | | Ref | | Ref | | Ref | |
| Choice of few | -0.18 | (-0.28, -0.08) | -0.18 | (-0.28, -0.08) | -0.14 | (-0.24, -0.04) | -0.12 | (-0.22, -0.01) |
| Parent decides | -0.19 | (-0.30, -0.08) | -0.19 | (-0.30, -0.08) | -0.16 | (-0.27, -0.05) | -0.12 | (-0.23, -0.02) |
| **Fat mass, 15 years (kg)** | | | | | | | | |
| **Feeding control** | | | | | | | | |
| Can choose any | Ref | | Ref | | Ref | | Ref | |
| Choice of few | -1.28 | (-2.13, -0.42) | -1.28 | (-2.13, -0.42) | -0.91 | (-1.78, -0.05) | -1.12 | (-1.89, -0.36) |
| Parent decides | -2.06 | (-2.95, -1.16) | -2.05 | (-2.95, -1.16) | -1.76 | (-2.67, -0.85) | -1.54 | (-2.35, -0.74) |

[a] Model 1: unadjusted
[b] Model 2: adjusted for infant temperament at 6 months
[c] Model 3: Model 2+ dietary patterns at 3 years
[d] Model 4: Model 3+ parental use of food to soothe, maternal education, sole parenting, financial difficulties, social class, home ownership, maternal age, birth weight z-score, alcohol, smoking, maternal BMI, crowding index, maternal depression, number of other children, ethnicity, child not eaten enough, child refuse right food, child choosy with food, child overeaten, and difficult to establish eating routine and age, sex and height (for model with fat mass outcome)

**Table 7.6 Associations between parental use of food to soothe (3.5 years) on BMI z-score at 7 and 15 years, and fat mass at 15 years in imputed sample (n=7312)**

| | Model 1[a] | | Model 2[b] | | Model 3[c] | | Model 4[d] | |
|---|---|---|---|---|---|---|---|---|
| | β | 95% CI | β | 95% CI | β | 95% CI | β | 95% CI |
| **BMI z-score at 7 years** | | | | | | | | |
| **Use of food to soothe** | | | | | | | | |
| Frequent | Ref | | Ref | | Ref | | Ref | |
| Infrequent | -0.03 | (-0.12, 0.04) | -0.04 | (-0.12, 0.04) | -0.03 | (-0.11, 0.05) | -0.03 | (-0.11, 0.05) |
| Never | -0.04 | (-0.12, 0.04) | -0.05 | (-0.13, 0.03) | -0.03 | (-0.11, 0.06) | -0.04 | (-0.12, 0.04) |
| **BMI z-score at 15 years** | | | | | | | | |
| **Use of food to soothe** | | | | | | | | |
| Frequent | Ref | | Ref | | Ref | | Ref | |
| Infrequent | -0.02 | (-0.11, 0.08) | -0.02 | (-0.10, 0.08) | 0.03 | (-0.06, 0.12) | 0.03 | (-0.05, 0.12) |
| Never | -0.05 | (-0.15, 0.04) | -0.05 | (-0.15, 0.04) | 0.01 | (-0.08, 0.11) | 0.01 | (-0.09, 0.10) |
| **Fat mass, 15 years (kg)** | | | | | | | | |
| **Use of food to soothe** | | | | | | | | |
| Frequent | Ref | | Ref | | Ref | | Ref | |
| Infrequent | -0.57 | (-1.34, 0.21) | -0.56 | (-1.34, 0.22) | -0.18 | (-0.96, 0.60) | -0.03 | (-0.70, 0.65) |
| Never | -0.85 | (-1.65, 0.05) | -0.83 | (-1.63, 0.03) | -0.33 | (-1.15, 0.48) | -0.05 | (-0.66, 0.76) |

[a] Model 1: unadjusted
[b] Model 2: adjusted for infant temperament at 6 months
[c] Model 3: Model 2+ dietary patterns at 38 months
[d] Model 4: Model 3+ parental use of food to soothe, maternal education, sole parenting, financial difficulties, social class, home ownership, maternal age, birth weight z-score, alcohol, smoking, maternal BMI, crowding index, maternal depression, number of other children, ethnicity, child not eaten enough, child refuse right food, child choosy with food, child overeaten, and difficult to establish eating routine and age, sex and height (for model with fat mass outcome)

### 7.3.6 Discussion

Using data from a population-based sample of UK infants, this study found a weak association that higher parental feeding control, in terms of restricting children's food choices, at age 3.5 years was associated with 0.07 lower BMI z-score at 7 years. The association of parental control and BMI was stronger at 15 years old with 0.12 lower BMI z-scores and 1.54 kg less body fat mass in children with the least food choice compared to children with the most food choice. A reduction in BMI z-score of 0.07 corresponds with a difference of 0.22 kg for a 7-year-old boys of average height (1.27 m) while a reduction in BMI z-score of 0.12 corresponds with a difference of 1.09 kg weight for a 15-year-old male youth of average height (1.65 m), which is slightly lower than the fat mass estimated at 15 years. Findings from the current study are in contrast to a US longitudinal study in year 2004 by Faith *et al*.[327] who found that restriction at age 5 was associated with a higher BMI z-score at age 7 years in high-risk children (maternal BMI≥66$^{th}$ percentile) (n=57). On the other hand, Spruijt-Metz *et al.*[187] reported no association between restriction at 11 years of age and fat mass at 14 years of age in a small sample of African American adolescents (n=47). Only two studies have reported a negative association between parental feeding control (food restriction) and BMI. Campbell *et al*.[106] examined food restriction in younger (5 to 6 years) and older (10 to 12 years) children and their BMI z-scores 3 years later and found that higher restrictive feeding was associated with lower BMI-scores in the younger but not older children (n=392). Farrow and Blisett[322] reported that parental restriction at age 1 year was associated with lower BMI at age 2 years (n=62).

It is not clear why parental feeding control had a negative association with BMI and fat mass. It may be that the parental feeding control item measured in the ALSPAC was more about the child's choice of foods and it is different from other questionnaires that measure mainly parental over controlling practices.[190,328] A cross-sectional study that measured feeding control using an item 'how much choice is your child allowed in deciding what foods he/she eats at breakfast and lunch?', which is similar to the feeding control item in this current study, also reported that child BMI z-scores were the lowest among children who were given no food choice than in children given little or some choice (n=1790).[329] Parents who allowed their children to have more choices may be less concerned about the child eating and dietary intakes, or had a more globally indulgent parenting style which has been associated with more energy-dense food intake and higher weight status in children.[314,315,330] Parents who considered themselves as more responsible in feeding were less likely to allow their child to have control over their own food intake.[331]

As opposed to Campbell *et al.*[106] who reported a reduction in BMI z-score of 0.013 of higher feeding control in children aged 5 to 6 years but no effect in children aged 10 to11 years, the current study found a greater reduction in BMI in older children (15 years) than in younger children (7 years). The weaker evidence on BMI z-score at age 7 years may be because dietary intake of younger children depends more on the food availability at home whereas older children may have more opportunities to eat out and have greater access to unhealthy foods, suggesting that higher parental use of control over food choice among older children might encourage a healthier diet and could be beneficial for their

weight. Findings from this current study suggested that future studies may need to follow children longer to discover the effect of feeding control on adiposity.

The current study found no evidence of parent use of food to soothe on adiposity outcomes at either age. Most studies that reported a positive association between using food to soothe and BMI involved very young children, *i.e.* 0 to 2 years.[102,103,324] We measured food to soothe at 3.5 years and findings are consistent with other cross-sectional studies that measured food to soothe at a similar or older ages.[111,188] Carnell and Wardle[188] measured food to soothe in 3 to 5 year old children and reported no association between food to soothe and BMI z-score. Rodenburg *et al.*[111] found that using food to soothe at age of 9 was associated with more energy-dense snacking but not with BMI z-scores.

Secondary analyses of the current study showed that mothers who allowed their child to have free choice were more socially disadvantaged and had higher BMI than mothers who decided what their child eat. Previous studies have also reported similar findings,[110,323,332] *e.g.* Ogden *et al.*[110] found that parents from higher social classes were more likely to use overt control (*i.e.* telling the child what, when or how much to eat). Consistent with other studies,[173,310,333] we found that mothers who used food to soothe frequently were more likely to consider their child as temperamentally difficult. These findings reinforce the need for taking into account the parents perception of their child's temperament to better understand the associations between parental feeding practices and child eating behaviours.

There have been a few child obesity interventions that target parenting feeding practices such as controlling feeding (*e.g*. the NOURISH, Prevention of Overweight in Infant).[192,334] These interventions showed small changes on parental feeding practices and children's diet but no reduction on BMI at ages 2 to 4 years.[192,334] Given that there is no a standard guideline that is specific on feeding control, it remains unknown how positive feeding practices may "protect" children from overweight. Greater understanding on the contexts and specific feeding behaviours are needed to design obesity prevention interventions, and observational data such as ours showing that parental control might influence adiposity can help inform such interventions.

The limitations of the current study include the use of one-item measure of feeding control which assessed the extent that the child was allowed to have food choices, but it was unclear how much, when and which choices were given to the child. However, the feeding control item used in this study reflected the extent to which parents allowed their child to have food choices which is an important indicator of restriction or authoritarian parenting style.[329,332] No universal definition of feeding control is available and no single measure is accepted as a gold standard.[104] The strengths of the current study include the use of a large, population-based study and adjustment for a wide range of confounders to provide a robust estimate of the association between feeding control, use of food to soothe and children's adiposity. In addition, we strengthened the evidence for associations with BMI by using data on fat mass as a more direct measure of adiposity.[335]

In conclusion, higher parental feeding control at 3.5 years was associated with 0.07 lower BMI z-score at 7 years and 0.12 lower BMI z-score and 1.54 kg less fat mass at 15 years.

There was little evidence that food to soothe at 3.5 years was associated with adiposity in childhood or adolescence.

### 7.3.7 Online appendices

**Appendix I Details of measurements of covariables included in analyses**

***Socio-economic position***

Information on maternal education, home ownership, household crowding, parental social class, and financial difficulties were collected by postal questionnaires sent to the mothers between 8 and 32 weeks of gestation. Maternal education was reported as the highest completed level of the mothers and categorised into four levels: none/certificate of secondary education (CSE)/ vocational training, O level, A level, and degree. CSE, O, and A levels were high school qualifications. O-levels were usually studied at age 16 and A-levels at age 18. Home ownership was categorised into owned or mortgaged and rented or other. Household crowding was categorised according to whether there were ≤1 or >1 person per room. Parental social class was categorised based on the highest occupation using standard United Kingdom classification of occupation, ranging from class I (highest) to class V (lowest).[336] Financial difficulties were measured using 5 items of how difficult mothers found it to afford food, clothing, heating, rent or mortgage, and things they will need for their babies. Each response ranged from 1 (very difficult) to 4 (not difficult). The sum of the scores of the five items was subtracted from 20 to derive a total financial difficulties score. Mothers with a score of 9 and above were defined as experiencing financial difficulties.

***Dietary factors***

Child's dietary data was measured at age 3 using Food Frequency Questionnaire (FFQ). Mothers reported the child's consumption of each foods and beverages item as 'never or rarely', 'one in 2 weeks', 'one to three times per week', 'four to seven times per week' or 'more than once a day'. Using principal component analysis, four dietary patterns were

identified: junk/processed, traditional, health conscious, and snacks. The methods have been described in details elsewhere.[326]

***Child and parent factors***

Child factors included temperament, birth weight z-score, ethnicity, and eating difficulties. Child temperament was measured at 6 months using the Revised Infant Temperament Questionnaire[42] and categorised into easy or difficult based on the ALSPAC-specific norms.[244] Birth weight was collected by ALSPAC staff from obstetric data and standardized based on the 1990 British growth chart stratified by sex and gestational age.[337] Child eating difficulties at age 24 months were reported by mothers (yes/no) on the following items: refused to eat the right food, been choosy with food, over-eaten, and difficult to establish an eating routine.

Parental factors included maternal age, maternal pre-pregnancy BMI, sole parenting, postnatal depression, smoking (yes/no) and alcohol intake during pregnancy, and number of other children. Maternal pre-pregnancy weight and height were self-reported at 12 weeks gestation and BMI was calculated as weight (kg)/height(m)$^2$. Sole parenting was categorised as: 'married or living with partner' and 'not married or not living with partner'. Maternal depression was indicated by scores ≥13 from the Edinburgh Postnatal Depression Scale measured at 8 weeks postpartum.[297] Mothers' alcohol intake during the first 3 months of pregnancy was categorised as: never, ≤ 1 glass per week, ≥ 1 glasses per week, and ≥ 1 glasses per day. Number of other children living in the family was measured at 6 months of age and categorised as: none, 1, 2, 3, and ≥4.

