Irritability : A Study of Its Origin, Nature and Role In Relation to Disorder

Kathryn J. Hudson

Thesis submitted to Cardiff University for the degree of Doctor of Philosophy

2010

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Contents

CHAPTER 1
GENERAL INTRODUCTION

PAGE

CHAPTER 2
Irritability: The Measurement of a Dimension of Infant Temperament, a Symptom of Childhood Disorders, and the Relationship between Them

CHAPTER 3
STUDY 1
Irritability and Childhood Disorder Symptoms: Testing the Measurement Confounding Hypothesis

CHAPTER 4
STUDY 2
Understanding the Role of Irritability in Adult Disorders, and the Potential Intergenerational Transmission of Irritability from Mother to Infant

CHAPTER 5
General Discussion

Index of tables

Index of figures

Figure 5.1

The Proposed Cycle of Irritability in Relation to Disorder and Generation across the Lifespan (page 204).

Dedication

This thesis is dedicated to the children and their families from both the Starting School Study and the Cardiff Child Development Study. This work would not have been possible without their commitment to this research.

Acknowledgements

I wish to sincerely thank my supervisor, Professor Dale Hay, for inspiring me in both her teaching and research, and for guiding and supporting me throughout my journey to complete this thesis. I would also like to thank Professor Stephanie VanGoozen for her time and support given to me during my studies, and to Professor Rob Honey for being calm and supportive. To my colleagues on the Cardiff Child Development Study team, I would like to express my gratitude for all the efforts and support that have gone into making the CCDS a special study that will always remain a part of us. Particularly, I thank Cerith, Lisa, Siwan, Sarah-Louise, Naomi, Cerian, Jane, Becca, Oliver, and Bryony.

I owe a great deal of thanks to my wonderful husband, Bob, who has encouraged me and supported me with love and care over these years, and who together with my beautiful family, Craig, Oliver and Gabrielle, have allowed me to realise my dream. Oliver – you have challenged me with your intellect, Gabrielle – you have been wonderfully insightful, and Craig – you inspire calm. I also thank my mum, for being there to allow me to work and know that my children are in safe hands, and my dad for giving me an exemplary childhood. Finally, thank you to all my friends and family who have been very patient and supportive towards me. In particular, I thank Gail, my friend and inspiration.

This work was supported by a Medical Research Council Grant, number GO400086.

Summary

This thesis explores the origin, nature and role of irritability and disorder across the psychology and psychiatry literature. Within two empirical studies, irritability was examined at different stages of the lifespan, at two transition points. Study 1, the Starting School Study, explored irritability in preschool-aged children, in relation to clinical symptoms of disorder. The measurement confounding hypothesis was tested for the relationship between irritability and internalising and externalising symptoms. Whilst some measurement confounding was found between irritability and symptoms of Oppositional Defiant Disorder and internalising symptoms (depression and anxiety), irritability remained significantly associated with the pure scales for both ODD and internalising symptoms. Irritability mediated the relationship between internalising and externalising symptoms, suggesting that irritability plays a role in co morbidity. In Study 2, a parallel investigation was carried out into the role of irritability and disorder in adult women at their transition to motherhood. Mothers' irritability predicted both her conduct symptoms and emotional disorders. The mother-infant subsystem was used as the focus for exploring the potential influence of mothers' characteristics and mental health on the infant's irritability. Mothers' irritability predicted infant irritability at 6 months, when mothers' mental health was taken into account. Additionally, mothers' irritability after childbirth mediated the relationship between mothers' antenatal irritability and infant irritability, suggesting an intergenerational transmission of irritability between mother and infant by 6 months.

The findings from these two empirical studies serve to inform the psychology and psychiatry literature about the need to define temperament constructs within studies and

to assess for potential confound items across measures. The importance of irritability in relation to emotional and behavioural problems at different points of the lifespan, and the potential for intergenerational transmission of irritability from mother to child, suggests that irritability could be an early indicator for possible intervention to prevent long-term disorders.

CHAPTER 1

General Introduction

1.1. *The Focus of the Thesis*

Post millennium, there has been a high degree of concern about the level of anger and aggression amongst children and young people in the UK. This is borne out in the media coverage and statistics that are oft quoted. In the 2006 *Offending, Crime and Justice Survey* 12% of 10- to 25-year-olds reported committing an assault in the past 12 months, and 17% of those committing assault reported doing so six or seven times in the last 12 months (Home Office, 2008). In schools in England, police officers in 29 out of 35 forces reported being called out over 7000 times in one year to incidents at schools (*Guardian*, 22nd December, 2008). The Association of Teachers and Lecturers surveyed 800 members across the UK and found that two thirds believed standards of behaviour were getting worse, with 30% of teachers and lecturers reporting physical aggression from pupils, and 10% reporting physical harm as a result (*Guardian*, 17th March, 2008). These newspaper reports and statistics lead politicians and academics to debate the causes and potential solutions to control such behaviours.

This recent UK concern for the behaviour of children and young people is not a new UK-only phenomenon, as illustrated by a Report to the President of the USA in 1971, in a White House Conference on Children (Bronfenbrenner, 1972). The conference considered evidence on the increasing alienation of children and youth, and the rising levels of juvenile drug abuse, delinquency, and violence. There appears to be little

difference between the 1971 USA experience and today's UK experience. What is surprising is the observation of an influential researcher comparing the two worlds of childhood in the USA and the USSR at that time,

"It is noteworthy that, of all the countries in which my colleagues and I are working, ..., the only one which exceeds the United States in the willingness of children to engage in antisocial behaviour is the nation closest to us in our Anglo-Saxon traditions of individualism. That country is England (UK)..... The only country in our sample which shows a level of parental involvement lower than our own." (p. 116. Bronfenbrenner, 1972).

The implication from the recent media reports and statistics is that in the UK angry behaviours of children and young people have continued from the 1970's to the present day and, evidently, these behaviours have increased across that time period. Using psychopathology criteria to assess the prevalence of antisocial children in the UK, a study of three British cohorts of 15-16 year olds across three time periods found increased reports from parents between 1974 and 1986, and between 1986 and 1999 (Collishaw, Maughan, Goodman, & Pickles, 2004), but little increase since 1999 (Green, McGinnity, Ford, & Goodman, 2005).

Understanding why some children are quick and/or prone to angry responses is the primary aim of this thesis. Irritability is the term used to describe "quickness and/or proneness to anger," and is a construct of interest to psychiatrists and psychologists, due to its dual role as both a main symptom of emotional and behavioural disorders in children, and as a psychological risk factor in the development of such disorders. Teasing apart this dual role of irritability is the second aim of this thesis.

Investigation of the dual role of irritability as a construct that straddles the different domains of psychology and psychiatry requires a research paradigm that reflects this interdisciplinary nature of studies on irritability. I have therefore chosen 'Developmental Psychopathology' as a research framework for this thesis, as it provides a broad church for disciplines to contribute to the knowledge about pathways to disorder. Such pathways in children are complex and involve multiple influences from biological, social and psychological factors. Studying an element of the complex pathways within this framework brings us closer to finding ways we may prevent children from suffering further. Within this chapter, I outline the research paradigm I have chosen, and the format of the thesis.

In Chapter 2, I compare the psychiatric evidence for the role of irritability as a symptom of childhood disorder, with the evidence for irritability as a temperament construct in the psychological literature. Unfortunately, irritability has not escaped one of the pitfalls of temperament studies, and has been studied under different descriptors within different contemporary approaches to temperament (Goldsmith, Buss, Plomin, Rothbart, Thomas, Chess, Hinde, & McCall, 1987). I have mapped out these approaches in Chapter 2, drawing together the research on the nature of irritability to determine the appropriate measure of irritability for use within this thesis, and to direct the focus of potential predictors of irritability in children.

Irritability has been a construct of interest for 2000 years, since the Greek physician Galen (130-201AD) suggested that an excess of yellow bile made a person irritable. Defining irritability for the purpose of this thesis requires acknowledgement of

its use across the different disciplines, psychiatry and psychology. Within psychiatry, irritability is a symptom of disorder, which suggests that it may be episodic and temporary. Within psychology, irritability is considered as an enduring behavioural response on a temperament dimension. Irritability exists within both contexts and is defined for this thesis as 'an episode and/or an enduring behaviour characterised by reduced control over temper which usually results in irascible verbal or behavioural outbursts' (adapted from Snaith & Taylor, 1985).

As an important factor in the development of childhood disorders, there is a need to understand the early signs of irritability in childhood starting in infancy, and its relationship with other known predictors of childhood disorder. Knowing that a combination of risk factors can significantly increase the likelihood of childhood disorder (Rutter, Tizard, Yule, Graham, & Whitmore, 1976), I would anticipate that there are potentially irritable infants for whom a combination of early risk factors could be identified and interventions made to prevent possible future disorder.

Infants are not born into a vacuum but into a complex system of relationships, and therefore I draw upon the ecological perspective proposed by Bronfenbrenner (1977). This ecological framework sits well with developmental psychopathology, as it lays out the detailed set of interacting systems that a child will develop within. Bronfenbrenner's ecological framework includes the individual as a dynamic organism set within the system of the family, school, parents' world of work and friends, and the social and cultural system. Within this thesis I aim to examine the mother-infant subsystem of the family system to explore the origin of infant irritability. In keeping with the developmental psychopathology perspective, I examine the infants' degree of irritability

in relation to potential maternal factors, both before and after the infant is born, that may add to risk for future childhood disorder. In doing this I aim to explore irritability in infants and why this occurs.

1.2. *Developmental Psychopathology – The Research Paradigm*

Developmental psychopathology uses a developmental framework that is based on systems thinking and organisation, to understand the adaptation of the individual across the lifespan. The research paradigm focuses on the interplay between normal and atypical development across diverse domains of functioning and as such, provides a broad perspective that encourages contributions from many different disciplines to help unravel the complex processes leading to adjustment and maladjustment. Developmental psychopathologists hence investigate functioning through the assessment of ontogenetic, biochemical, genetic, biological, physiological, cognitive, social-cognitive, representational, socioemotional, environmental, cultural, and societal influences on behaviour (Cicchetti & Cohen, 1995).