## Appendix J Characteristics of ALSPAC study participants

### Table 7.7 Characteristics of ALSPAC study participants

| | Response sample[a] | | Complete case sample (n=3082)[b] | Imputed sample (n=7312)[c] |
|---|---|---|---|---|
| | N | Mean (SD) or % | Mean (SD) or % | Mean (SD) or % |
| ***Outcomes*** | | | | |
| BMI z-score 7 years | 7759 | 0.13 (1.04) | 0.10 (1.02) | 0.12 (1.03) |
| BMI z-score 15 years | 5159 | 0.36 (1.09) | 0.34 (1.07) | 0.38 (1.09) |
| Fat mass 15 years (kg) | 4905 | 15.32 (9.17) | 15.17 (9.03) | 15.31 (9.22) |
| ***Exposures*** | | | | |
| Feeding control 42months | 9542 | | | |
| Can choose any | 981 | 9.9 | 7.7 | 8.9 |
| Choice of few | 5941 | 60.1 | 63.1 | 61.2 |
| Parent decides | 2959 | 30.0 | 29.2 | 29.9 |
| Use food to soothe 42 months | | | | |
| Frequent | 1209 | 12.1 | 10.5 | 11.4 |
| Infrequent | 5112 | 51.2 | 52.1 | 51.2 |
| Never | 3671 | 36.7 | 37.4 | 37.4 |
| ***Covariables*** | | | | |
| Maternal age (year) | 11815 | 28.36 (4.82) | 29.72 (4.27) | 29.16 (4.51) |
| Maternal BMI (kg/m$^2$) | 11534 | 22.93 (3.84) | 22.84 (3.63) | 22.92 (3.73) |
| Maternal education[d] | 12418 | | | |
| None/CSE/Vocational | 3728 | 30.0 | 14.7 | 21.7 |
| O level | 4296 | 34.6 | 35.1 | 35.7 |
| A level | 2794 | 22.5 | 30.2 | 26.6 |
| Degree | 1600 | 12.9 | 20.0 | 16.1 |
| Parental social class | 11505 | | | |
| I/II (highest) | 6342 | 55.1 | 66.9 | 60.1 |
| III | 4481 | 39.0 | 30.2 | 35.5 |
| IV/V (lowest) | 682 | 5.9 | 2.9 | 4.4 |
| Home ownership, rented/other | 13027 | 26.6 | 11.8 | 16.7 |
| Household crowding | 12804 | 6.9 | 2.5 | 4.0 |
| Financial difficulties | 12088 | 10.0 | 6.0 | 7.9 |
| Sole parenting | 13151 | 8.5 | 3.6 | 5.6 |
| Maternal depression | 11697 | 10.2 | 7.3 | 8.5 |
| Smoking in the first 3 months of pregnancy | 13158 | 25.0 | 13.0 | 18.3 |
| Alcohol intake in the first 3 months of pregnancy | 13005 | | | |
| Never | 5917 | 45.5 | 43.6 | 44.0 |
| <1 glass a week | 5034 | 38.7 | 41.2 | 41.0 |
| 1-6 glasses a week | 1804 | 13.9 | 13.7 | 13.3 |
| 1+ glass a day | 250 | 1.9 | 1.5 | 1.7 |
| Number of other children | 11386 | | | |
| None | 4925 | 43.3 | 49.0 | 45.4 |
| 1 | 4201 | 36.9 | 36.3 | 37.1 |
| 2 | 1675 | 14.7 | 11.8 | 13.7 |

| | Response sample[a] | | Complete case sample (n=3082)[b] | Imputed sample (n=7312)[c] |
|---|---|---|---|---|
| | N | Mean (SD) or % | Mean (SD) or % | Mean (SD) or % |
| 3 | 440 | 3.9 | 2.4 | 3.0 |
| ≥4 | 145 | 1.3 | 0.6 | 0.8 |
| Ethnicity, non-White | 13641 | 5.4 | 3.3 | 3.9 |
| Difficult temperament 6 months | 11097 | 14.8 | 14.6 | 14.8 |
| Sex, male | 13976 | 51.7 | 48.2 | 50.8 |
| Birth weight z-score | 13891 | 0.08 (1.00) | 0.14 (0.98) | 0.13 (0.99) |
| Dietary patterns at 38 months[e] | 10014 | | | |
| Junk/Processed | 10014 | -0.00 (1.00) | -0.24 (0.89) | -0.10 (0.95) |
| Health conscious | 10014 | 0.00 (1.00) | 0.07 (1.01) | 0.02 (1.00) |
| Traditional | 10014 | -0.00 (1.00) | -0.01 (1.00) | -0.00 (0.99) |
| Snacks | 10014 | -0.00 (1.00) | 0.13 (0.93) | 0.06 (0.98) |
| Child not eaten enough | 10422 | 64.2 | 65.0 | 64.6 |
| Child refused right food | 10422 | 62.3 | 64.6 | 63.4 |
| Child choosy with food | 10422 | 68.2 | 69.6 | 69.0 |
| Child overeaten | 10422 | 16.6 | 15.1 | 15.8 |
| Child has difficulty to establish eating routine | 10422 | 22.1 | 20.3 | 21.5 |

[a] Response sample is the number who responded to specific assessment for each child exposure, outcome, or covariable.

[b] Complete case sample includes participants who have data on all outcomes, exposures, and covariables.

[c] Imputed sample includes participants who have data on at least one of the three outcomes and one of the two exposures.

[d] CSE= Certificate of Secondary Education. CSE, O, and A levels were high school qualifications. O-levels were usually studied at age 16 and A-levels at age 18.

[e] 'Junk/Processed' pattern was described by foods with high fat and sugar content, and processed and convenience foods. 'Health conscious' pattern was associated with consumption of salads, fruit, vegetables, fish, pasta and rice. 'Traditional' pattern was highly associated with meat, poultry, potato and vegetable consumption. 'Snack' pattern was described as diet consisting primarily of snack and finger foods as opposed to foods where cooking is required.[326]

## 7.4 PART 3: Does temperament modify the association between parental feeding control and adiposity?

### 7.4.1 Introduction

In Part 2 of Chapter 7, the association between parental feeding control and food to soothe and children's adiposity outcomes at ages 7 and 15 years were examined. There was some evidence that lower parental feeding control (but not food to soothe) at age 3.5 years was associated with a higher BMI z-score and fat mass in children at age 7 and 15 years. Part 3 of this Chapter extends these analyses to examine the effect-measure modification by temperament of the association between parental feeding control and adiposity outcomes, *i.e.* whether the association between parental feeding control and adiposity outcomes differs in children with easy versus difficult temperament. The rationale to this analysis has been reviewed in Section 2.5.3. In brief, studies suggested that temperament may predispose an individual's risk of adiposity; temperamentally difficult children whose parents used high feeding control may be more likely to be overweight than temperamentally easy children whose parents used high feeding control.[186]

With regard to the parenting practice of food to soothe, the previous analyses was indicative of a null association with adiposity, and therefore an analysis of effect measure modification by temperament was not undertaken. This section is not included in the article considering the suitability for readers of the journal and the word limitation.

### 7.4.2 Method

The analysis of effect-measure modification approach used in this chapter is similar to the effect-measure modification approach used in Chapter 6. However, there are three categories of parenting in this analysis, instead of the two categories of parenting (*i.e.* 'low warmth and 'high 'warmth') presented in Chapter 6. The three categories are; the child can choose any available food, the child is given a few choices and the parent decides food intake). BMI at 15 years was categorised into two categories: 'healthy' and 'overweight/obese' using the International Obesity Task Force (IOTF) definition.[201] The sex and age specific cut points for BMI were included in Chapter 2. There is no standard definition of overweight using fat mass. For this analysis, excess fat mass was defined as scores above the sex-specific 90$^{th}$ percentile, adjusted for age and height.[338]

### 7.4.3 Result

Table 7.8 shows effect-measure modification by temperament on the association between parental feeding control and overweight (BMI). In the stratum of easy temperament, compared to children in the 'can choose any' group, children in the 'few choices' group were associated with 15% (95% CI 0.71, 1.01) lower risk of overweight while children in the 'parent decides' group were associated with 18% (95% CI 0.68, 1.00) lower risk of overweight. In the stratum of difficult temperament, children in the 'few choices' group had a similar association with overweight as children in the 'can choose any' group (RR=0.99, 95% CI 0.57, 1.67) but children in the 'parent decides' group were associated with 13% lower risk of overweight than children in the 'can choose any' group (RR=0.87, 95% CI 0.48, 1.56). However, the confidence intervals were wide, possibly due to a small number of children in the difficult temperament category.

There was little evidence of an effect-measure modification on the risk-difference and risk-ratio scale of the association between parental feeding control and overweight by temperament. For instance, on the risk-difference scale, the combined effect of parent decided food intake and difficult temperament (0.81) was slightly lower than expected (0.84) when summing the individual effects of parent decides and difficult temperament. On the risk-ratio scale, the combined effect of parent decided food intake and difficult temperament (0.81) was lower than expected (0.84) when multiplying the individual effects of parent decides and difficult temperament. However, the confidence intervals were wide, indicating high variability and inconclusive findings.

Table 7.9 shows effect-measure modification by temperament on the association between parental feeding control and excess fat mass. Results for the fat mass outcome were generally consistent with the overweight outcome in Table 7.8. However, as the proportion of children with excess fat mass was small (10%) and there were fewer children in the difficult temperament category, confidence intervals were very wide in the stratum of difficult temperament.

**Table 7.8 Effect-measure modification by temperament on the association between parental feeding control and overweight[a] at age 15 years (n=7312)**

| | | Parental feeding control | | | | | | RRs (95% CI) for choice of few within strata of temperament | RRs (95% CI) for parent decides within strata of temperament |
|---|---|---|---|---|---|---|---|---|---|
| | | Can choose any | | Choice of few | | Parent decides | | | |
| | | N overweight/healthy | RR (95%CI) | N overweight/healthy | RR (95%CI) | N overweight/healthy | RR (95%CI) | | |
| Temperament | Easy | 131/422 | 1.00 (Ref) | 717/3069 | 0.87 (0.73, 1.04) | 345/1552 | 0.83 (0.69, 1.00) | 0.85 (0.71, 1.01) | 0.82 (0.68, 1.00) |
| | Difficult | 20/77 | 1.01 (0.65, 1.60) | 137/555 | 0.98 (0.77, 1.24) | 50/237 | 0.81 (0.58, 1.14) | 0.99 (0.57, 1.67) | 0.87 (0.48, 1.56) |
| Measure of interaction on risk-difference scale: RERI (95%CI) | | | | 0.98-0.87-1.01+1= 0.09 (-0.39, 0.58) | | 0.81-0.83-1.01+1= -0.03 (-0.56, 0.50) | | | |
| Measure of interaction on risk-ratio scale: Ratio of RRs (95%CI) | | | | 0.98/(0.87*1.01)= 1.10 (0.64, 1.89) | | 0.81/(0.83*1.01)= 0.96 (0.52, 1.77) | | | |

RRs are adjusted for infant temperament, dietary patterns at 3 years, parental use of food to soothe, maternal education, sole parenting, financial difficulties, social class, home ownership, maternal age, gestational age, birth weight, alcohol, smoking, maternal BMI, crowding index, maternal depression, number of other children, ethnicity, child not eaten enough, child refuse right food, child choosy with food, child overeaten, and difficult to establish eating routine.

[a] Overweight was defined using the International Task Force of Obesity (IOTF) age and sex specific BMI cut-offs (at age 15 years, 20% of the imputed sample were overweight/obese)

**Table 7.9 Effect-measure modification by temperament on the association between parental feeding control and excess fat mass[a] at age 15 years (n=7312)**

| | | Parental feeding control | | | | | | RRs (95%CI) for choice of few within strata of temperament | RRs (95%CI) for parent decides within strata of temperament |
|---|---|---|---|---|---|---|---|---|---|
| | | Can choose any | | Choice of few | | Parent decides | | | |
| | | N excess fat mass/healthy | RR (95%CI) | N excess fat mass/healthy | RR (95%CI) | N excess fat mass/healthy | RR (95%CI) | | |
| Temperament | Easy | 82/471 | 1.00 (Ref) | 377/3409 | 0.80 (0.62, 1.03) | 174/1723 | 0.69 (0.52, 0.92) | 0.77 (0.59, 1.00) | 0.69 (0.52, 0.92) |
| | Difficult | 11/86 | 0.81 (0.38, 1.72) | 78/614 | 0.98 (0.71, 1.35) | 23/264 | 0.66 (0.41, 1.07) | 1.33 (0.59, 3.01) | 0.98 (0.41, 2.31) |
| Measure of interaction on risk-difference scale: RERI (95%CI) | | | | 0.98-0.80-0.81+1= 0.37 (-0.27, 1.01) | | 0.66-0.69-0.81+1= 0.16 (-0.53, 0.85) | | | |
| Measure of interaction on risk-ratio scale: Ratio of RRs (95%CI) | | | | 0.98/(0.80*0.81)= 1.52 (0.67, 3.49) | | 0.66/(0.69*0.81)= 1.17 (0.46, 3.01) | | | |

RRs are adjusted for infant temperament, dietary patterns at 3 years, parental use of food to soothe, maternal education, sole parenting, financial difficulties, social class, home ownership, maternal age, gestational age, birth weight, alcohol, smoking, maternal BMI, crowding index, maternal depression, number of other children, ethnicity, child not eaten enough, child refuse right food, child choosy with food, child overeaten, and difficult to establish eating routine.

[a] Excess fat mass was defined as above the sex-specific 90$^{th}$ percentile, adjusted for age and height[338]

### 7.4.4 Discussion

Using a large population-based cohort, the current analysis found little evidence of effect-measure modification by temperament on the association between parental feeding control at age 3.5 years and adiposity outcomes (overweight and excess fat mass) at age 15 years. Results confirmed that children with the least food choice at 3.5 years had lower BMI and fat mass at 15 years than children with the most choice, but this effect was comparable for temperamentally difficult and easy children.

The results presented here and those shown in Part 1 of Chapter 7, found very little evidence to support an association between temperament and adiposity using two different large population samples (ALSPAC and LSAC cohorts). Although the current study included over 7,000 participants, there were problems with small numbers of children in some subgroups, reducing the ability to draw meaningful conclusions. The two other studies in this area have even smaller sample sizes (n=180 and n=197), highlighting an inability of any study in this field to make strong inference.

One limitation of this analysis may be that self-regulation or inhibitory control, which were temperament dimensions that are believed to have an impact on children's eating behaviours, were not measured.[194,307] There is some evidence that the self-regulatory aspects of children's temperament (e.g. emotional, behavioural and attentional control) are modifiable. For example, interventions to improve self-regulation of obese children resulted in reductions to children's body weight.[339,340] However, in summing up, the evidence from this chapter indicates that childhood adiposity is not strongly influenced by

the temperament dimensions used here and that any future interventions would need to involve different targets.

# CHAPTER 8. Summary and Conclusions

The research described in this thesis has investigated the influence of temperament (0 to 5 years) and parenting (2 to 5 years) on children's cognitive, academic and adiposity outcomes in childhood and adolescence. The uniqueness of this body of work is that it applied an epidemiological approach to examine effects of constructs from the field of developmental psychology on cognitive ability, academic achievement and nutrition outcomes. To enhance the robustness and rigor of this research, considerable effort was taken to adopt advanced methodological approaches. The outcomes, IQ, school achievement and adiposity, were collected by different informants (IQ measured by psychologists, school grades reported by teachers and anthropometry collected by health researchers), which are less subject to recall or social desirability biases compared with self-reported outcomes. Furthermore, the research spans two population-based cohort studies from different countries (Australia and the UK), and data were collected almost two decades apart, with two commonly used measures of temperament.

The similarities in the results from these different cohorts, time periods and temperament measures add robustness to the findings. It suggested that the weak or null associations between temperament, parenting and children's outcomes are consistent over time and place, suggesting some generalizability to other similar contexts. In this final chapter, key findings and contributions of this thesis are summarised, limitations are discussed, and potential areas for future research are highlighted.

## 8.1 Key findings and contributions

At the commencement of this doctoral program, the overarching goal was to examine the long-held idea that the construct of innate or constitutional 'temperament' influences outcomes that have life impact (cognitive ability, academic achievement and adiposity), while also accounting for the influence of parenting. Although others have reported that temperament was associated with cognitive ability and academic achievement, these associations may be due to the influence of parenting, *i.e.* child temperament influences parenting, which in turn, affects children's cognitive and academic outcomes. Thus, this goal comprised a main effect, effect measure modifying, and potentially mediation investigations.

The association between temperament and cognitive and academic outcomes was unexpectedly weak in the cohort of Australian children (Chapter 5), with the only convincing association for the temperamental dimension of persistence. The exploration of this study is timely because of considerable interest in temperament as a means to improve academic achievement, for instance, by improving persistence.[9,117] Nevertheless, the small associations between temperament and both cognitive and academic achievement became the motivation to test the concept of differential susceptibility. In epidemiology this is equivalent to effect measure modification. That is, the theory that temperamentally difficult children are more susceptible to the negative effect of parenting practices on a range of different outcomes than temperamentally easy children. The concept of differential susceptibility was explored in a UK cohort because of its large sample, diversity of variables and longitudinal follow-up (Chapter 6). Results were somewhat consistent with the differential susceptibility theory, in that the impacts

of parenting on IQ were different for temperamentally easy compared with difficult children, but the effects were in the opposite direction to those expected. The results indicated that parental warmth and control did not substantially alter the increased risk of having a low IQ among temperamentally difficult children, but lower warmth and higher controlling parenting did appear to increase the risk of a low IQ among temperamentally easy children. While this is not consistent with the differential susceptibility model, it is logical that children with easy temperament benefit from warmth and less controlling parenting. The lack of support for the differential susceptibility model in children with difficult temperament may be due to a small sample of temperamentally difficult children that resulted in wide confidence intervals. Studies with larger samples may be required to further investigate the differential effect of parental warmth and control amongst temperamentally easy and difficult children.

The next focus of the research extended investigations to examine the influences of temperament and parenting on adiposity outcomes. Despite examining temperament and BMI in both the Australian and UK cohorts, there was little to no effect of temperament on BMI (Chapter 7, Part 1). With little evidence of temperament on BMI, the focus of this thesis shifted from temperament to parenting practices as the exposure of interest. Parents play important roles in children's eating, and early feeding practices are believed to shape children's later eating habits and food preferences, which may then have a longer term impact on health.[316] Although effect sizes were small, findings suggested higher parental feeding control was associated with lower BMI and fat mass and this association became stronger with age (Chapter 7, Part 2). This was another unexpected finding as parental feeding practices have been the focus of many childhood

obesity prevention interventions[192] with the general sentiment that lower parental control of child eating is better for adiposity.[318] The mechanism behind this idea is that children may have lower self-regulation in energy intake and are more likely to eat foods that are restricted by parents if parents are highly controlling.[341] This contradiction in findings could be due to the way in which parental control was measured in the ALSPAC cohort 15 years ago, which was more about the child's choice of foods and did not encompass recent, more structured questionnaires for measuring parental control that include other overt and covert controlling practices. As the finding on the association between parental control and BMI was inconsistent with the literature, this thesis then further examined whether the association between parental feeding control and adiposity may differ in children with different temperaments. Findings showed null effects of parental feeding control on adiposity among temperamentally difficult children, but higher parental feeding control was associated with lower BMI and fat mass among temperamentally easy children (Chapter 7, Part 3). These findings are contrary to the literature which suggests that children with difficult temperament are more likely to have higher BMI when parents are highly controlling because they are less able to resist the temptations of palatable foods than children with easy temperament.

One of the most notable contributions of the research in this thesis was the large number of null and weak associations. This is in contrast to the generally stronger findings reported in the fields of developmental psychology and nutrition. There are several possible reasons for the differences in the current findings and previous research which are discussed below:

1. Confounding adjustment

Studies of temperament and cognitive and adiposity outcomes often have poor adjustment for confounding,[63,342] which is likely to result in inflated (biased) effect estimates. Adjusting for confounding in observational studies is important so that the exposed and the unexposed groups are 'exchangeable' and excludes the confounders as an alternative explanation to the observed association between the exposure and the outcome.[217] While randomised trials have the advantages of making causal inference, in some situations observational studies are needed to answer certain research questions.[279] With the goal of achieving conditional exchangeability, observational data are analysed as if the exposure had been randomly assigned, conditional on all the measured covariates.[217] In most situations, adjustment for confounding attenuates an effect estimate. By not adjusting for all known confounders, the effect estimate remains biased.[343] Residual confounding can also be a problem with confounder measurement, as poorly measured confounders can also leave effect estimates open to bias.[344]

2. Sample size and power

Many studies in psychology have small sample sizes that generate low statistical power and therefore, a lower probability of discovering effects that are genuinely true.[345] Low-powered studies are likely to result in widely varying effect estimates and are more likely to have inflated effect sizes.[346,347] This is because in small studies, results are more uncertain and fluctuate more in response to analytical changes.[345] Added to this, there is also a potential of publication bias in smaller, underpowered studies where stronger positive results are more likely to be reported and published.[347,348] Interestingly, the problem of publication bias from small underpowered studies has led to the 'replication' movement in psychology,

where some studies have been independently tested for reproducibility.[349] From 100 studies across the field of psychology, as little as 39% of studies were replicated, which raises concerns more broadly and might explain the null effects observed in this thesis.[349] Studies that reported associations between temperament and cognitive, academic or adiposity outcomes typically involved sample sizes of less than 200.[63,139,171,350]

3. Measurement of temperament

It is possible that the associations between temperament and outcomes mainly lie in sub-dimensions such as 'inhibitory control' that are not measured in this research. However, this thesis attempted to explore the associations between different dimensions of temperament (*e.g.* persistence, reactivity) on cognitive and adiposity outcomes where possible and the small effect sizes were consistent across different temperament dimensions. The psychometric properties of the temperament questionnaires used in the thesis (the RITQ and the STST) have shown moderate to good reliability (see Section 3.1.1.3 and 3.1.2.4). The RITQ has been used in many large-scale, longitudinal studies including the Millennium Cohort Study and the Helsinki Longitudinal Temperament Study.[254,256] The STST was developed from a large sample of Australian children.[59] Typically parents or carers report temperament profiles for young children and although this could be a source of information bias, there is some evidence of consistencies between parent-reported temperament and independent observer ratings.[55] Thus, although there may be differences in measures, it seems unlikely that using self-reported temperament data would attenuate or nullify the association.

4. Construct validity

   The idea that different children have different 'temperaments' has a common sense meaning that is easily recognisable by parents, family and the wider public. Indeed, it is almost obvious that in some sense different 'temperament' does exist in the world. Given that the importance of the concept of temperament has been demonstrated in decades of research, the failure of this thesis to find any strong associations with valued outcomes in cognitive ability, academic achievement and adiposity outcomes may due to the possible reasons suggested above, but it is also possible that temperament does not actually influence the outcomes under investigation in this thesis to a great extent, or at least the broad-brush concept of temperament. While decades of research have generated many plausible theories or models of temperament, it is not clear whether the findings (*e.g.* the associations between temperament and cognitive ability) might differ if other temperament tools were used. However, in part 1 of Chapter 7, associations between temperament and adiposity were consistent even when different temperament tools were used. While it is true that parents and carers can identify different temperaments constructs exhibited by children and that the psychometric properties of temperament can be reliably measured, it is possible that the temperament of infants and young children has no lasting effects on the outcomes measured in this research.

During the course of this doctoral work, it became apparent that the different theories of temperament continue to be highly contested by clinicians, researchers and academics. Users of temperament questionnaires and reviewers often favour one particular theory

or approach and are critical of the use of other theories or tools. A case-in-point is the criticism of the first publication from this research (Chapter 4) by the developers of the RITQ questionnaire which we include as Appendix K. This work was fundamental for understanding how the RITQ performed in a different sample from which it was validated, and was crucial to be able to answer the other research questions of the thesis. The findings suggested the UK sample were more "difficult" (*i.e.* had higher means) on numerous temperamental dimensions than the US standardization sample, which would have resulted in more children being misclassified as "difficult" if US norms were applied (24%) than using empirically-derived ALSPAC norms (15%). Historically, much temperament research has been conducted in clinical samples and in responding to criticisms of this work (Appendix L), it became apparent that there is little appreciation of data collection processes for large-scale population-based studies like LSAC and ALSPAC. Decisions about which temperament data are to be collected occurs many years in advance of outcome measurements – this has to occur for researchers to investigate long-term effects of early life exposures. While large population-based longitudinal cohort studies and contemporary epidemiological methods have not been widely used in developmental psychology and nutrition, the flip-side of this view is that there remains an enormous opportunity to generate more rigorous work in this area through multi-disciplinary collaborations and applying more advanced statistical methods.

## 8.2 Limitations and future directions

Limitations of each individual study have been discussed in the relevant chapter and will not be repeated here. This section focuses on general limitations of the thesis and potential areas for future research.

Although this thesis has the advantage of using data from two population-based studies to help triangulate the findings, both ALSPAC and LSAC shared some similarities in cultural backgrounds. The associations between parenting and temperament on cognitive and adiposity outcomes may vary in different cultural groups. For instance, the effect of parental control on IQ may be different in European compared with Asian cultures.[95] Hence, future research may consider using cohorts with different cultural and ethnicity backgrounds to give more insight.

As with other longitudinal studies, parenting practices and child temperament were solely dependent upon parents' (usually mothers') self-report and this may be subject to response bias. For instance, most parents reported that they showed high levels of warmth and affection to their children. Moreover, using a single informant (mother) to report parenting and child temperament may be influenced by the mother's personality or experience. Using multiple informants and direct observation may increase the validity of the parenting and temperament data, however, it is more practical to use parent-reported questionnaires than direct observations in large-scale population studies due to cost, time and labour constraints.

In studies 3 and 4, although the associations between parenting and IQ and adiposity were adjusted for child temperament, these adjustments may not rule out the possibility that parenting and child temperament may influence each other bi-directionally (*i.e.* parenting may be an effect of child temperament and parenting also affect child temperament).[351] However, analyses in this thesis have included a wide range of

confounding factors, it is unlikely that results would have differed greatly from what is observed here.

In Study 4, the parental feeding control was measured using a single item about the extent of the child was given food choices. There is not a standard definition and measure of feeding control. A better clarification of what is meant by feeding control is needed and questionnaires that cover overt and covert control may give more insight into which practices are considered "bad" for the child's BMI and which practices are "good" for the child's BMI.

## 8.3 Implications and recommendations

Despite the evidence from previous studies that suggested temperament may be a means to improve academic achievement,[9,117] findings from this thesis suggested the effect sizes of temperament dimensions of approach, reactivity and persistence and cognitive and academic outcomes are very small. Although persistence has the largest effect on cognitive and academic outcomes, temperament encompasses multiple dimensions including approach and reactivity that have little effect on outcomes. Improving child temperament *e.g.* persistence per se may not be the key approach to improve academic skills. Interventions that encourage supportive parenting and reduce harsh parenting (*e.g.* use of verbal and physical punishments) may be beneficial for children's cognitive outcomes. Efforts to reduce childhood overweight may need to adopt a more direct approach that intervenes on nutritional intake or food environment rather than focusing on improving child temperament or parental feeding control itself. This is particularly evident given that the high-quality and high-intensity randomised controlled trials (*e.g.*

involving home visits by nutrition-trained health care workers), which target parenting feeding behaviours have shown negligible effects on children's BMI.[192,334] Stronger evidence, preferably from randomised controlled trials of parenting interventions, are needed before any further recommendations can be made.

## 8.4 Concluding remarks

Using two large, longitudinal observational studies from different countries, different temperament tools, and measures of temperament at different ages, the research in this thesis indicated that the effect sizes for temperament on cognitive, academic and adiposity outcomes are at best, very small. Parental warmth and control had small effect sizes on children's IQ while higher parental feeding control was associated with lower adiposity. The differential susceptibility theory, suggested by previous psychological studies that temperamentally difficult children were more vulnerable to the detrimental effects of negative parenting, was not supported. Interventions to improve cognitive, academic, and adiposity outcomes may need to involve a broader environmental context than improving temperament per se.

## 8.5 Appendices

Appendix A is a commentary from the RITQ developers on the first paper in this thesis (Chapter 4). The commentary was published in the journal *Infant Behavior and Development* in August 2015. The two RITQ developers have shown strong reactions on the first paper. Appendix B is the authors' response to the commentary which was published in the journal *Infant Behavior and Development* in August 2015.

**Appendix K Carey WB, McDevitt SC. Commentary on paper by Chong et al. "How many infants are temperamentally difficult?"40 (2015) 20–28. *Infant Behav Dev*. 2015;41:167-168.**

A recent paper in this Journal by Chong *et al.* "How many infants are temperamentally difficult?" (40:20-28;2015) concluded that the scoring procedures for the Revised Infant Temperament Questionnaire (RITQ), which were based on a smaller sample 40 years ago, are incorrect and should be replaced by their recent data from their larger population of over 10,000 infants in the UK. They also claim that they found a larger percentage of difficult infants than was established by our earlier data. The study is of interest to temperament researchers but has not accomplished its stated goal of providing better norms for the scale. There are several reasons for their failure.