The developmental part of the paradigm follows the system organisational model (von Bertalanffy, 1968), suggesting that the individual at each stage of development is faced with new challenges to which they must adapt. Adaptation involves reorganisation within the individual that may result in the achievement of competence within and among the various sub-systems of the individual, i.e. emotional, cognitive, social, representational and biological systems. Maladaptation may also occur, and subsequent to adaptation or maladaptation these reorganised structures are incorporated into later structures in a successive process of hierarchical integration. The future consequence for development within the individual is that early competence tends to promote later

competence and, conversely, maladaptation may lead to later problems. The process is not necessarily deterministic, as illustrated by Sroufe's (1997) useful tree metaphor of the process of adaptation and maladaptation in development. This metaphor originates from the phylogenic tree metaphor used in Darwin's theory of evolution. An adapted illustration of the tree metaphor of development is shown in Figure 1.1 (reproduced with kind permission of A. Sroufe, personal communication, August 12, 2009). Normal development can be viewed as continuous growth at or near the main body or trunk of the tree. Pathways involving large groups of individuals are represented as branches diverting only slightly from the tree trunk and reflecting approximations of normality. The different branches represent multifinality, in which one component may function differently resulting in different outcomes (von Bertalanffy, 1968). Abnormality is seen as a succession of branches away from the main trunk, with the further away the branch the greater the deviation from common pathways. Some branches may after an initial divergence grow closer to the trunk with secondary branches. This indicates the potential for individuals to reach a common outcome following different pathways, equifinality (vonBertalanffy, 1968). Each nodal point of the tree is considered as a developmental transition. Within this system model of development, the notion of multiple pathways in development is known as 'developmental pluralism'. The process of development is dynamic and may involve numerous pathways with many starts and stops and changes of direction towards disorder or competence as the child gets older.

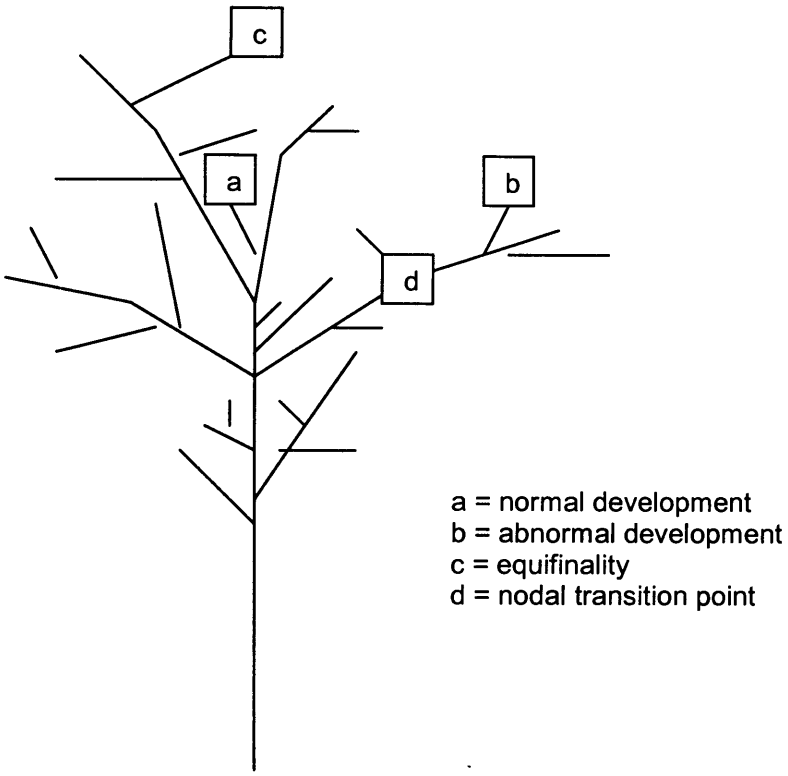


Figure 1.1. An illustration of Sroufe's tree metaphor for development.

The psychopathology part of the research paradigm may be thought of as a distortion or disturbance or degeneration of normal functioning (Cicchetti & Cohen, 1995). Developmental psychopathologists are interested in the whole continuum of functioning from normal to abnormal and associated pathways, and therefore research within this paradigm is usually carried out on high-risk and disordered populations, and on populations at risk who do not develop disorder. Developmental psychopathologists are interested in understanding the mechanisms and processes that moderate the outcome of risk factors based on the hypothesis that psychopathology will result from the continual process of inter-relations between the individual and the environment. The risk factors may come from a multitude of domains and interactions between the individual

and their environment. Bronfenbrenner's ecological model of development provides an insight into the influences and interactions that individuals may have with their environment, and as such maps out the methodological issues for consideration within this thesis.

1.3. *Bronfenbrenner's Ecological Model of Human Development*

Born out of concern that research on development had been moving in divergent ways towards either naturalistic or experimental approaches that were not to the benefit of scientific progress, Bronfenbrenner (1977) proposed a broader approach to research in human development which he called the ecology of human development, thus defined,

'The ecology of human development is the scientific study of the progressive mutual accommodation, throughout the lifespan, between a growing human organism and the changing immediate environments in which it lives, as this process is affected by relations obtaining within and between these immediate settings, as well as the larger social contexts, both formal and informal, in which the settings are embedded.' (Bronfenbrenner, 1977, p. 514). An adapted illustration of the model is shown in Figure 1.2.

This model was based on systems theory, with the relation between person and environment conceived in systems terms. The ecological environment was proposed as a nested arrangement of structures each contained within the next, a hierarchy of systems, microsystems, mesosystems, exosystems, and macrosystems, illustrated in Figure 1.2 as concentric circles with the child at the centre. A microsystem is the complex of relations between the developing person and environment in an immediate setting containing that person (e.g. home or school). The complex interactions in the microsystem are illustrated

with the two-way arrows between child and family. Within the model, a setting is defined as a place with particular physical features in which the participants engage in particular activities in particular roles for particular periods of time. A mesosystem comprises the interrelations among major settings containing the developing person at a particular point in the individual's development, e.g. the interrelations between home and school, again illustrated in Figure 1.2 using the two-way arrows. The exosystem is an extension of the mesosystem and includes settings in which the child does not participate, but which impinge upon the immediate settings in which the child is found e.g. parent's work, informal social networks. The macrosystem differs fundamentally from the previous systems as it does not refer to specific contexts that affect the life of the developing person but to general prototypes in the culture that set the pattern of structures and activities, such as the economic, social, and political systems of which the microsystems, mesosystems, and exosystems are concrete expressions. Macrosystems are examined not just in structural terms but also as carriers of information and ideology, such as the place or priority of children in a culture, and will determine how children and their carers are treated in society.

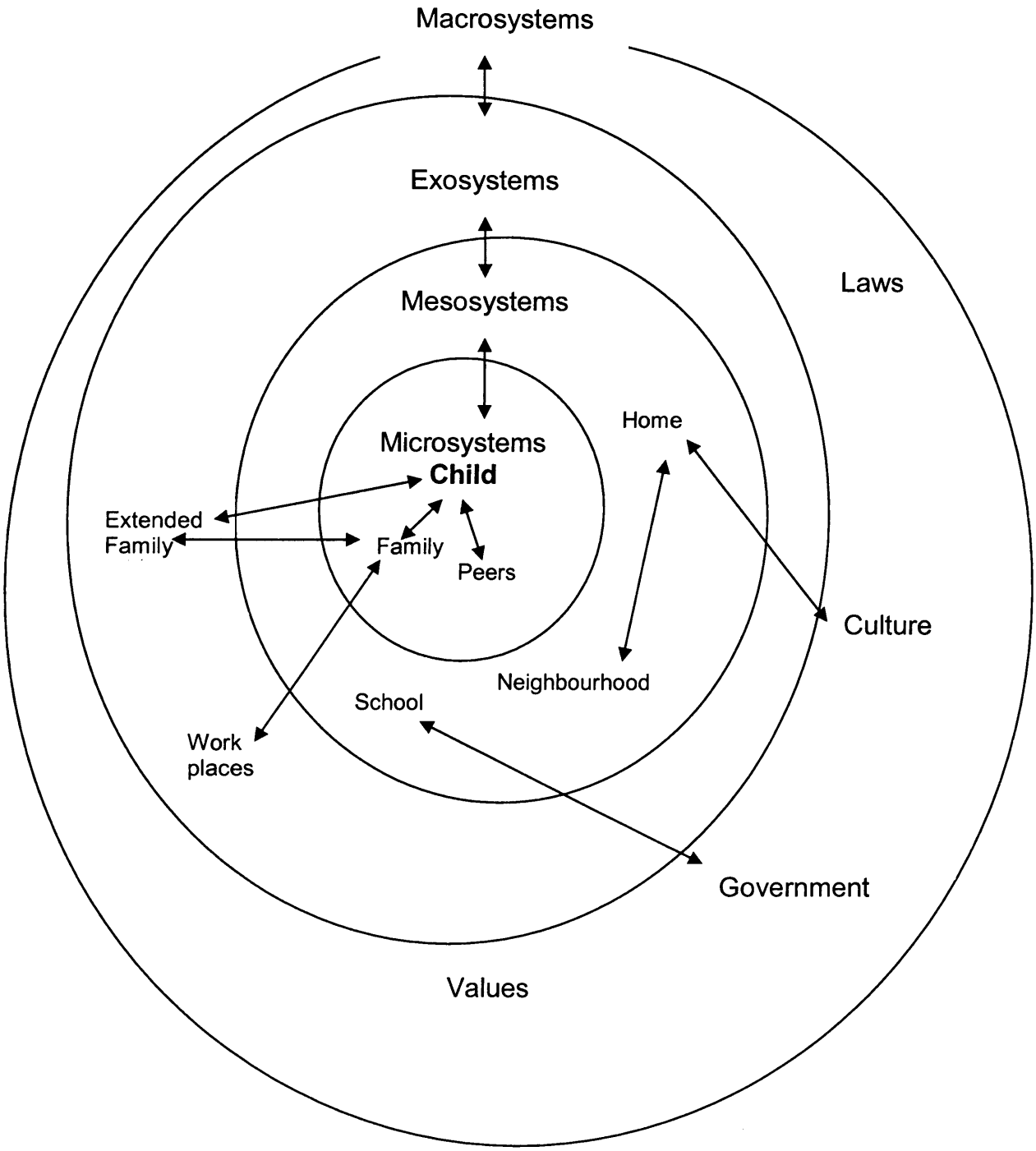


Figure 1.2. Bronfenbrenner's ecological model of development. Adapted from Bronfenbrenner (1977).

Bronfenbrenner's ecological model of development was outlined as a research approach with a set of nine propositions to consider when using this approach. It is not a requirement of the research approach that all nine propositions are adhered to, as the propositions relate to the different nested systems, and not all the systems have to be studied simultaneously. Despite this, it is important to consider the propositions when making inferences from research findings within this ecological framework.

Firstly, the research should consider the potential for reciprocal processes to occur at the microsystem level of analysis, such that investigation into the mother-child relationship should consider the bidirectional potential for influence. The second proposition refers to the importance of examining the bidirectional influences on the developing child from all individuals within the child's immediate setting. The third proposition suggests that the research should recognise the social system operative in the research setting, for example, taking account of the influence of the mother-father relationship in a family system when investigating the child. Proposition 4 is concerned with taking account of the physical environment as possible indirect influences on the social processes taking place within the setting. The fifth proposition suggests that to understand the development of a child, researchers need to consider interactions between different settings within the mesosystem and the subsequent influence of these interactions on the child. Using the model illustrated in Figure 1.2, a school's approach to encouraging good peer relations may not only influence the child directly but may also influence the approach to relations within the home setting.

Proposition 6 is an extension of the fifth proposition that directs researchers to consider possible sub-systems that may develop between settings and thus the possible

higher-order effects that may result. Within Figure 1.2, such a sub-system may occur between parent and teacher, or between parent and grandparent. Proposition 7 highlights the natural experiments that occur with ecological transitions, which are associated with changes in role and setting as a function of a person's maturation or of events in the life cycle of others responsible for the person's care or development, (e.g. childbirth, transition to school, relationship breakdown). Bronfenbrenner (1977) suggests that these transitions provide a ready-made experiment of nature with a built in before-after design in which each subject serves as his or her own control. Such transitions will result in changes over time of role, activity and often place for the developing individual and their carer, (e.g. wife to mother, or, child at home to school pupil). Proposition 8 is concerned with the exosystem and the macrosystem, examining the larger contexts that affect the immediate setting, such as socioeconomic and demographic factors. Finally, Bronfenbrenner (1977) proposed the idea of the transforming experiment, in which action is taken to depart from cultural or institutional ideologies. This may be illustrated with an intervention study, such as introducing a flexible work schedule for employees with families in a business and comparing child outcomes. Bronfenbrenner's ecological model of development posed specific challenges for researchers of child development. In the next section I consider the methodological challenges to inform the present study design.

1.4. *Methodological Issues in Examining Pathways to Disorder using an Ecological Approach*

Developmental psychopathology suggests that internal and external processes implicated in maladaptation do not occur in isolation but within complex processes, that provide a challenge for research and analyses to disentangle mediating and moderating influences

on an outcome. It is likely that a multitude of factors across broad domains of biology, psychology, and sociology will be at least indirectly related to the aetiology, course and sequelae of risk conditions and psychopathology. Sameroff (1995) provides theoretical support for the ecological model proposition in which the development of a child is seen as the product of continuous dynamic interaction of the child and the experience provided by his or her family and social context, i.e. a transactional process. The transactional process emphasises the effect of the child on the environment, in that experiences provided by the environment are not independent of the child. It is such that the child's function at any point in time is not a function of the initial state of the child or the initial state of the environment but a complex function of the interplay of child and environment over time. Patterson's theory of coercive behaviour is a good example of transactional processes in development (Patterson, 1982). Patterson suggested that children normally engage in some noncompliance, but if parents are inept at disciplining they create a context where the child is reinforced for learning a set of coercive behaviours. Poor discipline includes lack of monitoring, harsh discipline, lack of positive reinforcement and lack of involvement with the child. The child develops non-compliant behaviours such as whining, teasing, yelling and disapproval, which escalate parents' negative coercive responses and promotes further child noncompliance eventuating in aggressive behaviours. The poor parenting leads itself to a lack of social strategies learnt by the child to cope with peers and others, leading to rejection by their peers, poor academic performance, delinquent peer association and subsequent delinquent behaviour. These findings suggest a transactional process existing between the individual and the environment.

Analysis of transactional processes requires mediational and moderator models. Mediational models answer the questions 'how' and 'why' risk conditions lead to maladaptive outcomes. Mediators are the generative mechanism by which an independent variable, such as, parental depression influences the outcome of child adjustment. Moderator models tell us 'who' is at risk and 'when', working from the assumption that the nature and degree of risk for a child to develop psychopathology is not uniform across people or conditions. There are two most common moderator models used in developmental psychopathology, the synergistic or multiplicative effects model, and the organism-environment interaction model.

The first of these models, the 'synergistic model', specifies that the occurrence of two or more factors incurs a greater deleterious impact than the sum of the factors considered in isolation from each other, i.e. there are significant interaction effects. Rutter and colleagues (1976) demonstrated this model when studying children exposed to any one of six family risk factors, (e.g. family discord, maternal psychiatric disorder, family dissolution). Rutter and colleagues found that when children were exposed to 2 or 3 risk factors there was a threefold increase in the incidence of psychiatric disturbance in the children, compared to children exposed to individual risk factors.

The second model, the organism-environment interaction is also known as the 'diathesis-stress' model and seeks to answer who is specifically at risk. The organism component consists of vulnerable, personological characteristics such as temperament, with the environment stressor reflecting some stressful event in the environment for the individual. When these are integrated together, a 'diathesis-stress' model suggests that individuals with particular personal attributes respond differently or more specifically

with greater maladaptation to similar environmental contexts, e.g. children with difficult temperaments (diathesis) may exhibit greater vulnerability in depressive family contexts (stress).

As a means to examine transactions occurring within the developing child's microsystem and mesosystem, cross-contextual information is required from the different members of the child's systems. Developmental psychopathology and the ecological model of development recommend the use of multi-informant, multi-method approaches. This may include questionnaire and interview report measures from those who interact with the developing child in the micro- and meso-systems, plus observational data. The use of questionnaires requires consideration of who is being asked to report on the child's behaviour. Data have been shown to be affected by who is reporting on the child's behaviour, which should not be surprising, as the informants will have interactions with the child in different contexts and within a different sub-system. Loeber and colleagues (1993) found that parents or teachers are better at reporting externalising problems than the children themselves, whereas children are better at reporting internalising problems, compared to parents or teachers. Self-report questionnaires are limited by age-appropriateness. Agreement between different informants is not usually at 100%, but meta-analyses of cross-informant correlations found that the correlations are often significant, although they can be modest, especially when informants play different roles in different contexts with the children (Achenbach, 1997). Cummings and colleagues (2000) suggest that the modest correlations between informants is not necessarily evidence for poor validity or reliability, but rather results from informants' different

characterisations of children's functioning, partly due to observing the child in different settings.

Structured and semi-structured interviews are often used to obtain data from parents, teachers, and children about a child's behaviour. A degree of standardisation and data from studies indicates that structured and semi-structured interviews tend to be reliable and valid indicators of child distress (Shaffer, Fisher, Dulcan, & Davies,1996). Observation of child behaviour allows the recording and measurement of multiple dimensions of behavioural and emotional responses within different contexts, although naturalistic observation may provide a particular challenge for the rigours of coding.

The ecological model for research allows for the blending of the strengths of the different approaches together and can be demonstrated in the use of the standard procedures in the home environment (Clark, Kochanska, & Ready, 2000). Clark and colleagues carried out a longitudinal study using multimethod investigation of the mother's personality and its interaction with infants' negative emotionality as predictors of parenting behaviour. To measure children's negative emotionality within the home, Clark and colleagues used a standard observational procedure, the Laboratory Temperament Assessment Battery (Lab-TAB; Goldsmith & Rothbart, 1994), as no props or setups that needed a lab setting were required. In this example the strength of the standardised observational task was blended with the naturalistic setting of the home.

Multi-informant reports and observations are crucial to examining the interactions and the pathways to disorder within an ecological model. Most often the informants used are those within the family of the developing child. As described earlier within the Bronfenbrenner model, the microsystem of the family is subject to the general system

theory principles (von Bertalanffy, 1968). In particular the principles of wholeness and order, hierarchical structure, adaptive self-stabilisation, and adaptive self-organisation are embedded within the family system. Firstly, the family is considered as an integral whole, from the view that the whole is greater than the sum of its parts and has properties that cannot merely be understood by combining the characteristics of each part together. It is the interactional processes that are important, because the relationships within a family place constraints on the behaviour of individual family members. The relationships are therefore considered as the most important unit of observation and intervention. Secondly, the family itself consists of a hierarchy of systems, comprising subsystems such as parental, marital, sibling, and which are embedded in larger meso- , exo- , and macro-systems. Thirdly, families act to maintain stability known as 'family homeostasis.' The homeostasis may be positive or negative, positive when the family adjusts to cope with new demands, e.g., moving house, and, negative when the family does not allow others to provide social support in times of crises (Cox & Paley, 1997).

The ecological model suggests that periods of change and transitions in an individual's life may act as a natural experiment. The transition to parenthood is an example of a natural experiment. It has been suggested that new parents are at more risk of psychosis, depression and the blues (Campbell, Cohn, Flanagan, Popper et al.,1992). Adaptation during the transition to parenthood might meet the needs of the new infant but might not meet the needs of the marital relationship (Cox, 1995). It has also been suggested that the quality of the care given to the new infant may be dependent upon the adaptation in the marital subsystem (Cox, Owen, Lewis, & Henderson, 1989).

Finally, the ecological model proposes the need to consider the larger contexts that affect events within the immediate setting, such as the socioeconomic and demographic status of the child and family. Developmental psychopathology provides a research environment to consider these social and demographic factors and their potential influence upon the outcomes for children. Socioeconomic status refers to the capital that the individuals and the family have to draw upon, i.e. financial capital (income and occupational status), human capital (education), and social capital (social support). Whether the socioeconomic measure is a composite of all this capital or used as individual factors is still open to debate (Bradley & Corwyn, 2002).

1.5. *Summary*

This review of developmental psychopathology and Bronfenbrenner's ecological model of development provides a rich research paradigm within which to explore the nature of irritability and its dual role as a dimension of temperament and a symptom of childhood disorder. Sameroff (1995) suggests that detailed analysis of the environment within such an ecological framework could result in the identification of an 'environ-type', which places an individual at potential risk of disorder. It also provides a framework to understand the contribution that the infant brings to the interaction. This thesis will focus the empirical studies of irritability within the family system, the microsystem, exploring the intergenerational transmission of irritability from mother to child and the effect of maternal influences on irritability in infants. Using a multi-method, multi-informant approach, this thesis aims to identify informants' reports of infants' irritability across settings within the microsystem of the family and the extended social support system. The research will use the natural before-after experiment in the transition to parenthood

for the mother to understand the potential risk factors of irritability in infants, and will take account of the exosystem through analyses of the sociodemographic factors.

The role of irritability as a behavioural response in the temperament literature and as a symptom of disorder in the psychopathology literature will be examined in Chapter 2. Such a detailed examination will identify the potential risk and protective factors of the development of irritability and disorders in childhood, and will shape the methodological detail for the two empirical studies described in Chapters 3 and 4.