1) Need to restandardize for use abroad. We have maintained and urged for decades that when one of our temperament scales is used in another culture, even where the language is superficially similar, the scale should be restandardized. (Carey 2009). Differences in the meanings of words and varying cultural values of behaviors cause differing interpretations of the items with resulting shifts in the outcome of the measure. Small variations in understanding of the words may make an item load onto a different trait. There are also some differing cultural views as to what constitutes temperamental difficulty. See Super et al. (2009) (Most differences in scores of the traits discovered by this study were small, less than ½ SD in all. but one trait.

2) Removal of items. The elimination of 7 items because they were not rated frequently enough both confirms the difference in their study populations and helps to explain some of the variation in results.

3) Selection of subjects. We are not told how the subjects were selected. What was the refusal and dropout rate? Were worried parents more likely to participate? By contrast, our standardization sample was drawn from two diverse but compliant suburban American pediatric practices. Those participating from Dr. Carey's practice in a suburb of Philadelphia were 95% of the eligible infants invited

4) The elimination of the "slow to warm up" (STWU) diagnostic option. The avoidance of the original slow-to-warm-up classification, urged by the original Chess and Thomas (1986) system and mentioned in the original publication of the ITQ (Carey, 1970) probably inflated their count of designated difficult infants in their results. The two groups difficult and STWU, are very closely related.

5) Misunderstanding of the phenomenon of temperamental difficulty. The early work of Chess and Thomas (1986) and Carey (1970) made use of the several diagnostic clusters for their studies of their behavioral problem outcomes. However, we clinical users of these clusters learned long ago that a questionnaire rating of difficult or easy often does not agree with the parental general perception of difficulty, an evaluation revealed by the rater in the final question on the scale. We have come to realize that, while certain traits are temperament risk factors, what matters clinically is how the parents view and interact with them, the goodness or poorness of fit. Unfortunately some users, especially academic researchers, continue to view difficult temperament as a numerical score on a scale, not as a demonstrable significant mismatch between a normal but aversive child and parents who did not get a temperament they like and want.

Any researchers, clinical or academic, wanting further clarification of these points or guidance in constructing definitive studies elsewhere, should feel welcome to contact us for help and support. We have gladly been responsive to many attempting to use our

scales in other places. One of us (WBC) helped to evaluate the translation and backtranslation check, which we advise for moving to a different language, and to select the appropriate items in the foreign language version in Italian. (Axia 1993).

**Appendix L Smithers LG, Chong SY, Chittleborough CR, Gregory T, Lynch JW. Authors respond to the commentary on Chong et al. "How many infants are temperamentally difficult?" (40:20-28; 2015) *Infant Behav Dev*. 2015; 41: 164-166.**

| | |
|---|---|
| Title of Paper | Authors respond to the commentary on Chong et al. "How many infants are temperamentally difficult?" (40:20-28; 2015) |
| Publication Status | ☑ Published<br>☐ Accepted for Publication<br>☐ Submitted for Publication<br>☐ Unpublished and Unsubmitted w ork w ritten in manuscript style |
| Publication Details | Smithers LG, Chong SY, Chittleborough CR, Gregory T, Lynch JW. Authors respond to the commentary on Chong et al. "How many infants are temperamentally difficult?" (40:20-28; 2015) *Infant Behav Dev*. 2015; 41: 164-166. |

## Principal Author

| | | | |
|---|---|---|---|
| Name of Principal Author | Lisa Smithers | | |
| Contribution to the Paper | Contributed to the design of the study, interpreted the data, and drafted the manuscript. I give consent for Shiau Yun Chong to present this paper for examination towards the Doctor of Philosophy. | | |
| Overall percentage (%) | 85 | | |
| Certification: | I am the primary author of this paper. | | |
| Signature |  | Date | 15-08-2016 |

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

xiii. the candidate's stated contribution to the publication is accurate (as detailed above);

xiv. permission is granted for the candidate in include the publication in the thesis; and

xv. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

| | | | |
|---|---|---|---|
| Name of Co-Author (Candidate) | Shiau Yun Chong | | |
| Contribution to the Paper | Contributed to the design of the study and reviewed the manuscript. This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. | | |
| Signature |  | Date | 15-08-2016 |

| | | | |
|---|---|---|---|
| Name of Co-Author | Catherine Chittleborough | | |
| Contribution to the Paper | Contributed to the design of the study and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| | | | |
|---|---|---|---|
| Name of Co-Author | Tess Gregory | | |
| Contribution to the Paper | Contributed to the design of the study and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

| | | | |
|---|---|---|---|
| Name of Co-Author | John Lynch | | |
| Contribution to the Paper | Contributed to the design of the study and reviewed the manuscript. | | |
| Signature | | Date | 15-08-2016 |

As stated in our manuscript,[244] our aim was to compare the original norms (means, standard deviations) published for the Revised Infant Temperament Questionnaire (RITQ)[42,202] with norms empirically derived from a large (n=10937) population sample of UK infants, and to compare the categorisation of temperament using these two different norms. Our paper highlights problems with applying the norms published in the RITQ to the UK population, such that misclassification of temperament types (*e.g.* easy, difficult) must be occurring for the two norms to arrive at profoundly different temperament profiles within our sample. Here we first respond by correcting errors, and then by providing discussion of each of the five points made in a commentary on our article by two renowned temperament researchers.[352]

***Errors in the Carey & McDevitt Commentary***

Carey & McDevitt[352] state that we wanted to replace the scoring procedures of the RITQ with our UK data and that our goal was to provide "better" norms. We made no such claims. Rather, we suggested that the norms from our study may be more appropriate for use with UK infants, and across all countries and contexts, we advocated a need for more up to date norms than those derived in 1978 from a small clinical sample.

Carey & McDevitt [352] state that we found a larger percentage of difficult infants than was established by their original data. This is patently incorrect and in fact, the opposite is true. This point is easily resolved by reading the abstract. To reiterate here, when applying the original RITS norms to our dataset, 24% of children were categorised as having difficult temperament, whereas only 15% were categorised as difficult when empirically-derived norms are applied.[244]

***Response to the Carey & McDevitt Commentary***

Need to re-standardise the RITQ: We agree with Carey & McDevitt that culturally appropriate and contemporary norms are urgently required for the RITQ to be useful for current paediatric practice. Super *et al.*[353] show that there are differences in temperament across settings, and Smart & Sanson[68] show secular changes in temperament measured within a country over a 20-year period (also known as cohort effects). Our data agree with Carey & McDevitt that the differences in temperament subscale scores between our data and the original RITQ were less than 0.5 SD across all RITQ subscales (ranging from 0.1SD to 0.4SD) except Adaptability (which was 0.7 SD). However, our research went further to clearly show that seemingly small differences in the distribution of temperamental subscales can lead to large differences when categorising children to temperament types (*e.g.* difficult).

Removal of items: Some items in the RITQ are unanswerable for some infants. For example, the question, '*The infant is fussy or moody throughout a cold or an intestinal virus?*' cannot be answered by carers of 6-month-old infants who have not yet experienced a cold or an intestinal virus. In our study, this and other questions were removed following 10% non-response during pilot testing. The values for internal consistency that we reported were similar to those published for the RITQ. Super *et al.*[353] put forward a strong case for removing questionnaire items to obtain more culturally appropriate measures of temperament. Carey & McDevitt[352] cite Super *et al.*[353] to argue the need to re-standardise the RITQ but they failed to acknowledge that Super *et al*[353] also removed items for cultural reasons, as we did in our study.

Selection of subjects: Detailed information about the cohort is provided in the methods section of our paper. To recap here, the 10937 children included in our study were from the Avon Longitudinal Study of Parents and Children (ALSPAC). ALSPAC is an extremely well-established, population-based cohort from the Avon area in the UK. Enrolment to ALSPAC was open to all women living in the Avon area and is estimated to include 72% of the local population.[197] The sample is considered representative of the area at the time of enrolment in 1991-1992.[197] The recruitment did not specifically seek carers who were worried about their child's temperament, although such participants will have been included due to the population-based nature of the sample. By comparison, the RITQ norm sample involved only 203 children from advantaged backgrounds who were brought to the clinics of one paediatrician known for his temperament research. The sample of "eligible infants" in the RITQ norm sample was small and not representative of the population, and thus a high participation rate of 95% will still have produced a small non-representative sample of children. Furthermore, compliance of 95% by Carey & McDevitt[352] could be indicative of a highly motivated sample with respect to concerns about children's temperament and/or clinical care. Thus ALSPAC is a very large relatively unselected study population in comparison to the small clinical sample used to norm the RITQ.

'Slow-to-warm up' (STWU) infants: We did not include the STWU category in our study for two reasons. First, information on the categorisation of temperament types is not provided in the RITQ manual and therefore we used the categorisation recommended by Carey,[41] which does not include a well-defined STWU category. Second, Carey[291] has

previously argued that staying true to the original concept of “difficult” temperament is needed for progress on research and practice.

The phenomenon of temperamental difficulty: Knowledge about temperament is not the exclusive domain of clinicians. Academics have advanced the study of temperament even when they do not practice as clinicians. We acknowledged the ‘goodness of fit’ in our paper and how parenting advice might be tailored according to the temperament of the child. As we see it, a carer’s response to the final question on the RITQ about ‘goodness of fit’ is only part of the characterisation of child temperament. If it were the whole picture, then the RITQ in its current form would no longer be needed, as we could simply ask this one question.

Regular re-norming of psychometric tools is standard practice. Indeed, the adult version of the RITQ (the Adult Temperament Questionnaire 2nd Edition) was updated in 2008 and 2013 (https://www.b-di.com/atqweb.html). Given that the norms for the children’s RITQ haven’t been updated since 1978 it would be wise for current and future users of the RITQ to consider whether their use of the RITQ is appropriate for contemporary settings (depending of course on the intended use of the RITQ). Re-norming is both important within a country over time, and across countries and cultures at a point in time. Ironically, Carey & McDevitt claim to endorse the collection of culturally-specific norms but disparage research (such as ours), which generates the evidence to show that more appropriate norms are required.

# References


1. Bronfenbrenner U. *The ecology of human development: experiments by nature and design*. Cambridge, MA: Harvard University Press; 1979.
2. Zubrick SR, Smith G, Nicholson JM, Sanson A, Jackiewicz T. *Parenting and families in Australia*. Australia: Australian Government Department of Families, Housing, Community Services and Indigenous Affairs; 2008.
3. Prior M, Sanson A, Smart D, Oberklaid F. *Pathways from infancy to adolescence: Australian Temperament Project 1983-2000*. Melbourne: Australian Institute of Family Studies– Commonwealth of Australia; 2000.
4. Kagan J. *Galen's prophecy*. New York: Basic Books; 1994.
5. Rothbart MK. *Becoming Who We Are: Temperament and Personality in Development*. New York, United States of America: The Guilford Press; 2011.
6. Rothbart M, Bates J. Temperament. In: Damon W, Lerner R, editors. *Handbook of child psychology*. 6th ed. Hoboken, New Jersey: John Wiley & Sons Inc.; 2006. p. 99-164.
7. Sanson A, Hemphill SA, Smart D. Connections between temperament and social development: A review. *Soc Dev*. 2004;13:142-170.
8. Kiff C, Lengua L, Zalewski M. Nature and nurturing: Parenting in the context of child temperament. *Clin Child Fam Psychol Rev*. 2011;14:251-301.
9. Blair C, Diamond A. Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Dev Psychopathol*. 2008;20:899-911.
10. Diamond A, Barnett WS, Thomas J, Munro S. Preschool program improves cognitive control. *Science*. 2007;318:1387-1388.
11. Raver CC, Jones SM, Li-Grining C, Zhai F, Bub K, Pressler E. CSRP's impact on low-income preschoolers' preacademic skills: Self-regulation as a mediating mechanism. *Child Dev*. 2011;82:362-378.
12. Hutchings J, Bywater T, Daley D*, et al.* Parenting intervention in Sure Start services for children at risk of developing conduct disorder: Pragmatic controlled trial. *Br Med J*. 2007;334:1-7.
13. Lindsay G, Strand S, Davis H. A comparison of the effectiveness of three parenting programmes in improving parenting skills, parent mental-well being and children's behaviour when implemented on a large scale in community settings in 18 English local authorities: The parenting early intervention pathfinder (PEIP). *BMC Public Health*. 2011;11:962.
14. Eshel N, Daelmans B, de Mello M, Martines J. Responsive parenting: Interventions and outcomes. *Public Health Rev*. 2006;84:992-999.
15. Walker A, Chang S, Powell C, Grantham-McGregor S. Effects of early childhood psychosocial stimulation and nutritional supplementation on cognitive and education in growth-stunted Jamican children: Prospective cohort study. *Lancet*. 2005;366:1804-1807.
16. Sweet M, Appelbaum M. Is home visiting an effective strategy? A meta-analysis review of home visiting programs for families with young children. *Child Dev*. 2004;75:1435-1456.
17. Fergusson DM, Horwood LJ, Ridder EM. Show me the child at seven II: Childhood intelligence and later outcomes in adolescence and young adulthood. *J Child Psychol & Psychiatry*. 2005;46:850-858.