CHAPTER 2

Irritability: The Measurement of a Dimension of Infant Temperament, a Symptom of Childhood Disorders, and the Relationship between Them

---

2.1. *Introduction*

The previous chapter set out the research framework for the present thesis, developmental psychopathology and ecological development theory (Bronfenbrenner, 1977). This complementary approach emphasises the development of an individual within a nested system in which a complex series of interactions occur between the individual and others within the system. In this thesis, the focus is on the family system. Interactions within the family system may lead to either adjustment or maladjustment in the individual over the lifespan, and the pathways toward normal and atypical development are the focus of developmental psychopathologists. Within this chapter the focus is on the development of irritability, which has been seen both as an early temperament predictor of later childhood adjustment and as a symptom of maladjustment in childhood psychopathology. Thus I propose that a focus on irritability will enable further understanding of the relationship between temperament and psychopathology.

Temperament is defined within this thesis according to the definition provided by Rothbart and Derryberry (1981), "constitutional differences in reactivity and self-regulation . . . influenced over time by heredity, maturation, and experience. (p.37)"

In a roundtable discussion about temperament, prominent temperament researchers agreed that there is evidence that temperament is heritable (Goldsmith et al., 1987). Extensive twin studies and adoption studies have indicated that children born to

parents with particular temperaments are likely to have the same temperament (Plomin, Defries, & Fulker, 1988). In a review of heritability studies of temperament dimensions, irritability as part of the neuroticism factor was reported to have heritability estimates larger than .50 (Henderson, 1982).

The knowledge that temperament, and irritability in particular, is likely to be hereditary does not imply that the environment that the infant is born into is unimportant. On the contrary, there is evidence from longitudinal studies of temperament that genetic influence increases over time, suggesting that 'the family environment can augment familial resemblance when family members share environment as well as heredity' (Plomin et al., 1988, p.28). As heredity and family environment both influence temperament over time, the stability of irritability and potential family system factors influencing stability will be explored.

Four models have been proposed as explanations for the relationship between temperament and psychopathology (Nigg, 2006): (1) a *spectrum model* in which normal and abnormal are at different points on a continuum, with psychopathology being a clinical manifestation of temperament; (2) a *vulnerability or resilience model* in which certain temperament types or traits predispose individuals to certain kinds of psychopathology, with some traits protecting individuals; (3) a *pathoplastic model* in which temperament alters the course of a disorder once it occurs; and (4) a *scar model* in which pathological processes alter temperament (Shiner & Caspi, 2003; Tackett & Krueger, 2005; Watson, Gamez, & Simms, 2006). For the purpose of this thesis, I will focus on the first two models, the spectrum and the vulnerability models, to investigate the relationship between irritability and early signs of psychopathology in young

children. Both the pathoplastic and scar models would require longitudinal investigation following the confirmed presence of psychopathology in children, which is beyond the scope of this thesis.

This chapter begins by presenting the evidence on irritability as a temperament dimension in infancy and its measurement. Then the stability of irritability as a temperament dimension from infancy to childhood is explored. Next, the discussion will proceed with a review of the evidence of irritability as a symptom of childhood psychopathology. Following this, the evidence of a relationship between temperament and psychopathology will be presented in relation to irritability. Finally, this chapter will conclude by outlining the research questions that will be addressed in the subsequent empirical chapters of this thesis.

2.2. *Irritability as a Temperament Dimension*

Historical references to irritability as a factor in the temperament construct can be traced back to the Greco-Roman physicians' description of a fourfold typology for "temperamentum" which included a choleric type, i.e. "irritable and quick to anger". In psychology, research interest into irritability is relatively young, stimulated by the work of Shirley (1933) and Bergman and Escalona (1949), and subsequently by Thomas and Chess's landmark studies of temperament in infancy (Thomas & Chess, 1977; Thomas, Chess, & Birch, 1968; Thomas, Chess, Birch, Hertzig, & Korn,1963). Infant studies have been the main focus for research on temperament with the infant considered as the model system for the study of temperament, enabling researchers to observe behaviours before extensive socialisation and development of higher-order controls (Goldsmith et al., 1987).

One of the earliest investigations into irritability in infancy was carried out by Shirley (1933) in a classic longitudinal study of infant development. The study followed the physical and psychological development of 24 infants from 24 hours after birth until the children were 4 ½ years old. The infants were recruited through obstetricians in Minnesota, who sent letters to the expectant mothers. Following a favourable response to the letters, Shirley and her research colleague visited the expectant mothers to initiate the study. Firstborn infants were excluded from the study because Shirley considered that firstborn babies at this time were on average somewhat smaller than children born later, and that firstborn babies were more likely to cause upset in the family routine. There was no attempt made to gain a representative sample, because Shirley's aim was to enlist the cooperation of intelligent and interested mothers who were prepared to remain in the study and participate in the significant time commitments. As a result, the sample was drawn from the three upper occupational classes as defined by the Minnesota Institute of Child Welfare, with the three lower classes unrepresented (50.8% of the Minnesota population at that time).

The data collection took place on a daily basis in the first week after birth, every other day in the second week after birth, then at weekly intervals during the first year and biweekly during the second year. Apart from the first few observations made whilst the mother and baby were in hospital following birth, all the observations of the infants were made within the home environment by two researchers. In total during the two years, 1,370 visits were made, 1,944 examinations were made by the observers, and a total of 4,181 records were obtained from the mothers.

Shirley examined irritability, as it was recognised as an important element of personality. The method adopted by Shirley was to record just what the baby did (i.e. grasped the toy and threw it on the floor), and also to supplement these behavioural records with non-scorable reactions, such as fussing, crying, or watching the examiner. Irritability was recorded for each type of examination, physical, anthropometric and psychological, by writing the word 'screaming' or 'crying' or 'fussing' at the appropriate point. The total irritability score was taken as the total number of screaming, crying and fussing counts for each examination. A percentage total irritability score was used to enable comparisons between infants, because the number of items and the time length of each examination differed. Complete data consisted of screaming, fussing and total irritability scores from more than sixty records of each of the infants throughout the entire first year.

The study found that irritability in the infants was much greater from birth to 6 months than from 6 months to a year, and that in general the younger babies are, the more irritable they are. The peak irritability in infants occurred during the first 8 weeks, in which the average amount of irritability was about 25%. In these early weeks the median score very closely approximated the mean, but after 8 weeks there was considerable discrepancy between the mean and median scores, indicating that the mean was greatly affected by a few babies who were very irritable. Further investigation showed that 2 out of the 24 babies studied were very irritable from the hospital period onwards, whereas the majority of infants were initially irritable but this irritability reduced throughout the first year. Infants who were irritable showed irritability across the different types of examination, suggesting that irritability is more a function of the infant than the situation.

Shirley concluded that it seemed likely that a tendency towards marked irritability is inborn.

Shirley's early descriptive work on irritability in infancy was a lone study at the time, and throughout the next two decades, there was very little research carried out on individual differences in temperament. The lack of research on temperament between the 1930's and 1950's occurred because of concerns about constitutionalist views at the time that linked constitution to psychopathic tendencies, and because of the predominant view held that behavioural differences were considered products of the environment (Thomas et al., 1968). Against this backdrop, Thomas and colleagues believed that they could not ignore their clinical observations of the lack of simple relationships between environmental circumstances and their consequences, and differential patterns of responses to similar stresses and parental care. Their response in 1956 was to mount a longitudinal study, the New York Longitudinal Study (NYLS), with one of its central aims to define the temperament characteristics in children (Thomas et al.,1968). This study was the launch of significant work in the temperament field.

Thomas and colleagues were inspired by the individual differences among their own children to study individual differences in the primary reaction patterns of infants. Within the NYLS, 85 families were recruited over 6 years, and 141 children studied. In 1968, following 12 years of the study, 136 of the original 141 children were retained in the sample, with 5 children lost due to long-distance changes to residence. To identify the temperament characteristics of infants, Thomas and colleagues interviewed parents when the infant was 3 months old and then at 3 month intervals until the child was 18 months. The temperament characteristics of the infants described by parents during interview

were collated for the first 22 infants. From the parental descriptions of the infant behaviours, nine dimensions of temperament were proposed: *rhythmicity* of biological functions; *activity*; *approach* to or *withdrawal* from new stimuli; *adaptability*; *threshold of responsiveness*; predominant quality of *mood*; *intensity of reaction*; *distractibility*; and *persistence/ attention span*. Irritability, as defined earlier by Shirley as fussing and crying, was part of the predominant mood dimension.

A validity check was made of the parent reports of the infants' behaviours by carrying out 2-3 hour behavioural observations of the infants at home. The observations were scored for the temperament characteristics using the same criteria derived from the parent interviews. The agreement between the independent observers and the parents' reports of infant temperament characteristics was significant at the $p < 0.01$ level of confidence, with independent observers' agreement with each other also significant at the $p < 0.05$ level of confidence (Thomas, et al., 1968). Thomas and colleagues concluded that the parent interview reports were a valid reflection of the infants' behaviour, and that a valid report is possible if descriptive, factual information is requested from the parents and that the behaviours referred to are not too remote in time. Thus a new set of criteria had been developed to measure temperament in infants, and sparked renewed interest in the identification of the most appropriate infant temperament dimensions.

Subsequent studies did not find Thomas and colleagues nine dimensions of infant temperament to be as reliable as required for research purposes, as there was overlap between some of the dimensions and some of the individual items on dimensions did not correlate with each other (Rothbart & Bates, 1998). It was at this stage that different groups of temperament researchers embarked upon programmes of work that resulted in

the development of varying descriptors of infant temperament and varied measures (Bates & Bayles, 1984; Buss & Plomin, 1975; Carey & McDevitt, 1978; Rothbart, 1981). In the case of Shirley's concept of infant irritability, Thomas and colleagues reframed irritability as negative mood in their predominant mood dimension. Other researchers also included the phenomenon of irritability as a temperament dimension but each used different terms.

Rothbart has clearly identified infant irritability as irritable distress measured as fussing and crying in a wide variety of situations and experiences, and named the dimension as 'distress to limitations', operationalised in the carer's questionnaire on infant temperament, the Infant Behaviour Questionnaire (IBQ; Rothbart, 1981). Carey and McDevitt (1978) chose to adopt Thomas and colleagues' nine dimensions and operationalised the measurement of these dimensions as the Revised Infant Temperament Scale, measuring irritability as predominant negative mood that is further combined into a fussy-difficult composite (RITQ; Carey & McDevitt, 1978).

An added complication to the debate about the most appropriate infant temperament dimensions to use within research was the tendency for researchers to group the dimensions together to form composites or higher-factors. Thomas and colleagues combined dimensions that co-varied into three temperament constellations, one of which was termed 'difficult temperament' – irregularity in biological functions, a predominance of negative (withdrawal) responses to new stimuli, slowness to adapting to changes in the environment, a high frequency of expression of negative mood, and a predominance of intense reactions (Thomas et al., 1968). Irritability, now termed negative mood, was

therefore subsumed into this difficult temperament constellation within the Thomas and Chess theoretical tradition.

Lack of replication of Thomas and colleagues' findings on the reliability of a 'difficult' construct has led to alternative definitions of the 'difficult' construct (Bates & Bayles,1984). Using factor analysis on parent reports of infant and toddler temperament, Bates and Bayles found that the core of what parents mean when they rate a child as difficult is frequent and intense expression of negative emotion. A new index of 'difficultness' construct was thus developed to reflect this finding, excluding the dimensions of adaptability and approach used in Thomas and colleagues' original 'difficult' construct (Bates, 1987). In other words, the 'difficultness' construct from the Bates theoretical tradition stripped down the difficult temperament constellation from Thomas and colleagues to the descriptions that reflect irritability, 'frequent and intense expression of negative emotion'.

The use of different terminology to describe similar constructs in temperament research creates problems when trying to compare studies (Rothbart, Ahadi, & Evans, 2000), although evidence that different measures tap the same constructs suggest that comparison is possible. A good example of this was demonstrated by Goldsmith and Rieser-Danner (1986), who had both mothers and day care teachers complete three different temperament scales for the same infants aged 4- to 8-months. The three scales used were the Revised Infant Temperament Questionnaire (RITQ; Carey & McDevitt, 1978), the Infant Characteristics Questionnaire (ICQ; Bates, Freeland, & Lounsbury, 1979), and the Infant Behaviour Questionnaire (IBQ; Rothbart, 1981). The results indicated high intercorrelations between the three scales. Irritability was assessed by the

IBQ distress to limitations scale, the RITQ negative mood scale, and the ICQ fussy-difficult scale. Intercorrelations among these scales for mothers and day care teachers were .54 and .71 respectively.

Within this thesis I have chosen to adopt both the definition of temperament and the dimensions of infant temperament from the Rothbart theoretical tradition, as in that school of thought irritability is traced back to the theoretical traditions of Shirley and is comparable with other contemporary temperament dimensions that reflect irritability (Goldsmith & Rieser-Danner, 1990). Table 2.1 details the various temperament dimensions that reflect irritability within different theoretical traditions, and the composites that are used within temperament studies.

Table 2.1

Temperament Dimensions for Irritability by Name, Composite, Origin, Measures, and Research Authors

Dimension	Description	Composite	Origins	Infant Measures	Developmental Measures
Distress to Limitations (frustration)	Irritable distress	Negative reactivity (includes fear and distress to limitations)	Influenced by Shirley (1933), Diamond (1957), Escalona (1968) and Thomas & Chess (1977)	Infant Behaviour Questionnaire (IBQ; 3-12 months Rothbart 1986)	Toddler Behaviour Questionnaire (TBAQ; Goldsmith, 1996) Child Behaviour Questionnaire (CBQ; Rothbart, Ahadi, & Hershey, 1994)
Predominant quality of mood	Negative mood	Difficult temperament	New York Longitudinal Study (NYLS)	Thomas & Chess (1977); Thomas, Chess & Birch (1968)	
Emotionality	Equivalent to distress, e.g. crying, tantrums, difficulty being soothed, low threshold for aversive stimuli			Revised EAS or EASI (emotionality, activity, sociability, impulsivity. Impulsivity removed from revised version Buss & Plomin 1975; 1984)	
Irritability	Irritable negative affect	Negative emotionality	NYLS Thomas & Chess nine dimensions	Revised and Short-Form Infant Temperament Questionnaire (RITQ; Carey & McDevitt, 1978; SITQ; 4-8 months; Sanson, Prior, Garino, Oberklaid, & Sewell 1987)	Toddler Temperament Scale (TTS; Fullard, McDevitt, & Carey, 1984) Childhood Temperament Questionnaire (Thomas & Chess, 1977) Middle Childhood Temperament Questionnaire (8- to 12-year-olds; Hegvik, McDevitt, & Carey, 1982)
Fussy-difficult	Irritable distress	Difficult perceived as demandingness accompanying irritability (Bates, 1987)	New York Longitudinal Study (NYLS)	Infant Characteristics Questionnaire (ICQ; 4-6 months; Bates, Freeland, & Lounsbury, 1979)	ICQ version for 13- and 24-month-olds (Bates & Bayles, 1984)

2.3. *The Stability of Irritability from Infancy to Childhood*

Research interest in the stability of temperament arose from the debates on the definition of temperament and the resultant view from several researchers of a constitutional basis for temperament (Buss & Plomin, 1975, 1984; Rothbart & Derryberry, 1981). Whilst a constitutional basis of temperament would suggest stability in temperament across the lifespan, research findings have been mixed (Fish & Crockenberg, 1981; Pedlow, Sanson, Prior, & Oberklaid, 1993). Temperament exists within a developmental framework and the behavioural manifestations of temperament may change according to age. To examine stability of temperament across development, measures that reflect age-appropriate behavioural manifestations of the temperament construct are essential. Some temperament researchers have embraced the conundrum of continuity of temperament and the maturational changes that occur with development, and have described temperament as developing over time with influences from maturation within the context of experience (Rothbart & Derryberry, 1981). The Rothbart theoretical tradition thus developed temperament measures to assess the same construct in different age appropriate behaviours, such that for irritability, the distress to limitations dimension of the IBQ maps onto the proneness to anger dimension for early childhood on the CBQ (Rothbart, Ahadi, Hershey, & Fisher, 2001).

Stability of individual differences in irritability across the lifespan would suggest similar rank orderings of the same individual on the same variable at another age, demonstrated through significant correlations across the ages (McCall, 1986). Studies examining the stability of irritability within the Rothbart framework of temperament, have shown continuity of irritability from infancy to childhood when both observation

and parent ratings of infant irritability are assessed. In a longitudinal study (Rothbart et al., 2000), 26 infants were assessed in a laboratory and using mothers' IBQ reports at ages 3, 6.5, 10 and 13.5 months, and then followed up to 7 years of age when the temperament was assessed again by mothers using the CBQ (Rothbart, Chew, & Gartstein, 2001). Significant associations were found in mother reports of distress to limitations (irritability) from 3 months to 7 years, $r = .50$, and between laboratory observations in infancy at 6.5 and 10 months and maternal reports of irritability at 7 years, $r = .36$, and $r = .59$, respectively.

As Rothbart and colleagues used only a small sample size a more extensive study was carried out using parent ratings within the same theoretical framework (Komsi, Raikkonen, Pesonen, Heinonen, Keskivaara, Jarvenpaa, & Strandberg, 2006). Within this more extensive study, Komsi and colleagues (2006) explored the continuity of temperament in Finnish children (n = 231) from 6 months to 5.5 years, using mothers' reports of temperament with the IBQ in infancy and the CBQ in childhood. Mothers' reports of infant distress to limitations (irritability) at 6 months were significantly related with mothers' reports of anger proneness (irritability) at 5 years, $r = .16$, $p < 0.05$. Similar results were reported in a study using Rothbart measures of irritability at infancy, at toddler age and at early childhood (Gartstein & Rothbart, 2003). Caregiver reports of irritability were significantly stable from infancy to preschool age $r = .29$, $p < 0.05$, and for the interim years, i.e. between infancy and toddler age, $r = .22$, $p < 0.05$, and between toddler and preschool age, $r = .52$, $p < 0.05$. Thus using the age-equivalent Rothbart temperament measures there is evidence of continuity for irritability between infancy and childhood, both using parent reports of temperament and independent observation of

infant temperament (Rothbart et al., 2000). The next step is to explore the continuity and discontinuity of irritability within different theoretical traditions.

Using independent examiners as an alternative means of measuring temperament in infancy and childhood, and to provide a way of controlling for subjective parental perceptions, a study of the continuity of infant irritability was carried out using different measures to the Rothbart tradition (Riese, 1987). In a study of temperament from birth to 24 months, neonatal irritability was measured in 67 infants between 1 and 4 days of age, and then at 24 months emotional tone was used as the construct to describe the principal emotional state exhibited during the rating period, ranging from extreme distress to animated laughter (Riese, 1987). The emotional tone construct was assessed in relation to limitation and frustration tasks that the toddlers were engaged with in the laboratory. The emotional tone construct therefore appears to have face validity for mapping on to the infant irritability construct within the Rothbart tradition described as 'distress to limitations'. Firstly, using simple correlations, neonatal irritability was significantly related to 24-month emotional tone, $r = -.36$, $p < 0.05$. Further multiple regression analyses revealed that irritability provided a consistent link between newborn and older infants. When the differences between those with extreme emotional tone scores were compared (i.e., upper quartile = high versus lower quartile = low), there was a significant difference in the neonatal irritability scores, $t\ (64) = 2.58$, $p < 0.01$. Toddlers who were distressed at 24 months had high irritability ratings as neonates. It therefore appears that independent examination also reveals stability of irritability from infancy to toddler age (Riese, 1987). From this evidence we can conclude thus far that irritability appears to show moderate stability between infancy and toddlerhood and infancy and early

childhood, when measured by parent report and independent observation across different theoretical traditions. The next step is to consider the stability of irritability during infancy.

An excellent study was carried out to try to answer these concerns (Lemery, Goldsmith, Klinnert, & Mrazek, 1999). Lemery and colleagues examined the relative stability of different temperament attributes, including irritability, in 180 children, using maternal ratings obtained from either 2 or 3 different temperament questionnaires at 3, 6, 12, 18, 24, 36 and 48 months of age. Using temperament questionnaires from three theoretical traditions (see Table 2.1), Rothbart and Goldsmith (IBQ & TBAQ), Bates and colleagues (ICQ), and Carey and colleagues (R-ITQ, TTS & BSQ), the mean interscale correlation of the irritability measures from the different questionnaires (across all ages) was .50, with a range from .40 – .60. These interscale correlations suggest that the measures are assessing the same construct, in this case, irritability, and that this appears to hold across the age-appropriate measures.

Lemery and colleagues used the different questionnaire measures to construct a composite measure of 'distress-anger' at each age of assessment as a way to minimise measurement bias. The distress-anger composite consisted of all the items that measure irritability in each of the measures. For the 3- to18-months age groups this included IBQ distress to limitations, ICQ fussy, ITQ mood (3 and 6 months only) and TTS mood (12 and 18 months only). For the 24- to 48-months age groups the distress-anger composite consisted of ICQ difficult, TTS mood (24 and 36 months only), BSQ mood (48 months only). Composites were also constructed for the temperament attributes positive emotionality, fear, and activity level. This allowed the authors to explore the stability of

the constructs within a developmental framework. Using different statistical models, the 'distress-anger' temperament attribute was assessed along with other temperament attributes to find the model of stability-change that best fit the data.