18. D'Angiulli A, Warburton W, Dahinten S, Hertzman C. Population-level associations between preschool vulnerability and grade-four basic skills. *PLoS One*. 2009;4:e7692.
19. Jokela M, Batty GD, Deary IJ, Gale CR, Kivimäki M. Low childhood IQ and early adult mortality: The role of explanatory factors in the 1958 British Birth Cohort. *Pediatrics*. 2009;124:e380-e388.
20. Kerckhoff A, Raaudenbush S, Glennie E. Education, cognitive skills, and labor force outcomes. *Socio Edu*. 2001;74:1-24.
21. Jones G, Schneider W. Intelligence, human capital, and economic growth: A Bayesian averaging of classical estimates (BACE) approach. *J Econ Growth*. 2006;11:71-93.
22. Word Health Organization. Obesity and overweight. http://www.who.int/mediacentre/factsheets/fs311/en/. Published 2015. Accessed 23th June 2015.
23. Commonwealth of Australia. 2007 Australian National Children's Nutrition and Physical Activity Survey. Australia; 2008.
24. Owen C, Whincup P, Orfei L*, et al.* Is body mass index before middle age related to coronary heart disease risk in later life? Evidence from observational studies. *Int J Obes*. 2009;33:866-877.
25. Reilly JJ, Bonataki M, Leary S*, et al.* Progression from childhood overweight to adolescent obesity in a large comtemporary cohort. *Int J Obes*. 2011;6:e138-143.
26. Shiner RL, DeYoung CG. The structure of temperament and personality traits: A developmental perspective. University of Chicago; 2011.
27. Thomas A, Chess S, Birch H, Hertzig M, Korn S. *Behavioral individuality in early childhood*. New York: New York University Press; 1963.
28. Thomas A, Chess S, Birch H. The origin of personality. *Sci Am*. 1970;222:102-109.
29. Hartzig M, Birch H, Thomas A, Mendez O. Class and ethnic differences in the responsiveness of preschool children to cognitive demands *Monogr Soc Res Child Dev*. 1968;33:1-69.
30. Hertzig M. Neurologic findings in prematurely born children at school age. In: *Life history research in psychopathology* Minneapolis: University of Minnesota Press; 1974. p. 42-52.
31. Stella C, Mahin H. Behavior deviations in mentally retarded children. *J Am Acad Child Psychiatry*. 1970;9:282-297.
32. Chess S, Korn S, Fernandez P. *Psychiatric disorders of children with congenital rubella*. New York: Brunner/Mazel; 1971.
33. Thomas A, Chess S. *Temperament and development*. Brunner/Mazel: New York; 1977.
34. Rowe D, Plomin R. Temperament in early Childhood. *J Pers Assess*. 1977;41:150-156.
35. Goldsmith HH, Buss AH, Plomin R*, et al.* Roundtable: What Is Temperament? Four Approaches. *Child Dev*. 1987;58:505-529.
36. Garcia-Coll C, Kagan J, Reznick J. Behavioral inhibition in young children. *Child Dev*. 1984;55:1005-1019.
37. Sanson A, Hemphill SA, Yagmurlu B, McClowry S. Temperament and social development. In: Smith P, Hart S, editors. *The Wiley-Blackwell Handbook of Childhood Social Development*. Oxford, UK: Wiley-Blackwell; 2011. p. 227-245.

38. Chess S, Thomas A. *Origins and evolution of behavior disorders*. Cambridge, MA: Harvard University Press; 1987.
39. Steinberg L. Early temperament antecedents of adult Type A behaviors. *Dev Psychol*. 1985;21:1171-1180.
40. Gartstein MA, Bridgett DJ, Low CM. Chapter 10: Asking questions about temperament. In: Zentner M, Shiner R, editors. *Handbook of temperament*. New York: The Guilford Press; 2012. p. 183-207.
41. Carey WB. A simplified method for measuring infant temperament. *J Pediatr*. 1970;77:188-194.
42. Carey WB, McDevitt SC. Revision of the Infant Temperament Questionnaire. *Pediatrics*. 1978;61:735-739.
43. Rothbart M, Ellis LK, Rosario Rueda M, Posner M. Developing mechanisms of temperamental effortful control. *J Pers*. 2003;71:1113-1143.
44. Rothbart MK, Mauro JA. Questionnaire Approaches to the Study of Infant Temperament. In: Fagen JW, Colombo J, editors. *Individual differences in infancy: Reliability, stability,and prediction*. Hillsdale, NJ: Erlbaum; 1990. p. 411-429.
45. Buss A, Plomin R. *A temperament theory of personality temperament*. New York Wiley; 1975.
46. Plomin R, Daniels D. Why are children in the same family so different from one another? *Int J Epidemiol*. 2011;40:563-582.
47. Saudino KJ, McGuire S, Reiss D, Hetherington EM, Plomin R. Parent ratings of EAS temperaments in twins, full siblings, half siblings, and step siblings. *J Pers Soc Psychol*. 1995;68:723-733.
48. Saudino KJ. Behavioral genetics and child temperament. *J Dev Behav Pediatr*. 2005;26:214-223.
49. Saudino KJ, Micalizzi L. Emerging trends in behavioral genetic studies of child temperament. *Child Dev Perspect*. 2015;9:144-148.
50. Goldsmith H. Studying temperament via Construction of the Toddler Behavior Assessment Questionnaire. *Child Dev*. 1996;67:218-235.
51. Gagne J, Van Hulle CA, Aksan N, Essex MJ, Goldsmith H. Deriving childhood temperament measures from emotion-eliciting behavioral episodes: Scale construction and initial validation. *Psychol Assess*. 2011;23:337-353.
52. Kagan J. Temperament and the reactions to unfamiliarity. *Child Dev*. 1997;68:139-143.
53. Kagan J, Snidman N, Kahn V, Towsley S. The perseverance of two infant temperaments into adolescence. *Monogr Soc Res Child Dev*. 2007;72:1-95.
54. Rothbart M, Alansky J. Temperament, behavioral inhibition, and shyness in childhood. In: Leitenberg H, editor. *Handbook of social and evaluation anxiety*. New York: Plenum; 1990. p. 139-160.
55. Bishop G, Spence S, McDonald C. Can parents and teachers provide a reliable and valid report of behavioral inhibition? *Child Dev*. 2003;74:1899-1917.
56. Vassallo S, Sanson A. *The Australian Temperament Project: the first 30 years*. Australia: Australian Institute of Family Studies; 2013.
57. Sanson A, Prior M, Garino E, Oberklaid F, Sewell J. The structure of infant temperament: Factor analysis of the Revised Infant Temperament Questionnaire. *Infant Behav Dev*. 1987;10:97-104.
58. Fullard W, McDevitt S, Carey W. Assessing temperament in one- to three-year old children. *J Pediatr Psychol*. 1984;9:205-217.

59. Prior M, Sanson A, Oberklaid F. The Australian Temperament Project. In: Kohnstamm G, Bates J, Rothbart M, editors. *Temperament in childhood*. Chichester, England: Wiley; 1989. p. 537-554.
60. Mervielde I, Clercq B, Fruyt F, Leeuwen K. Temperament, personality, and developmental psychopathology as childhood antecedentcs of personality disorders. *J Pers Disord*. 2005;19:171-201.
61. Van Leeuwen K, Mervielde I, Braet C, Bosman G. Child personality and parental behavior as moderators of problem behavior: variable- and person-centered approaches. *Dev Psychol*. 2004;40:1028-1046.
62. Palisin H. Preschool temperament and performance on achievement tests. *Dev Psychol*. 1986;22:766-770.
63. Martin RP, Holbrook J. Relationship of temperament characteristics to the academic achievement of first-grade children. *J Psychoeduc Assess*. 1985;3:131-140.
64. Rothbart M, Posner M. Temperament, attention, and developmental psychopathology. In: Cicchetti D, Cohen D, editors. *Developmental psychopathology: Vol 1 Theory and methods*. New York: Wiley. p. 465-501.
65. Asendorpf J. Development of inhibition during childhood: Evidence for situational specificity and a two-factor model. *Dev Psychol*. 1990;26:721-730.
66. Bishop G, Spence SH, McDonald C. Can parents and teachers provide a reliable and valid report of behavioral inhibition? *Child Dev*. 2003;74:1899-1917.
67. Roberts B, DelVecchio W. The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. *Psychol Bull*. 2000;126:3-25.
68. Smart D, Sanson A. A comparison of children's temperament and adjustment across 20 years. *Fam Matters*. 2005;7:50-57.
69. Oberklaid F, Sanson A, Prior M. The development of Australian normative data for infant temperament. *Aust Paediatr J*. 1986;22:185-187.
70. Mednick B, Hocevar D, Schulsinger C, Baker R. Personality and demographic characteristics of mothers and their ratings of their 3-to 10-year-old children's temperament. *Merrill Palmer Q*. 1996;42:397-417.
71. Sanson A, Letcher P, Smart D, Prior M, Toumbourou JW, Oberklaid F. Associations between early childhood temperament clusters and later psychosocial adjustment. *Merrill Palmer Q*. 2009;55:27-54.
72. Kim P, Evans D, Angstadt M*, et al.* Effects of childhood poverty and chronic stress on emotion regulatory brian function in adulthood. *Proc Natl Acad Sci U S A*. 2013;110:18442-18447.
73. Blair C, Raver CC. Child development in the context of adversity. *Am Psychol*. 2012;67:309-318.
74. Pickett K, Wood C, Adamson J, DeSouza L, Wakschlag L. Meaningful differences in maternal smoking behaviour during pregnancy: implications for infant behavioral vulnerability. *J Epidemiol Community Health*. 2007;62:318-324.
75. Xinyin C, Fan Y, Rui F. Chapter 22: Culture and temperament. In: Zentner M, Shiner R, editors. *Handbook of temperament*. New York: The Guilford Press; 2012. p. 462-478.
76. Hsu C, Soong W, Stigler J, Hong C, Liang C. The temperamental characteristics of Chinese babies. *Child Dev*. 1981;52:1337-1340.

77. Maccoby E, Martin J. Socialization in the context of the family: parent-child interaction. In: Heatherington E, editor. *Mannual of Child Psychology*. New York: Wiley; 1983. p. 1-102.
78. Repetti RL, Taylor SE, Seeman TE. Risky families: family social environments and the mental and physical health of offspring. *Psychol Bull*. 2002;128:330-366.
79. Belsky J, Bell B, Bradley RH, Stallard N, Stewart-Brown SL. Socioeconomic risk, parenting during the preschool years and child health age 6 years. *Eur J Public Health*. 2007;17:508-513.
80. Schaefer E. A configurational analysis of children's reports of parent behavior. *J Consult Psychol*. 1965;29:552-557.
81. Becker W, Peterson D, Luria Z, Shoemaker D, Hellmer L. Relations of factors derived from parent-interview ratings to behavior problems of five-year-olds. *Child Dev*. 1962;33:509-535.
82. Rohner R, Pettengill S. Perceived parental acceptance-rejection and parental control among Koren adolescents. *Child Dev*. 1985;56:524-528.
83. Schaefer E, Bell R, Bayley N. Development of a maternal behavior research instrument. *J Genet Psychol*. 1959;95:83-104.
84. Amato PR. Dimensions of the family environment as perceived by children: A multidimensional scaling analysis. *J Marriage Fam*. 1990;52:613-620.
85. Zubrick SR, Lucas N, Westrupp EM, Nicholson J. *Parenting measures in the Longitudinal Study of Australian Children: Construct validity and measurement quality, Waves 1 to 4*. Canberra: Department of Social Services; 2014.
86. Fergusson D, Horwood L, Lynskey M. The childhoods of multiple problem adolescents: A 15-year longitudinal study. *J Child Psychol Psychiatry*. 1994;35:1123-1140.
87. Landry S, Smith K, Swank P. Responsive parenting: Establishing early foundations for social, communication, and independent problem-solving skills. *Dev Psychol*. 2006;42:627-642.
88. Barber B, Stoltz H, Olsen J. Parental support, psychological control, and behavioral control: Assessing relevance across time, culture, and method. *Monogr Soc Res Child Dev*. 2005;70:1-151.
89. Grolnick W. *The psychology of parental control: How well-meant parenting backfires*. New York: Psychology Press; 2002.
90. Grolnick WS, Price CE, Beiswenger KL, Sauck CC. Evaluative pressure in mothers: Effects of situation, maternal, and characteristics on autonomy supportive versus controlling behavior. *Dev Psychol*. 2007;43:991-1002.
91. Grolnick W, Ryan R. Parent styles associated with children's self-regulation and competence in school. *J Educ Psychol*. 1989;81:143-154.
92. Qin L, Pomerantz EM, Wang Q. Are gains in decision-making autonomy during early adolescence beneficial for emotional functioning? The case of the United States and China. *Child Dev*. 2009;80:1705-1721.
93. Joussemet M, Vitaro F, Barker E*, et al.* Controlling parenting and physical aggression during elementary school. *Child Dev*. 2008;79:411-425.
94. Wang Q, Pomerantz EM, Chen H. The role of parents' control in early adolescents' psychological functioning: a longitudinal investigation in the United States and China. *Child Dev*. 2007;78:1592-1610.
95. Byford M, Kuh D, Richards M. Parenting practices and intergenerational associations in cognitive ability. *Int J Epidemiol*. 2011;41:263–272.

96. Tamis-LeMonda CS, Shannon JD, Cabrera NJ, Lamb ME. Fathers and mothers at play with their 2- and 3-year-olds: Contributions to language and cognitive development. *Child Dev*. 2004;75:1806-1820.
97. Grantham-McGregor S, Schofield W, Powell C. Development of severely malnourished children who received psychosocial stimulation: Six-year follow-up. *Pediatrics*. 1987;79:247-254.
98. Desforges C, Abouchaar A. The impact of parental involvement, parental support and family education on pupil achievements and adjustment: A literature review: Department for Education and Skills; 2003. Report No.: DfES Research Report 433.
99. Fisher J, Birch L. Parents' restrictive feeding practices are associated with young girls' negative self-evaluation of eating. *J Am Diet Assoc*. 2000;100:1341-1346.
100. Birch L, Fisher J, Davison K. Learning to overeat: Maternal use of restrictive feeding practices promotes girls' eating in the absence of hunger. *Am J Clin Nutr*. 2003;78:215-220.
101. Birch L, Fisher J. Mothers' child-feeding practices influence daughters' eating and weight. *Am J Clin Nutr*. 2000;71:1054-1061.
102. Stifter CA, Moding K. Understanding and measuring parent use of food to soothe infant and toddler distress: A longitudinal study from 6 to 18 months of age. *Appetite*. 2015;95:188-196.
103. Stifter CA, Anzman-Frasca S, Birch LL, Voegtline K. Parent use of food to soothe infant/toddler distress and child weight status. An exploratory study. *Appetite*. 2011;57:693-699.
104. Vaughn A, Tabak R, Bryant M, Ward D. Measuring parent food practices: a systematic review of existing measures and examination of instruments. *Int J Behav Nutr Phys Act*. 2013;10:1-27.
105. Hughes S, Frankel L, Beltran A*, et al.* Food parenting measurement issues: Working group consensus report. *Child Obes*. 2013;9:S95-S102.
106. Campbell K, Andrianopoulos N, Hesketh K*, et al.* Parental use of restrictive feeding practices and child BMI z-score. A 3-year prospective cohort study. *Appetite*. 2010;55:84-88.
107. Galloway AT, Fiorito LM, Francis LA, Birch LL. 'Finish your soup': counterproductive effects of pressuring children to eat on intake and affect. *Appetite*. 2006;46:318-323.
108. Wehrly S, Bonilla C, Perez M, Liew J. Controlling parental feeding practices and child body composition in ethnically and economically diverse preschool children. *Appetite*. 2014;73:163-171.
109. Birch LL, Fisher JO, Grimm-Thomas K, Markey CN, Sawyer R, Johnson SL. Confirmatory factor analysis of the Child Feeding Questionnaire: a measure of parental attitudes, beliefs and practices about child feeding and obesity proneness. *Appetite*. 2001;36:201-210.
110. Odgen J, Reynolds R, Smith A. Expanding the concept of parental control: A role for overt and covert control in children's snacking behaviour? *Appetite*. 2006;47:100-106.
111. Rodenburg G, Kremers SP, Oenema A, Van De Mheen D. Associations of parental feeding styles with child snacking behaviour and weight in the context of a general parenting. *Public Health Nutr*. 2013;17:960-969.