Using structural equation modelling, Lemery and colleagues concluded that there was a pattern of increasing stability of distress-anger from infancy to the toddler-preschooler period (2- to 4-years of age). Prediction of continued distress-anger from infancy to toddler-preschooler period was the strongest amongst the temperament attributes studied. There was some evidence of change in the stability of distress-anger during the infancy period, but the toddler-preschooler period had a high pattern of stability. The statistical model that best fit the longitudinal data was a mediational model in which the reported distress-anger assessed at time points between 6 and 18 months partially mediated the relationship between distress-anger reported at 3 months and distress anger reported at 24 months. This model suggests progressive change in the development of distress-anger rather than absolute stability, due to a causal process operating only early in life. This raises the possibility of environmental influences that may influence the infant between 6 and 18 months, resulting in the development of temperament irritability that becomes increasingly stable between 18 months and beyond.

Previous concern of only modest stability of irritability in the first year of life has brought questions about the conceptualising and measurement of infant temperament (Hubert, Wachs, Peters-Martin, & Gandour, 1982), but evidence discussed here points to a pattern of change and continuity within a developmental framework. Irritability in infants appears to change from birth to 18 months, becoming more stable with age and particularly stable between 2 to 4 years of age. The change in infant irritability may

represent lawful discontinuity as has been previously suggested (Belsky & Pensky, 1988; Lemery et al., 1999), perhaps occurring as a result of family system factors (Belsky, Fish, & Isabella, 1991; Crockenberg, 1986).

2.4 *Gender Differences in Temperament*

Differences in temperament, between boys and girls, have not been consistently identified before 12 months of age (Pauli-Pott, Mertesacker, Bade, Haverlock, & Beckermann, 2003; Hane, Fox, Polak-Toste, Ghera, & Ghuner, 2006). In one study of parental perceptions and temperament development, 43 girls and 58 boys were observed for temperament at infant ages of 4, 8 and 12 months and at the same time, mothers completed the IBQ (Rothbart, 1981). At 8 months mothers described their male infants as more negative in emotionality than female infants (Pauli-Pott et al., 2003). In a further study, from the Australian Temperament Project (ATP), it was suggested that systematic differences in temperament between boys and girls were unlikely to appear before 4 years of age (Buss & Plomin, 1975), but surprisingly significant sex differences were found for 6 out of the 9 temperament dimensions for the younger toddlers (mean age = 20.5 months), with boys having significantly more negative mood than girls $t = 3.16$, $p < 0.002$, (Oberklaid, Prior, Sanson, Sewell, & Kyrios,1990). A recent meta-analysis of 189 studies on sex differences in temperament, concluded that there was no sex differences reported for the temperament dimensions: difficulty, emotionality, anger and frustration, or distress to limitations, all terms used to describe irritability (Else-Quest, Shibley Hyde, Goldsmith, & Van Hulle, 2006). From these studies there appears to be some debate about the likelihood of sex differences in infant irritability.

A comprehensive study of irritability in infancy would therefore consider potential gender differences and focus on possible family system factors that may influence infant irritability at an early stage of development, and preferably before the period of greater stability in temperament, i.e. between 6 and 18 months. The possible family system factors that may influence the development of irritability are now explored.

2.5. *The Mother-Infant Relationship*

2.5.1. *Maternal Perception of Infant Irritability*

Concern about the validity of using parents to report on their child's temperament is fuelled by evidence of moderate correlations between parent reports of temperament and laboratory or naturalistic measures (Kagan, 1998; Saudino, Cherny, & Plomin, 2000). Yet parent reports provide a useful perspective on the children, not only because parents witness a rich range of child behaviours, but also because the perception a parent holds about a child may influence the interactions between that parent and child. Indeed evidence suggests that mothers' reports of infant temperament are influenced by maternal perception (Pauli-Pott et al., 2003).

Research has demonstrated that infant temperament develops according to maternal perceptions of the child, and affects the quality of the mother-child relationship, and the child's adjustment in early childhood (Pauli-Pott et al., 2003; Crockenberg & Acredelo, 1983; Olson, Bates, & Bayles, 1989). In a longitudinal study of 101 infants and their mothers, which assessed the relationship between mothers' perceptions of their infants' irritability and observed irritability, the mothers' perception of irritability in their infants at 8 months was predictive of the observed irritability in their infants at 12 months (Pauli-Pott et al., 2003). What affects the mothers' perception of their infants has also

been subject to analysis. Both maternal mental state and partner relationship have been associated with parents' perception of their babies (Bates & Bayles, 1984; Coffman, Levitt, Guacci-Franco, & Silver, 1992).

2.5.2. *Maternal Mental State*

Maternal anxiety and depression may influence the mothers' perception of their infants' temperament (Edhborg, Seimyr, Lundh, & Widstrom, 2000; Vaughn, Bradley, Joffe, Seifer, & Barglow, 1987). A study of 304 Swedish women and their infants used the ICQ (Bates et al., 1979) to measure fussy/difficult temperament in infants at 2 months, and the Edinburgh Depression Scale to measure depression in mothers concurrently (Edhborg et al., 2000). The study found that depressed women were more likely than non-depressed women to report their infants as fussy/difficult, F (1,278) = 17.17, $p< 0.0001$. Similar results were found when mothers' anxiety was assessed (Vaughan et al., 1987). In a study to identify the influence of maternal characteristics on maternal ratings of infant temperament, prenatal assessments of the mothers' anxiety predicted reports of infant difficult temperament as measured by the ITQ and the revised ITQ (Carey, 1970; Carey & McDevitt, 1978) at 4 and 6 months (Vaughn et al., 1987). On the basis of these results some researchers have expressed concern about the validity of the Carey infant temperament scales in identifying temperamentally difficult infants (Vaughan et al., 1987), but whether or not the mothers' perceptions of their children are correct or affected by the mothers' mental state, the perceptions still exist as an important part of the mother-child system, and may serve to augment genetic predispositions to difficult or irritable temperament, and subsequent adjustment. For this reason parental reports of

infant irritability will be explored in this thesis in the context of the mothers' mental state in relation to anxiety and depression.

2.6. *Irritability as a Symptom of Childhood Psychopathology*

The temperament literature has provided clear evidence for stability and consolidation of irritable temperament in the first years of life. The spectrum model of the pathway from temperament to psychopathology (Nigg, 2006) suggests that irritability as a symptom of disorder is an exaggerated manifestation of the normal distribution of irritability on a continuum. The role of irritability as a symptom of childhood psychopathology is now considered, followed by an examination of the literature to understand better the relationship between early indications of irritability and psychopathology in children.

Irritability has been the subject of much recent attention in relation to childhood disorders (Baroni, Lunsford, Luckenbaugh, Tobin, & Leibenluft, 2009), and is a topic of debate in the development of diagnostic categories in DSM-V, the new diagnostic classification system. Within the classification of disorders in preschool children, DSM-IV-TR (APA, 2004), irritability has already been recognised as an age-appropriate symptom of depression in children as young as 2 years (Carlson & Kashani, 1988; Luby et al., 2002; 2003a; 2003b; Mitchell, McCauley, Burke, & Moss, 1988; Ryan et al., 1987). Irritability is a symptom of adult bipolar depression within the current DSM-IV-TR (APA; 2000) classification system. Recent studies of bipolar disorder in children have highlighted irritability as a symptom central to the debate in diagnosis (Biederman et al., 2005; Carlson, Loney, Salisbury, & Volpe, 1998; Fergus et al., 2003).

There is a difference between the temperament literature and the psychopathology literature in the way that they report on irritability. Within the temperament literature,

irritability is measured as a dimension of temperament but within the psychopathology literature irritability is reported as a symptom category. Carlson (1983) found that bipolar children between ages 15 months to 8 years experience irritability instead of euphoria as the predominant mood symptom, whereas for bipolar children aged 9- to 12-years irritability occurs with euphoria. Fergus and colleagues (2003) carried out a retrospective study using parents willing to rate the presence and severity of symptoms in each year of their children's lives for children with and without a diagnosis of bipolar disorder. The results indicated that children diagnosed with bipolar disorder had a 10% or greater incidence of irritability by age 1 and age 3, compared to children diagnosed with other non-bipolar disorders, or children with no disorder. Fergus and colleagues concluded that irritability was associated with later clusters of more classic manic and depressive symptoms, sufficient to lead to a diagnosis of bipolar illness. What was noteworthy from this study is that the children who had unipolar depression also experienced irritability but not to the same extent as the children diagnosed with bipolar depression.

Investigation of the symptoms of externalising disorders also reveals symptoms that are akin to irritability. Oppositional Defiant Disorder (ODD) has nine symptoms, of which three have face validity as descriptors of irritable behaviour: loses temper, is touchy/easily annoyed, and is angry/resentful (DSM-IV; 2000). It is therefore not clear where the distinction between irritability as a temperament dimension and ODD as a clinical diagnosis can be made (Loeber, Burke & Pardini, 2009). In a study of 92 boys aged 4- to 5.5-years with diagnoses of ODD, nearly all the boys had temper tantrums, 50% had angry/resentful and touchy/easily annoyed symptoms, with the latter group of symptoms being associated with poorer outcomes two years later, including higher levels

of internalising and externalising problems and higher probability of psychiatric disorder (Speltz, McClellan, DeKlyen, & Jones, 1999). Emerging evidence suggests that ODD is a precursor to externalising disorders such as Conduct Disorder (CD) and antisocial behaviour, and to internalising disorders, such as mood disorders and anxiety (Burke, Loeber, Lahey, & Rathouz, 2005; Nock, Kazdin, Hiripi, & Kessler, 2007). Evidence is also emerging that ODD explains the co-morbidity between externalising and internalising disorders (Loeber et al., 2009). The links amongst irritability, ODD and both internalising and externalising disorders raise the need to examine the pathway from early irritable temperament to both externalising and internalising disorders. The tendency for temperament researchers to move swiftly to higher-order factors (such as 'difficult temperament' or 'negative emotionality') rather than studying individual attributes, such as irritability, may mask important features of the developmental pathway from temperament to disorder.

The psychopathology literature has drawn attention to irritability as an important symptom of both internalising and externalising disorders. Researchers have even developed a set of criteria for a new diagnosis, Severe Mood Dysregulation (SMD; Leibenluft, James, Blair, Charney, & Pine, 2003), that categorises children with non-episodic irritability, to differentiate these children from episodic irritable moods present in bipolar depression. Irritability is thus an important symptom of disorder in children but the debate continues as to the nature of the pathway from temperament to disorder. Evidence on irritability as a risk factor for disorder is now considered.

2.7. *Irritability as a Risk Factor for Childhood Psychopathology*

Within the psychology literature, irritability has been examined as a dimension of temperament that influences both externalising and internalising trajectories during early childhood (Gilliom & Shaw, 2004). This hypothesis is in line with the vulnerability model of temperament (Gilliom & Shaw, 2004), in which irritability is seen as a vulnerability risk factor for later adjustment. Early indicators of future adjustment risks are important to assist interventions to prevent extreme behavioural and emotional difficulties in preschool developing into psychopathology in middle childhood and beyond (Gilliom & Shaw, 2004).