112. Braden A, Rhee K, Peterson C, Rydell S, Zucker N, Boutelle K. Associations between child emotional eating and general parenting style, feeding practices, and parent psychopathology. *Appetite*. 2014;80:35-40.
113. Little K, Sanson A, Zubrick S. Do individual differences in temperament matter for Indigenous children? The structure and function of temperament in Foortprints in Time. *Fam Matters*. 2012;91:92-105.
114. Putnam SP, Sanson A, Rothbart MK. Child temperament and parenting. In: Bornstein MH, editor. *Handbook of Parenting: Volume I: Children and Parenting*. United States of America: Lawrence Erlbaum Associates.
115. Heckman JJ, Kautz T. Hard evidence on soft skills. *Labour Econ*. 2012;19:451-464.
116. Duckworth AL, Allred KM. Temperament in the Classroom. In: Shiner RL, Zentner M, editors. *Handbook of temperament*. New York: Guilford Press; 2012. p. 627-644.
117. Heckman JJ, Rubinstein Y. The importance of noncognitive skills: Lessons from the GED testing program. *Am Econ Rev*. 2001;91:145-149.
118. Salley B, Dixon W. Temperamental and joint attentional predictors of language development. *Merrill Palmer Q*. 2007;53:131-154.
119. Rabiner D, Coie J. Early attention problems and children's reading achievement: A longitudinal investigation. *J Am Acad Child Adolesc Psychiatry*. 2000;39:859-867.
120. Baron R, Kenny D. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51:1173-1182.
121. Cole SR, Platt RW, Schisterman EF*, et al.* Illustrating bias due to conditioning on a collider. *Int J Epidemiol*. 2010;39:417-420.
122. Nandi A, Glymour MM, Kawachi I, VanderWeele TJ. Using marginal structural models to estimate the direct effect of adverse childhood social conditions on onset of heart disease, diabetes, and stroke. *Epidemiology*. 2012;23:223-232.
123. VanderWeele TJ. Invited commentary: structural equation models and epidemiologic analysis. *Am J Epidemiol*. 2012;176:608-612.
124. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. 2009;20.
125. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11:550-560.
126. Belsky J, Pluess M. Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychol Bull*. 2009;135:885-908.
127. Chess S, Thomas A, Rutter M, Birch H. Interaction of temperament and environment in the production of behavioral disturbance. *Am J Psychiatry*. 1963;120:142-148.
128. Aiken L, West S. *Multiple regression: Testing and interpreting interactions*. Newbury, CA: Sage; 1991.
129. Park S-Y, Belsky J, Putnam S, Crnic K. Infant emotionality, parenting, and 3-year inhibition: Exploring stability and lawful discontinuity in a male sample. *Dev Psychol*. 1997;33:218-227.
130. Stoltzfus J, Karraker K. Goodness of fit between parenting style and child temperament: Special challenges with slow to warm-up infants. In: Krause P, Dailey T, editors. *Handbook of parenting: Styles, stresses & strategies*. New York: Nova Science Publishers; 2009. p. 87-110.

131. Ellis BJ, Boyce WT, Belsky J, Bakermans-Kranenburg MJ, Van IJzendoorn MH. Differential susceptibility to the environment: an evolutionary-neurodevelopmental theory. *Dev Psychol*. 2011;23:7-28.
132. Pluess M, Belsky J. Children's differential susceptibility to effects of parenting. *Fam Sci*. 2010;1:14-25.
133. Kochanska G, Kim S. Difficult temperament moderates links between maternal responsiveness and children's compliance and behavior problems in low-income families. *J Child Psychol Psychiatry*. 2013;54:323-332.
134. Stright AD, Gallagher KC, Kelley K. Infant temperament moderates relations between maternal parenting in early childhood and children's adjustment in first grade. *Child Dev*. 2008;79:186-200.
135. VanderWeele TJ. On the distinction between interaction and effect modification. *Epidemiology*. 2009;20:863-871.
136. Knol MJ, Egger M, Scott P, Geerlings MI, Vandenbroucke JP. When one depends on the other: Reporting of interaction in case-control and cohort studies. *Epidemiology*. 2009;20:161-166.
137. Greenland S, Lash T, Rothman K. Concepts of interaction. In: Rothman K, Greenland S, Lash T, editors. *Modern epidemiology*. 3rd ed. Philadelphia: Lippincott, Williams & Wilkins; 2012. p. 71-83.
138. Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. *Int J Epidemiol*. 2012;41:514-520.
139. Carey WB, Fox M, McDevitt SC. Temperament as a factor in early school adjustment. *Pediatrics*. 1977;60:621-624.
140. Maziade M, Côté R, Boutin P, Bernier H. Temperament and intellectual development: A longitudinal study from infancy to four years. *Am J Psychiatry*. 1987;144:144-150.
141. Vaughn B, Joffe L, Bradley C, Seife R, Barglow P. Maternal characteristics measured prenatally are predictive of ratings of temperamental 'difficulty' on the Carey Infant Temperament Questionnaire. *Dev Psychol*. 1987;23:152-161.
142. Wachs TD, Gandour MJ. Temperament, environment, and six-month cognitive-intellectual development: a test of the organismic specificity hypothesis. *Int J Behav Dev*. 1983;6:135-152.
143. Field T, Hallock N, Ting G, Dempsey J, Dabiri C, Shuman HH. A first-year follow-up of high-risk infants: formulating a cumulative risk index. *Child Dev*. 1978;49:119-131.
144. Matheny A. Temperament and cognition: Relations between temperament and mental test scores. In: Kohnstamm G, Bates J, Rothbart M, editors. *Temperament in childhood*. West Sussex, England: John Wiley & Sons; 1989. p. 263-282.
145. Lawson K, Ruff H. Early attention and negative emotionality predict later cognitive and behavioural function. *Int J Behav Dev*. 2004;28:157-165.
146. Blair C, Razza RP. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev*. 2007;78:647-663.
147. Schoen MJ, Nagle RJ. Prediction of school readiness from kindergarten temperament scores. *J Sch Psychol*. 1994;32:135-147.
148. Keogh BK. Applying temperament research in school. In: Kohnstamm G, Bates J, Rothbart M, editors. *Temperament in childhood*. USA: John Wiley & Sons; 1989. p. 437-450.

149. Hughes K, Coplan R. Exploring processes linking shyness and academic achievement in childhood. *Sch Psychol Q*. 2010;25:213-222.
150. Coplan RJ, Barber AM, Lagacé-Séguin DG. The role of child temperament as a predictor of early literacy and numeracy skills in preschoolers. *Early Child Res Q*. 1999;14:537-553.
151. Molfese V, Rudasill KM, Beswick J, Jacobi-Vessels J, Ferguson M. Infant temperament, maternal personality, and parenting stress as contributors to infant developmental outcomes. *Merrill Palmer Q*. 2010;56:49-78.
152. DiLalla L, Plomin R, Fagan J*, et al.* Infant predictors of preschool and adult IQ: A study of infant twins and their parents. *Dev Psychol*. 1990;26:759-769.
153. Beck C. A meta-analysis of the relationship between postpartum depresion and infant temperament. *Nurs Res*. 1996;45:225-230.
154. Hanington L, Ramchandani P, Stein A. Parental depression and child temperament: Assessing child to parent effects in a longitudinal population study. *Infant Behav Dev*. 2010;33:88-95.
155. Murray L, Fiori-Cowley A, Hooper R. The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. *Child Dev*. 1996;67:2512-2526.
156. McBride BA, Schoppe S, Rane T. Child characteristics, parenting stress, and parental involvement: Fathers versus mothers. *J Marriage Fam*. 2002;64:998-1011.
157. Maccoby EE, Snow ME, Jacklin CN. Children's dispositions and mother–child interaction at 12 and 18 months: A short-term longitudinal study. *Dev Psychol*. 1984;20:459-472.
158. Lengua LJ. Growth in temperament and parenting as predictors of adjustment during children's transition to adolescence. *Dev Psychol*. 2006;42:819-832.
159. van den Boom DC. The influence of temperament and mothering on attachment and exploration: An experimental manipulation of sensitive responsiveness among lower-class mothers with irritable infants. *Child Dev*. 1994;65:1457-1477.
160. van den Boom DC. Do first-year intervention effects endure? Follow-up during toddlerhood of a sample of Dutch irritable infants. *Child Dev*. 1995;66:1798-1816.
161. O'Connor E, Rodriguez E, Cappella E, Morris J, McClowry S. Child distruptive behavior and parenting efficacy: a comparison of the effects of two models of insights. *J Community Psychol*. 2012;40:555-572.
162. O'Connor E, Cappella E, McCormick M, McClowry S. An examination of the efficacy of INSIGHTS in enhancing the academic and behavioral development of children in early grades. *J Educ Psychol*. 2014;106:1156-1169.
163. Sheeber LB, Johnson JH. Evaluation of a temperament-focused, parent-training program. *J Clin Child Psychol*. 1994;23:249-259.
164. Tamis-LeMonda CS, Bornstein MH, Baumwell L. Maternal responsiveness and children's achievement of language milestones. *Child Dev*. 2001;72:748.
165. Olds D, Kitzman H, Cole R*, et al.* Effects of nurse home-visiting on maternal life course and child development: Age 6 follow-up results of a randomised trial. *Pediatrics*. 2004;114:1550-1559.
166. Melhuish E, Belsky J, Heyland A, Barnes J. Effects of fully established Sure Start Local Programmes on 3-year-old children and their families living in England: A quasi-experimental observation study. *Lancet*. 2008;372:1641-1647.

167. Robling M, Bekkers M, Bell K*, et al.* Effectiveness of a nurse-led intensive home-visitation programme for the first-time teenage mothers (Buiding Blocks): A pragmatic randomised controlled trial. *Lancet*. 2015;387:146-155.
168. Poehlmann J, Hane A, Burnson C, Maleck S, Hamburger E, Shah PE. Preterm infants who are prone to distress: Differential effects of parenting on 36-month behavioral and cognitive outcomes. *J Child Psychol Psychiatry*. 2012;53:1018-1025.
169. Maupin A, Fine J. Differential effects of parenting in preterm and full-term children on developmental outcomes. *Early Hum Dev*. 2014;90:869-876.
170. Conway A, Stifter CA. Longitudinal antecedents of executive function in preschoolers. *Child Dev*. 2012;83:1022-1036.
171. Dixon W, Smith P. Links between early temperament and language acquisition. *Merrill Palmer Q*. 2000;46:417-440.
172. Liew J. Effortful control, executive functions, and education: Bringing self-regulatory and social-emotional competencies to the table. *Child Dev Perspect*. 2012;6:105-111.
173. Bergmeier H, Skouteris H, Horwood S, Hooley M, Richardson B. Associations between child temperament, maternal feeding practices and child body mass index during the preschool years: A systematic review of the literature. *Obes Rev*. 2013;15:9-18.
174. Carey W. Temperament and increased weight gains in infants. *J Dev Behav Pediatr*. 1985;6:128-131.
175. Niegel S, Ystrom E, Vollrath ME. Is difficult temperament related to overweight and rapid early weight gain in infants? A prospective cohort study. *J Dev Behav Pediatr*. 2007;28:462-466.
176. Wu T, Dixon WE, Jr., Dalton WT, 3rd, Tudiver F, Liu X. Joint effects of child temperament and maternal sensitivity on the development of childhood obesity. *Matern Child Health J*. 2011;15:469-477.
177. Pryor LE, Tremblay RE, Boivin M, et al. Developmental trajectories of body mass index in early childhood and their risk factors: An 8-year longitudinal study. *Arch Pediatr Adolesc Med*. 2011;165:906-912.
178. Slining M, Adair L, Goldman B, Borja J, Bentley M. Infant temperament contributes to early infant growth: A prospective cohort of African American infants. *Int J Behav Nutr Phys Act*. 2009;6:51.
179. Francis LA, Susman EJ. Self-regulation and rapid weight gain in children from age 3 to 12 years. *Arch Pediatr Adolesc Med*. 2009;163:297-302.
180. Vollrath ME, Tonstad S, Rothbart MK, Hampson SE. Infant temperament is associated with potentially obesogenic diet at 18 months. *Int J Pediatr Obes*. 2011;6:e408-e414.
181. Thompson AL, Adair LS, Bentley ME. Maternal characteristics and perception of temperament associated with infant TV exposure. *Pediatrics*. 2013;131:e390-e397.
182. Wasser H, Bentley M, Borja J*, et al.* Infants perceived as “fussy” are more likely to receive complementary foods before 4 months. *Pediatrics*. 2011;127:229-237.
183. Anderson SE, Gooze RA, Lemeshow S, Whitaker RC. Quality of early maternal-child relationship and risk of adolescent obesity. *Pediatrics*. 2012;129:132-140.
184. Wake M, Nicholson JM, Hardy P, Smith K. Preschooler obesity and parenting styles of mothers and fathers: Australian national population study. *Pediatrics*. 2007;120:e1520-e1527.

185. Jansen PW, Giallo R, Westrupp EM, Wake M, Nicholson JM. Bidirectional Associations Between Mothers' and Fathers' Parenting Consistency and Child BMI. *Pediatrics*. 2013;132:e1513-e1520.
186. Rollins BY, Loken E, Savage JS, Birch LL. Maternal controlling feeding practices and girls' inhibitory control interact to predict changes in BMI and eating in the absence of hunger from 5 to 7 y. *Am J Clin Nutr*. 2014;99:249-257.
187. Sprujit-Metz D, Li C, Cohen E, Birch LL, Goran M. Longitudinal influence of mothers's child feeding practices on adiposity in children. *J Peds*. 2006;148:314-320.
188. Carnell S, Wardle J. Associations between multiple measures of parental feeding and children's adiposity in United Kingdom preschoolers. *Obesity*. 2007;15:137-144.
189. Haycraft EM, Blissett JM. Maternal and paternal controlling feeding practices: Reliability and relationships with BMI. *Obesity*. 2008;16:1552-1558.
190. Baughum A, Powers S, Johnson S*, et al.* Maternal feeding pratices and beliefs and their relationships to overweight in early childhood. *J Dev Behav Pediatr*. 2001;22:391-408.
191. Campbell K, Lioret S, McNaughton S*, et al.* A parent-focused intervention to reduce infant obesity risk behaviors: A randomised trial. *Pediatrics*. 2013;131:652-660.
192. Daniels LA, Mallan KM, Nicholson JM, Battistutta D, Magarey A. Outcomes of an early feeding practices intervention to prevent childhood obesity. *Pediatrics*. 2013;132:e109-e118.
193. Rollins BY, Loken E, Savage JS, Birch LL. Effects of restriction on children's intake differ by child temperament, food reinforcement, and parent's chronic use of restriction. *Appetite*. 2014;73:31-39.
194. Anzman SL, Birch LL. Low inhibitory control and restrictive feeding practices predict weight outcomes. *J Peds*. 2009;155:651-656.
195. Chong SY, Chittleborough CR, Smithers L, Mittinty M, Lynch JW. Parenting influence on the association between temperament and IQ [Masters dissertation]. Adelaide, Australia: University of Adelaide; 2011.
196. Brownson R, Fielding J, Maylahn C. Evidence-based public health: A fundamental concept for public health practice. *Annu Rev Public Health*. 2009;30:175-201.
197. Boyd A, Golding J, Macleod J*, et al.* Cohort Profile: The 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2013;42:111-127.
198. Fraser A, Macdonald-Wallis C, Tilling K*, et al.* Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol*. 2013;42:97-110.
199. Weshsler D, Golombok S, Rust J. *Wechsler Intelligence Scale for Children – Third Edition UK Manual.* Sidcup, UK: The Psychological Corporation; 1992.
200. Cole T. Growth monitoring with the British 1990 growth reference. *Arch Dis Child*. 1997;76:47-49.
201. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*. 2000;320:1-6.
202. Carey W, McDevitt S. *Infant Temperament Questionnaire (4-8 months)*. Philadephia: Department Educational Psychology, Temple University; 1977.

203. Bould H, Joinson C, Sterne J, Araya R. The Emotionality Activity Sociability Temperament Survey: Factor analysis and temporal stability in a longitudinal cohort. *Pers Individ Dif*. 2013;54:628-633.
204. Edwards B. Growing Up in Australia: The Longitudinal Study of Australian Children: the first decade of life. *Fam Matters*. 2013;91:1-17.
205. Soloff C, Lawrence D, Johnstone R. *LSAC Technical Paper No. 1: Sample Design*. Melbourne: Australia Institute of Family Studies; 2005.
206. Sanson A, Nicholson J, Ungerer J*, et al. Introducing the Longitudinal Study of Australian Children*. Melbourne, Australia; 2002.
207. Rothman S. *Data Issues Paper No 2. Report on the Adapted PPVT-III and the Who Am I?* Melbourne, Australia: Australian Institute of Family Studies; 2005.
208. National Center for Education Statistics. *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book NCES 2002-135*. Washington, DC; 2002.
209. Rothman S. *The LSAC Academic Rating Scale Score*. Melbourne, Victoria: Australian Council for Educational Research; 2009.
210. Ogden C, Kurzmarski R, Flegal KM*, et al.* Centers for Disease Control And Prevention 2000 Growth Charts for the United States: Improvements to the 1977 National Center for Health Statistics Version. *Pediatrics*. 2002;109:45-60.
211. Prior M, Sanson A, Oberklaid F, Northam E. Measurement of temperament in one to three year old children. *Int J Behav Dev*. 1987;10:121-132.
212. Sewell J, Oberklaid F, Prior M, Sanson A, Kyrios M. Temperament in Australian toddlers. *Aust Paediatr J*. 1988;24:343-345.
213. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10:37-48.
214. Greenland S, Robins JM. Identifiability, exchangeability and confounding revisited. *Epidemiol Perspect Innov*. 2009;6:1-9.
215. Textor J, Hardt J, Knuppel S. DAGitty: A graphical tool for analyzing causal diagrams. *Epidemiology*. 2011;5:745.
216. Cole S, Hernan MA. Fallibility in estimating direct effects. *Int J Epidemiol*. 2002;31:163-165.
217. VanderWeele TJ. *Explantion in causal inference: Methods for mediation and interaction*. New York: Oxford University Press; 2015.
218. VanderWeele TJ, Hawkley L, Thisted R, Cacioppo J. A marginal structural model analysis for loneliness: Implications for intervention trials and clinical practice. *J Consult Clin Psychol*. 2011;79:225-235.
219. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168:656-664.
220. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000;11:561-570.
221. VanderWeele TJ. Bias formula for sensitivity analysis for direct and indirect effects. *Epidemiology*. 2010;21:540-551.
222. Grolnick W, Pomerantz EM. Issues and challenges in studying parental control: toward a new conceptualization. *Child Dev Perspect*. 2009;3:165-170.
223. Berge J, Wall M, Bauer K, Neumark-Sztainer D. Parenting characteristics in the home environment and adolescent overweight: A latent class analysis. *Obesity*. 2010;18:818-825.

224. Stern H, Arcus D, Kagan J, Rubin D, Snidman N. Using mixture models in temperament research. *Int J Behav Dev*. 1995;18:407-423.
225. Cleveland M, Collins LM, Lanza S, Greenberg M, Feinberg M. Does individual risk moderate the effect of contextual-level protective factors? A latent class analysis of substance use. *J Prev Interv Community*. 2010;38:213-228.
226. Collins LM, Lanza S. *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences.* Hoboken, NJ: John Wiley & Sons, Inc; 2010.
227. Lanza S, Bray B, Collins LM. An introduction to latent class and latent transition analysis. In: Weiner I, Schinka J, Velicer W, editors. *Handbook of pyshology: Research methods in psychology*. New York: John Wiley & Sons; 2012.
228. Dempster A, Laird N, Rubin D. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc*. 1977;39:1-38.
229. Goodman L. Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*. 1974;61:215-231.
230. Lanza S, X T, Bray B. Latent class analysis with distal outcomes: a flexible model-based approach. *Struct Equ Modeling*. 2013;20:1-26.
231. Lanza S, Collins LM, Lemmon D, Schafer J. PROC LCA: A SAS procedure for latent class analysis. *Struct Equ Modeling*. 2007;14:671-694.
232. Lanza ST, Dziak JJ, Huang L, Xu S, Collins LM. Proc LCA & Proc LTA user's guide (version 1.2.7). In: University Park: The Methodology Center, Penn State; 2011.
233. VanderWeele TJ, Knol MJ. A tutorial on interaction. *Epidemiol Methods*. 2014;3:1-40.
234. Spratt M, Carpenter J, Sterne JA*, et al.* Strategies for multiple imputation in longitudinal studies. *Am J Epidemiol*. 2010;172:478-487.
235. Rubin D. Inference and missing data. *Biometrika*. 1976;63:581-592.
236. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30:377-399.
237. Sterne JA, White IR, Carlin JB*, et al.* Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ*. 2009;338:b2393.
238. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: What is it and how does it work? *Int J Methods Psychiatr Res*. 2011;20:40-49.
239. Lee KJ, Carlin JB. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *Am J Epidemiol*. 2010;171:624-632.
240. Royston P. Multiple imputation of missing values. *Stata J*. 2004;4:227-241.
241. van Buuren S, Boshuizen H, Knook D. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med*. 1999;18:681-684.
242. von Hippel PT. Regression with missing ys: An improved strategy for analyzing multiply imputed data. *Sociol Methodol*. 2007;37:83-117.
243. Sullivan T, Salter A, Ryan P, Lee K. Bias and precision of the "multiple imputation, the deletion" method for dealing with missing outcome data. *Am J Epidemiol*. 2014:528-534.
244. Chong SY, Chittleborough CR, Gregory T, Lynch JW, Smithers LG. How many infants are temperamentally difficult? Comparing norms from the Revised Infant Temperament Questionnaire to a population sample of UK infants. *Infant Behav Dev*. 2015;40:20-28.

245. Younger E. Human growth and development. In: Silbert-Flagg J, Sloand E, editors. *Pediatric nurse practitioner certification: Review guide*. 5th ed. United States of America: Kevin Sullivan; 2011. p. 7-36.
246. Hertzig M, Snow M. The assessment of temperament. In: Kestenbaum C, Williams D, editors. *Handbook of clinical assessment of children and adolescents*. New York: New York Press; 1988. p. 133-153.
247. Stein M, Plonsky C, Zuckerman B, Carey W. Reformatting the 9-month health supervisionn visit to enhance developmental, behavioral and family concerns. *J Dev Behav Pediatr*. 2005;26:56-60.
248. Carey W, McDevitt S. *Child behavioral assessment & management in primary care*. Scottsdale, AZ: Behavioral-Developmental Initiatives; 2012.
249. Andersen CJ. *The temperament guides: Vol. I. Resources for early intervention professionals*. Scottsdale, AZ: Behavioral–Developmental Initiatives; 2000.
250. Andersen CJ. *The temperament guides: Vol. II. Resources for early intervention professionals*. Scottsdale, AZ: Behavioral-Developmental Initiatives; 2002.
251. Jellinek M, Patel B, Froechle M. *Bright Futures in Practice: Mental Health—Volume I. Practice Guide*. Arlington, VA: National Center for Education in Maternal and Child Health; 2002.
252. Kurcinka M. *Raising your spirited child*. New York: Harper; 1998.
253. Neville H, Williams J. *Is this a phase? Child development & parent strategies, birth to 6 years*. United States of America: Parenting press, Inc; 2007.
254. Pickett KE, Wood C, Adamson J, DeSouza L, Wakschlag LS. Meaningful differences in maternal smoking behaviour during pregnancy: implications for infant behavioural vulnerability. *J Epidemiol Community Health*. 2008;62:318-324.
255. Bradley RH, Corwyn RF. Infant temperament, parenting, and externalizing behavior in first grade: a test of the differential susceptibility hypothesis. *J Child Psychol Psychiatry*. 2008;49:124-131.
256. Martin RP, Wisenbaker J, Baker J, Huttunen MO. Gender differences in temperament at six months and five years. *Infant Behav Dev*. 1997;20:339-347.
257. Sayal K, Heron J, Maughan B, Rowe R, Ramchandani P. Infant temperament and childhood pyshicatric disorder: Longitudinal study. *Child Care Health Dev*. 2013;40:292-297.
258. Lewis AJ, Olsson CA. Early life stress and child temperament style as predictors of childhood anxiety and depressive symptoms: Findings from the Longitudinal Study of Australian Children. *Depress Res Treat*. 2011;2011:1-9.
259. Taylor CL, Christensen D, Lawrence D, Mitrou F, Zubrick SR. Risk factors for children's receptive vocabulary development from four to eight years in the Longitudinal Study of Australian Children. *PLoS One*. 2013;8:e73046.
260. Saudino KJ, Plomin R. Why are hyperactivity and academic achievement related? *Child Dev*. 2007;78:972-986.
261. Sasaki Y, Mizuno R, Kaneko H, Murase S, Honjo S. Application of the Revised Infant Temperament Questionnaire for evaluating temperament in the Japanese infant: creation of an abridged Japanese version. *Psychiatry Clin Neurosci*. 2006;60:9-17.
262. McClowry SG, Rodriguez ET, Koslowitz R. Temperament-based intervention: Re-examining goodness of fit. *Eur J Dev Sci*. 2008;2:120-135.
263. McClowry S, Snow D, Tamis-LeMonda C. An evaluation of the effects of INSIGHTS on the behavior of inner city primary school children. *J Prim Prev*. 2005;26:567-584.

264. Lee CL, Bates JE. Mother-child interaction at age two years and perceived difficult temperament. *Child Dev*. 1985;56:1314-1325.
265. Greenberg MT, Kusche C, Cook E, Quama J. Promoting emotional competence in school-aged children: The effects of the PATHS curriculum. *Dev Psychopathol*. 1995;7:117-136.
266. Dixon WE, Smith P. Who's controlling whom? Infant contributors to maternal play behaviour. *Infant Child Dev*. 2003;12:177-195.
267. Valiente C, Lemery-Chalfant K, Swanson J. Prediction of kindergartners' academic achievement from their effortful control and emotionality: Evidence for direct and moderated relations. *J Educ Psychol*. 2010;102:550-560.
268. Rudasill KM, Gallagher KC, White JM. Temperamental attention and activity, classroom emotional support, and academic achievement in third grade. *J Sch Psychol*. 2010;48:113-134.
269. Cunha F, Heckman J. The technology of skill formation. *Am Econ Rev*. 2007;97:31-47.
270. Wechsler D. *Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV)*. San Antonio, TX, USA: Psychological Corporation 2003.
271. Dunn L, Dunn L. *Peabody Picture Vocabulary Test, 3rd ed.*: American Guidance Service, Circle Pines, MN; 1997.
272. Rock D, Pollack J. *Early Childhood Longitudinal Study-Kindergarten Class of 1998–99 (ECLS–K), Psychometric Report for Kindergarten Through First Grade*. Washington, DC: US Dept of Education, National Center for Education Statistics; 2002.
273. United Nation's Children's Fund. Multiple Indicator Cluster Survey (MICS). http://www.unicef.org/statistics/index_24302.html. Accessed May 27, 2015.
274. Karrass J, Braungart-Rieker JM. Infant negative emotionality and attachment: Implications for preschool intelligence. *Int J Behav Dev*. 2004;28:221-229.
275. Posner MI, Rothbart MK. Developing mechanisms of self-regulation. *Dev Psychopathol*. 2000;12:427-441.
276. Prior M, Bavin E, Cini E*, et al.* Influences on communicative development at 24 months of age: Child temperament, behaviour problems, and maternal factors. *Infant Behav Dev*. 2008;31:270-279.
277. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992;3:143-155.
278. Daniel R, De Stavola B, Cousens S. gformula: Estimating causal effects in the presence of time-varying confounding or mediation using g-computation formula. *Stata J*. 2011:1-30.
279. Hernan MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health*. 2004;58:265-271.
280. Naimi A, Moodie EE, Auger N, Kaufman J. Constructing inverse probability weights for continous exposures: A comparison of methods. *Epidemiology*. 2014;25:292-299.
281. Dobbins T, Sullivan E, Roberts C, Simpson J. Australian national birthweight percentiles by sex and gestational age, 1998–2007. *Med J Aust*. 2012;197:291-294.
282. Kessler RC, Green JG, Gruber MJ*, et al.* Screening for serious mental illness in the general population with the K6 screening scale: Results from the WHO World Mental Health (WMH) survey initiative. *Int J Methods Psychiatr Res*. 2010;19:4-22.