A relationship between irritability and disorder was evident in a study of associations between emotionality, self-regulation, adjustment problems and positive adjustment in children aged 7-to-10-years (Lengua, 2003). Within this study, emotionality included the dimensions of irritability, fearfulness, and smiling/laughter, and self-regulation included the dimensions of attention, inhibitory control, and impulsivity. Adjustment problems in the study were described as internalising and externalising problems, with positive adjustment defined as well-being and social competence. Lengua hypothesised that the different components of negative emotionality would relate differently to adjustment indices, with irritability related to externalising problems and fearfulness related to internalising problems.

Seventy-nine families provided complete data that were collected using highly structured, scripted 2.5-hour interviews and structured tasks in the families' homes at two time points, 1 year apart. The children's mean age at Time 1 was 9.9 years. Mothers and children were interviewed by separate interviewers, and questionnaires were administered

as part of the structured interviews. Following the child interview, children engaged in structured tasks to measure emotionality and self-regulation, which were videotaped, and coded later by coders who were unaware of the study hypotheses. The Early Adolescent Temperament Questionnaire (EATQ; Capaldi & Rothbart, 1992) was used to measure irritability and fear. Mothers reported on children's externalising and internalising problems using the Child Behaviour Checklist (CBCL; Achenbach, 1991), and the children reported on their own depression and externalising behaviours, using the Child Depression Inventory (CDI, Kovacs, 1981) and the Youth Self-Report (YSR; Achenbach, 1991) respectively. From the initial zero-order correlations, irritability was related to higher levels of internalising and externalising problems and lower social competence, both concurrently and longitudinally. Using multiple regression the independent effects of irritability on higher levels of externalising and internalising problems remained significant after controlling for the effects of the other emotionality and self-regulation variables. Lengua (2003) concluded that irritability appears to relate to a broader range of outcomes than externalising behaviour alone, and that this may be as a result of problems in relationships with parents, peers, and teachers, leading to lower social competence as well as distress that can in turn result in internalising problems.

A second study examined the temperament profiles associated with internalising and externalising problems in Dutch preadolescents (Oldehinkel, Hartman, De Winter, Veenstra, & Ormel, 2004; N = 2230; mean age = 11.09; 50.8% girls). Oldehinkel and colleagues found that, whilst girls and boys differed in the types of adjustment problems experienced, the temperament profiles that related to adjustment problems were the same for both sexes. Preadolescent girls were reported by parents on the CBCL (Achenbach,

1991) as having more internalising problems than boys, and boys were reported as having more externalising and more co-morbid problems than girls. The temperament profiles were measured according to Rothbart's temperament definitions using the EATQ (Putnam, Rothbart, & Gartstein, 2008). The EATQ uses frustration as the adolescent descriptor of irritability (frustration being analogous to the distress to limitations construct in the IBQ). Overall frustration was found to be the main temperament factor that related to adjustment in general, whether the child with adjustment problems was a boy or girl. Oldehinkel and colleagues found that adolescents with internalising problems had temperament profiles with significantly higher scores for shyness, fear, and effortful control than the temperament profiles associated with externalising problems. The temperament profiles for adolescents with externalising problems showed significantly higher scores for high-intensity pleasure and frustration than the temperament profiles for those with internalising problems. Interestingly, the temperament profiles for those adolescents with co-morbid problems showed significantly higher scores for frustration and fear than those adolescents with either externalising or internalising problems.

From the studies reviewed there is evidence for a relationship between irritability and psychopathology within childhood and adolescence. The vulnerability model would suggest that the psychopathology has emerged in children who have irritability as a vulnerability trait that predisposes them to maladjustment (Gilliom & Shaw, 2004). However, when the relationship between temperament and disorder is measured concurrently, an alternative explanation for this relationship is provided by the spectrum model, which suggests that adjustment problems are the clinical manifestation of abnormally high levels of irritability, although this may be more relevant to irritability

and externalising problems such as ODD, rather than with internalising problems. Neither the vulnerability model nor the spectrum model can refute the criticism that such a relationship between irritability and psychopathology may be artefactual, and a result of confounding measures in which similar items measure irritability and psychopathology (Nigg, 2006). This concern about the measurement of similar constructs used in the temperament and psychopathology literature requires further examination.

2.8. *The Measurement Confounding Hypothesis*

The measurement confounding hypothesis has been tested in three previous studies (Lemery et al., 2002; Lengua, West, & Sandler, 1998; and Sanson, Prior, & Kyrios, 1990). In a two-stage study, Lemery and colleagues (2002) tested the measurement confounding hypothesis by taking a conceptual approach, asking experts to judge items from behaviour problem and temperament scales as to their best fit to temperament or behaviour problem constructs. The expert assessment led to the production of purified scales excluding 10% of the temperament items that were judged to be confounded. One of these temperament items was irritability, which was judged to be confounded with the clinical symptom of anger.

At the second stage of the study both the original and the purified scales were used to measure temperament and behaviour problem symptoms in an existing longitudinal sample, and comparisons were made between the original and purified scales. Mothers completed the CBQ (Rothbart, Ahadi, & Hershey, 1994) for assessment of temperament in the children at 3.5 and 4.5 years, and mothers, fathers and caregivers completed the Preschool Behaviour Questionnaire (PBQ; Behar & Stringfield, 1974), a measure of problem behaviours, also at 3.5 and 4.5 years. At 5.5 years, mothers and

fathers completed the MacArthur Health and Behaviour Questionnaire (HBQ; Ablow et al., 1999) to measure behaviour problem symptoms.

When the study examined the relationship between both the original and purified temperament scales for predicting behaviour problems, the magnitude of the associations using the purified scales was equivalent to those with the original scales, suggesting that measurement confounding did not account for the observed relation between temperament and behaviour problem symptoms (Lemery et al., 2002). This result corroborated the findings of a similar study in which temperament continued to correlate with psychopathology symptoms even when items present in both temperament and psychopathology scales were removed from the temperament scales (Lengua, West, & Sandler, 1998).

Whilst measurement confounding was not considered to explain the relationship between temperament and disorder in two studies, there was some evidence of confounding in a study that was part of the Australian Temperament Project (Sanson et al., 1990). The investigators asked a group of psychologists to rate the relative adequacy of questions from the Short Temperament Scale for Toddlers (Sanson, Prior, & Oberklaid, 1984) and two behaviour problem questionnaires, the Behaviour Checklist (Richman & Graham, 1971) and the PBQ (Behar & Stringfield, 1974), as measures of temperament and also as measures of behaviour problems. From the experts' judgements, the Temperament Scale for Toddlers was considered a better but not excellent measure of temperament than the behaviour questionnaires. The behaviour questionnaires were regarded as good or very good measures of behaviour problems (both externalising and internalising problems) and moderate measures of temperament.

The study identified two temperament scales that showed the most confounding, *Activity/Intensity* and *Irritability*, and the authors concluded that the association between these scales and behavioural problems was largely artefactual. From the studies reviewed here it appears that the jury has still to decide whether or not the relationship between irritability and disorder is a result of measurement confounding or that a relationship exists and can be explained either by the spectrum or the vulnerability models of the link between temperament and psychopathology.

Within this thesis I aim to investigate the measurement confounding hypothesis in relation to irritability and early symptoms of both internalising and externalising disorders. I shall adopt a similar approach to Lemery and colleagues and compare the relationship between irritability and purified and original measures of symptoms of disorder.

2.9. *Summary and Aims of the Thesis*

Irritability is an important temperament dimension that has been variously named and measured by different temperament researchers. From the literature reviewed, the irritable distress construct from the Rothbart theoretical tradition of infant temperament is the contemporary construct of infant irritability that most closely resembles and indeed is theoretically derived from the irritability construct used by Shirley in her classic study on infants. Measured using the IBQ *distress to limitations* scale, Rothbart's *irritable distress* construct has good validity against other measures of irritability in different research traditions. In addition, Rothbart has developed age-appropriate measures to record irritability within a developmental framework, a crucial requirement for any work on pathways from temperament to disorder.

The age-appropriate measures allow for the testing of the stability of irritability within a developmental framework. Studies on continuity and discontinuity of irritability thus far indicate that irritability may show lawful discontinuity in the presence of mediators, such as family system factors, such as the mothers' mental state. The influence of these factors upon irritability in infants is crucial in the mapping of the pathway from irritability to disorder.

The proposal of two models to explain the relationship between temperament and psychopathology has been used in this thesis to guide the literature review. The spectrum model would indicate that irritability as a symptom of disorder is an abnormal level of irritability on a continuum. Extreme manifestations of irritability in childhood are being classed as clinical conditions, with a new diagnosis proposed of Serious Mood Dysregulation (Leibenluft et al., 2003) for children with non-episodic irritability. ODD appears currently to be the clinically recognised childhood disorder that would reflect the spectrum model in relation to irritability. Both internalising and externalising childhood disorders include irritability as a symptom, and concern is expressed within the psychopathology literature about the role of irritability in the most difficult psychopathologies. The psychopathology literature is therefore leading the way in identifying irritability as a key component in childhood disorder. Using the developmental psychopathology framework I aim to bring the psychopathology and the temperament traditions together in this thesis through the investigation of irritability, as a dimension of temperament in infants and their mothers in relation to both externalising and internalising symptoms.

This thesis will enhance the temperament and psychopathology literature by taking the first step in mapping the pathway from irritability to disorder. Firstly I will address the measurement confounding concerns by repeating the comparison of associations between irritability and original and pure measures of symptoms of childhood disorders. As an important methodological contribution, this will enable further research into the appropriate models to explain the relationship between temperament and disorder. Then secondly, I approach the relationship between irritability and symptoms of internalising and externalising problems from the preschool child and adult stages of the pathway. I examine mothers' irritability in relation to maternal mental health. Following mothers' transition to first-time parent, I will describe infants' irritability at 6 months of age using Rothbart's measure of *distress to limitations* that is rooted in the literature to irritability, and finally, I examine the family system factors that may influence intergenerational continuity in irritability from parent to child. This will further the temperament literature in understanding the potential origin, nature and role of irritability in relation to disorders and within the context of the family system. A focus on the mother-infant subsystem will inform future research into the continuity of irritability and its relationship with disorder within a developmental framework and within the family system context. The specific aims of the thesis are set out below.