283. Chong SY, Chittleborough CR, Gregory T, Lynch JW, Mittinty M, Smithers L. Parenting practices at 24 to 47 months and IQ at age 8: effect-measure modification by infant temperament. *PLoS One*. 2015;11:e0152452.
284. Hanushek EA, Woessmann L. The role of cognitive skills in economic development. *J Econ Lit*. 2008;46:607-668.
285. Vygotsky L. *Thought and language*. Cambridge, MA: MIT Press; 1962.
286. Lugo-Gil J, Tamis-LeMonda CS. Family resources and parenting quality: Links to children's cognitive development across the first 3 years. *Child Dev*. 2008;79:1065-1085.
287. Grolnick WS, Gurland ST, DeCourcey W, Jacob K. Antecedents and consequences of mothers' autonomy support: An experimental investigation. *Dev Psychol*. 2002;38:143-155.
288. Canals J, Hernandez-Martinez C, Esparo G, Fernandez-Ballart J. Neonatal Behavioral Assessment Scale as a predictor of cognitive development and IQ in full-term infants: A 6-year longitudinal study. *Acta Paediatr*. 2011;100:1331-1337.
289. Sims DM, Lonigan CJ. Inattention, hyperactivity, and emergent literacy: Different facets of inattention relate uniquely to preschoolers' reading-related skills. *J Clin Child Adolesc Psychol*. 2013;42:208-219.
290. McClelland MM, Wanless SB. Growing up with assets and risks: The importance of self-regulation for academic achievement. *Res Hum Dev*. 2012;9:278-297.
291. Carey WB. Futher thoughts on temperament. *J Dev Behav Pediatr*. 1986;7:120-123.
292. Gardner J, Walker S, Powell C, Grantham-McGregor S. A randomised controlled trial of a home-visiting intervention on cognition and behavior in term low birth weight infants. *J Pediatr*. 2003;143:634-639.
293. McClowry S, Collins A. Temperament-based interventions. In: Zentner M, Shiner R, editors. *Handbook of temperament*. New York: Guilford Press; 2012. p. 607-670.
294. Thompson W. Effect modification and the limits of biological inference from epidemiologic data. *J Clin Epidemiol*. 1991;44:221-232.
295. Waylen A, Stewart-Brown S. *Parenting in ordinary families: Diversity, complexity and change*. York, UK: Joseph Rowntree Foundation; 2008.
296. Bowen E, Heron J, Waylen A, Wolke D, the ALSPAC study team. Domestic violence risk during and after pregnancy: Findings from a British longitudinal study. *BJOG*. 2005;112:1083-1089.
297. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry*. 1987;150:782-786.
298. Murray L, Carothers AD. The validation of the Edinburgh Post-natal Depression Scale on a community sample. *Br J Psychiatry*. 1990;157:288-290.
299. Hosmer D, Lemeshow S. Confidence interval estimation of interaction. *Epidemiology*. 1992;3:452-456.
300. Estrada P, Arsenio WF, Hess RD, Holloway SD. Affective quality of the mother–child relationship: Longitudinal consequences for children's school-relevant cognitive functioning. *Dev Psychol*. 1987;23:210-215.
301. Belsky J, Bakermans-Kranenburg MJ, Van IJzendoorn MH. For better and for worse: Differential susceptibility to environmental influence. *Curr Dir Psychol Sci*. 2007;16:300-304.

302. van Zeijl J, Mesman J, Stolk M*, et al.* Differential susceptibility to discipline: The moderating effect of child temperament on the association between maternal discipline and early childhood externalizing problems. *J Fam Psychol*. 2007;21:626-636.
303. Bates J, Pettit G, Dodge K, Ridge B. Interaction of temperamental resistance to control and restrictive parenting in the development of externalizing behavior. *Dev Psychol*. 1998;34:982-995.
304. Rothman K, Greenland S, Walker A. Concepts of interaction. *Am J Epidemiol*. 1980;112:467-470.
305. Ahlbom A, Alfredson L. Interaction: A word with two meanings create confusion. *Eur J Epidemiol*. 2005;20:563-564.
306. Siemiatycki J, Thomas D. Biological models and statistical interactions: An example from multistage carcinogenesis. *Int J Epidemiol*. 1981;10:383-387.
307. Bergmeier H, Skouteris H, Horwood S, Hooley M, Richardson B. Child temperament and maternal predictors of preschool children's eating and body mass index. A prospective study. *Appetite*. 2014;74:125-132.
308. Anzman SL, Stifter CA, Birch H. Temperament and childhood obesity risk: A review of the literature. *J Dev Behav Pediatr*. 2012;33:732-745.
309. Cole T, Freeman J, Preece M. Body Mass Index reference curves for the UK, 1990. *Arch Dis Child*. 1995;73:25-29.
310. Vollrath M, Stene-Larsen K, Tonstad S, Rothbart M, Hampson S. Association between temperament age 1.5 years and obesogenic diet at ages 3 and 7 years. *J Dev Behav Pediatr*. 2012;33:721-727.
311. Haycraft E, Farrow C, Meyer C, Powell F, Blissett J. Relationships between temperament and eating behaviours in young children. *Appetite*. 2011;56:689-692.
312. Faith MS, Hittner JB. Infant temperament and eating style predict change in standardized weight status and obesity risk at 6 years of age. *Int J Obes*. 2010;34:1515-1523.
313. Graziano PA, Calkins SD, Keane SP. Toddler self-regulation skills predict risk for pediatric obesity. *Int J Obes*. 2010;34:633-641.
314. Collins C, Duncanson K, Burrows T. A systematic review investigating associations between parenting style and child feeding behaviours. *J Hum Nutr Diet*. 2014;27:557-568.
315. Vollmer R, Mobley A. Parenting styles, feeding styles, and their influence on child obesogenic behaviors and body weight. A review. *Appetite*. 2013;71:232-241.
316. Ventura A, Birch L. Does parenting affect children's eating and weight status? *Int J Behav Nutr Phys Act*. 2008;5:1-12.
317. Sleddens EF, Gerards SM, Thijs C, de Vries NK, Kremers SP. General parenting, childhood overweight and obesity-inducing behaviors: A review. *Int J Pediatr Obes*. 2011;6:e12-27.
318. Hurley K, Cross M, Hughes S. A systematic review of responsive feeding and child obesity in high-income countries. *J Nutr*. 2011;141:495-501.
319. Shloim N, Edelson L, Martin N, Hetherington M. Parenting styles, feeding styles, feeding practices, and weight status in 4-12 year-old children: A systematic review of the literature. *Front Psychol*. 2015;6:1-20.
320. Fisher J, Birch L. Restricting access to foods and children's eating. *Appetite*. 1999;32:405-419.

321. Birch L, Fisher J. Development of eating behaviors among children and adolescents. *Pediatrics*. 1998;101:539-549.
322. Farrow CV, Blissett J. Controlling feeding practices: A cause or consequence of early child weight? *Pediatrics*. 2008;121:e164-e169.
323. Wardle J, Sanderson S, Guthrie CA, Rapoport L, Plomin R. Parental feeding style and the inter-generational transmission of obesity risk. *Obes Res*. 2002;10:453-462.
324. Anzman-Frasca S, Liu S, Gates K, Paul IM, Rovine M, Birch L. Infants' transitions out of a fussing/crying state are modifiable and are related to weight status. *Infancy*. 2013:662-686.
325. Patrick H, Hennessy E, McSpadden K, Oh A. Parenting styles and practices in children obesogenic behaviors: Scientific gaps and future research directions. *Child Obes*. 2013;9:S73-S86.
326. Northstone K, Emmett P. Are dietary patterns stable throughout early and mid-childhood? A birth cohort study. *Br J Nutr*. 2008;100:1069-1076.
327. Faith MS, Berkowitz RI, Stallings VA, Kerns J, Storey M, Stunkard AJ. Parental feeding attitudes and styles and child body mass index: Prospective analysis of a gene-environment interaction. *Pediatrics*. 2004;114:e429-e436.
328. Birch L, Fisher J, Grimm-Thomas K, Markey CN, Sawyer R, Johnson S. Confirmatory factor analysis of the Child Feeding Questionnaire: A measure of parental attitudes, beliefs and practices about child feeding and obesity proness. *Appetite*. 2001;36:201-210.
329. Faith MS, Heshka S, Keller KL. Maternal-child feeding patterns and child body weight: Findings from a population-based sample. *Arch Pediatr Adolesc Med*. 2003;157:926-932.
330. Hughes S, Power T, Fisher J, Mueller S, Nicklas TA. Revisiting a neglected construct: parenting styles in a child-feeding context. *Appetite*. 2005:83-92.
331. Musher-Eizenman D, Holub S. Comprehensive Feeding Practices Questionnaire: Validation of a new measure of parental feeding practices. *J Pediatr Psychol*. 2007;32:960-972.
332. Patrick H, Nicklas TA, Hughes SO, Morales M. The benefits of authoritative feeding style: caregiver feeding styles and children's food consumption patterns. *Appetite*. 2005;44:243-249.
333. McMeekin S, Jansen E, Mallan K, Nicholson J, Magarey A, Daniels L. Associations between infant temperament and early feeding practices. A cross-sectional study of Australian mother-infant dyads from the NOURISH randomised controlled trial. *Appetite*. 2013;60:239-245.
334. Fangupo L, Heath A, Williams S*, et al.* Impact of an early-life intervention on the nutrition behaviors of 2-y-old children: A randomized controlled trial. *Am J Clin Nutr*. 2015;102:704-712.
335. Freedman DS, Wang J, Maynard LM*, et al.* Relation of BMI to fat and fat-free mass among children and adolescents. *Int J Obes*. 2005;29:1-8.
336. Office of Population Censuses and Surveys. *Standard occupational classification*. London: Her Majesty's Stationery Office; 1991.
337. Freeman J, Cole T, Chinn S, Jones P, White E, Preece M. Cross sectional stature and weight reference curves for the UK, 1990. *Arch Dis Child*. 1995;73:17-24.

338. Lawlor D, Benfield L, Logue J*, et al.* Association between general and central adiposity in childhood, and change in these, with cardiovascular risk factors in adolescence: prospective cohort study. *BMJ*. 2010;341:c6224.
339. Israel AC, Guile CA, Baker JE, Silverman WK. An evaluation of enhanced self-regulation training in the treatment of childhood obesity. *J Pediatr Psychol*. 1994;19:737-749.
340. Braet C, Van Winckel M. Long-term follow-up of a cognitive behavioral treatment program for obese children. *Behav Ther*. 2000;31:55-74.
341. Fisher JO, Birch L. Restricting access to palatable foods affects response, food selection, and intake. *Am J Clin Nutr*. 1999;69:1264-1272.
342. Wells JCK, Stanley M, Laidlaw AS, Day JME, Stafford M, Davies PSW. Investigation of the relationship between infant temperament and later body composition. *Int J Obes Relat Metab Disord*. 1997;21:400-406.
343. Davey Smith G. Epidemiology-is it time to call it a day? *Int J Epidemiol*. 2001;30:1-11.
344. Kaufman J, Cooper R, McGee D. Socioeconomic status and health in Blacks and Whites: The problem of residual confounding and the resiliency of race. *Epidemiology*. 1997;8:621-628.
345. Button K, Ioannidis J, Mokrysz C*, et al.* Power failure: Why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013;14:365-376.
346. Ioannidis J. Why most published research findings are false. *PLoS Med*. 2005;2:e124.
347. Kühberger A, Fritz A, Scherndl T. Publication bias in psychology: A diagnosis based on the correlation between effect size and sample size. *PLoS One*. 2014;9:e105825.
348. Franco A, Malhotra N, Simonovits G. Publication bias in the social sciences: Unlocking the file drawer. *Science*. 2014;345:1502-1505.
349. Baker M. Over half of psychology studies fail reproducibility test. http://www.nature.com/news/over-half-of-psychology-studies-fail-reproducibility-test-1.18248. Published 2015. Accessed 1st September 2015.
350. Darlington AS, Wright CM. The influence of temperament on weight gain in early infancy. *J Dev Behav Pediatr*. 2006;27:329-335.
351. Lengua LJ, Kovacs EA. Bidirectional associations between temperament and parenting and the prediction of adjustment problems in middle childhood. *J Appl Dev Psychol*. 2005;26:21-38.
352. Carey WB, McDevitt SC. Commentary on paper by Chong et al. "How many infants are temperamentally difficult?"40 (2015) 20–28. *Infant Behav Dev*. 2015;41:167-168.
353. Super CM, Axia G, Harkness S*, et al.* Culture, temperament, and the "difficult child": A study in seven western cultures. *Eur J Dev Sci*. 2008;2:136-157.