2.9.1. *Aim 1*

Within this thesis I aim to follow the approach of Lemery et al. (2002) and test the measurement confounding hypothesis at an early stage of emerging symptoms of childhood disorder. In the first empirical study, Study 1 (Chapter 3), a community voluntary sample of families with children aged 3- to 5-years, one parent is interviewed

about the presence or absence of symptoms of both internalising and externalising disorders. The teacher and parent are also asked to report on children's symptoms using a questionnaire. A comparison of associations between a composite measure of irritability and both original and purified interview measures of symptoms of disorder is made to investigate the measurement confounding hypothesis. Without the exclusion of redundant items, future research to investigate irritability as either a dimension that is on a continuum with disorder as with the spectrum model, or as a vulnerability factor in the development of psychopathology would be difficult. Furthermore, it is only after controlling for measurement confounding that we can conclude that irritability is a potential predictor of both internalising and externalising disorders.

2.9.2. *Aim 2*

Having examined irritability in relation to both internalising and externalising disorders in early childhood, I aim to look more closely at irritability in infancy in relation to other dimensions of temperament and infant behaviours. Following from the classic studies by Shirley, Escalona, and Thomas and colleagues, I aim to adopt a descriptive approach to the study of infant irritability. The temperament literature has moved away from the investigation of individual temperament dimensions, but the psychopathology literature is emphasising the importance of irritability in childhood disorders. The spectrum and vulnerability models purported as explanations for the relationship between temperament and disorder argue for irritability being investigated more thoroughly as an individual temperament dimension. Various terminology used to describe the same constructs within the temperament literature has also served to mask the importance of irritability. Within the second empirical study of this thesis, Study 2 (Chapter 4), I aim to use a measure of

irritability that has good validity in the literature (the IBQ distress to limitations scale) and examine convergence across three informants' reports of irritability. In addition, I shall measure irritability in relation to other dimensions of infant temperament, such as activity level and fear, to identify possible groupings of temperament dimensions that may give insight into the different manifestations of disorder in which irritability is a symptom. I will also measure mothers' reports of infant irritability in relation to independently observed infant behaviours, as a means to test the reliability of mothers as informants of their infants' temperament and to understand better the behaviours associated with informants' reports of infant irritability.

2.9.3. ***Aim 3***

In Study 2 (Chapter 4), I shall examine the mother-infant subsystem for factors that may serve to influence the intergenerational transmission of irritability from mothers to their first-born infants. Through antenatal and postnatal interviews and questionnaires I will gain a picture of the mother's own irritability, in relation to her mental health, both emotional disorders and her own history of behavioural symptoms. I will explore the relationship between maternal irritability, maternal emotional and behavioural symptoms and maternal social circumstances, and whether any relationship between maternal irritability and infant irritability is influenced by these maternal factors. Examination of these family system factors as potential predictors of infant irritability will inform our understanding of the potential cycle of irritability within families and in turn the potential pathways to disorder.

CHAPTER 3

STUDY 1

Irritability and Childhood Disorder Symptoms: Testing the Measurement Confounding Hypothesis

---

3.1. *Introduction*

The review of the literature presented in Chapters 1 and 2 has highlighted the need to consider the psychological and psychiatric views of irritability together within the framework of developmental psychopathology to understand further the relationship between irritability and disorder. This relationship between temperament and disorder has been repeatedly demonstrated in studies examining the development of disorder in childhood (Bates & Bayles, 1988; Earls & Jung, 1987; Gilliom & Shaw, 2004; Gjone & Stevenson, 1997; Shaw, Owens, Giovannelli, & Winslow, 2001), and has stimulated research activity exploring the nature of this relationship (Goldsmith, Lemery & Essex, 2004; Levy et al., 1997; Nigg, 2006). Four models, detailed in Chapter 2, have been suggested as explanations for the conceptual and developmental relations between temperament and psychopathology: (a) a spectrum model; (b) a vulnerability model; (c) a pathoplastic effect; and (d) scar effects (Nigg, 2006). A further explanation for this evident relationship between temperament and psychopathology is that of measurement confounding (Frick, 2004; Lahey, 2004), which is tested in this present study.

Temperament measures and psychopathology measures use similar items to assess behaviours that may be deemed as both a temperament characteristic and as a symptom

of disorder, e.g. fear, irritability. Conceptually, temperament dimensions are considered as the normal range of character differences in affective responses and expression (Carey, 1990; Thomas & Chess, 1984), and symptoms of disorder are considered as extremes of these characteristics that result in distress to the individual and interference in the everyday life of the individual (Lengua, West & Sandler, 1998). The result is that one child may be generally fearful of meeting new people but continues to attend school and other activities without distress, whereas another child is so fearful of others that Separation Anxiety Disorder (SAD) is diagnosed.

3.1.1 *Previous Tests of the Measurement Confounding Hypothesis*

The concern about measurement confounding within assessment of temperament and disorder has led to a number of studies testing this hypothesis (Lemery, Essex, & Snider, 2002; Lengua, West, & Sandler, 1998; Sanson, Prior, & Kyrios, 1990). The first study took place as part of the Australian Temperament Project (ATP), a prospective, longitudinal study of 2,443 children, examining the influence of temperament on behavioural development and adjustment, from infancy to school age (Sanson et al., 1990). The authors were concerned that whilst the clinical implications of temperament were being questioned at this time, very little research effort was being made to resolve this measurement concern (Bates, 1986). The resultant study firstly asked 36 practicing psychologists to examine the degree of conceptual overlap between the Short Temperament Scale for Toddlers (Sanson, Prior, & Oberklaid, 1984) and two behaviour problem questionnaires, the Behaviour Checklist (Richman & Graham, 1971) and the PBQ (Behar & Stringfield, 1974). From the psychologists' judgement, the Short Temperament Scale for Toddlers was considered a better measure of temperament than

the behaviour questionnaires, and the behaviour questionnaires were regarded as good or very good measures of behaviour problems (both externalising and internalising problems) and moderate measures of temperament. Activity/intensity and irritability items were rated by the psychologists as reflecting both temperament and behaviour problems, and, using this criterion, items for these dimensions were removed from the temperament and the two behaviour questionnaires. The next stage of the study used data from the ATP to assess the relationship between temperament and both internalising and externalising problems, comparing the associations obtained before and after removal of the potentially confounding items. For irritability and activity/intensity, there was a significant difference in the pre- and post-removal relations with both internalising and externalising problems.

The authors concluded that the association between these temperament scales and behavioural problems is artefactual, but this study of measurement confounding was itself methodologically flawed (Bates, 1990). Using the Short Temperament Scale for Toddlers allowed only 5 items on both the irritability and activity/intensity scales that were subsequently reduced to only 1 item per scale following decontamination of the scales. Critics have suggested that such action will have reduced the reliability and validity of the temperament scales and thus caution should be exercised in interpreting these results (Bates, 1990). It was therefore considered essential to investigate the measurement confounding hypothesis further (Lemery et al., 2002; Lengua, West, & Sandler, 1998).

Using data from an experimental intervention trial for children of divorced parents (N = 232), Lengua and colleagues adopted two methods to investigate the measurement confounding hypothesis for the relationship between temperament and symptoms of

disorder. The first method was to use students and faculty clinical psychologists as experts to judge the relevance of items on temperament and symptom scales according to definitions provided. Secondly confirmatory factor analysis (CFA) was adopted as an empirical method to assess item overlap between measures. Following both these methods items deemed as potential confounders were removed from measures and the relationship between temperament and symptoms tested. Learning from the concerns relating to the ATP study on measurement confounding, Lengua and colleagues tested the scale reliabilities of both the temperament and behaviour measures before and after the removal of items. The scale reliabilities were not significantly reduced by the decontamination (e.g. for negative emotionality scale: mother report original scale alpha = .78; mother report uncontaminated scale alpha = .69). The negative emotionality dimension used by Lengua and colleagues was defined within the Buss and Plomin (1985) temperament tradition and is theoretically linked to irritability (see Table 2.1, Chapter 2), but also included fear and sadness.

In the second stage of the study, the relationships between temperament and symptoms of disorder for the original and the uncontaminated measures were compared. The resultant pattern and magnitude of correlations between temperament and symptoms using the uncontaminated measures were found to be very similar to the correlations using the original measures (Lengua et al., 1998). This result held for two different informants, the mother and the child. Further analysis to test the independent effects of the temperament variables revealed differences between the mother and child path models. Decontamination of the negative emotionality scale for mother reports resulted in a significant but decreased relation between negative emotionality and both depression

and conduct symptoms. For the child reports, decontamination of the negative emotionality scale resulted in a significant though decreased relationship between negative emotionality and depression, but a reduced and non-significant relation between negative emotionality and conduct problems (Lengua et al., 1998). This result may have occurred due to the removal of irritability items from the negative emotionality scale, but the use of negative emotionality as a higher-order temperament factor may have masked this result.

Lengua and colleagues improved upon previous studies that had examined questions about the clinical validity of temperament in predicting disorder (Sanson et al., 1990), firstly by comparing temperament scale reliabilities before and after decontamination and secondly by using two different methods to assess potential confounding, the CFA empirical method and the expert judgment method. The results indicated good scale reliability following removal of potentially confounding items on the temperament scales. Both the CFA and expert judgement methods identified different confounding items and both sets of items were subsequently removed from the temperament scales to produce the uncontaminated scales. The methods were considered to be complementary. The empirical CFA method identified items that had shared variance which was either not accounted for by the items' latent construct or the correlation between constructs, and the experts' method enabled judgements to be made on any conceptual overlap between items. Previous comparisons between empirical and experts' clinical judgements favoured empirical methods (Dawes, Faust, & Meehl, 1989), but Lengua and colleagues concluded that even empirical methods include some element of researcher judgement and that measurement confounding could be avoided if the

measurements for temperament and symptoms include clear definitions of constructs. Finally Lengua and colleagues tested the measurement confounding hypothesis using two different informants, the mother and the child, to answer criticism of shared method variance when testing the measurement confounding hypothesis.

Taking these improved research methods into account, the conclusion that measurement confounding does not account for the evident relation between temperament and symptoms of disorder appears robust. However, from the literature reviewed in Chapter 2 it is apparent that different measures of temperament have been developed within different theoretical traditions and that different temperament dimensions may relate differently to symptoms of disorder. Conceptually temperament dimensions also differ across instruments and therefore the measurement confounding hypothesis should be tested within different theoretical traditions to improve the validity of temperament measures. A further study explored measurement confounding within the Rothbart theoretical tradition of temperament (Lemery et al., 2002).

In a two-stage study, Lemery et al., (2002) tested the measurement confounding hypothesis by taking a conceptual approach, asking experts to judge items from behaviour problem and temperament scales as to their best fit to temperament or behaviour problem constructs. The expert assessment led to the production of purified scales excluding 10% of the temperament items that were judged to be confounded. One of these temperament items was irritability which was judged to be confounded with the clinical symptom of anger. At the second stage of the study both the original and the purified scales were used to measure temperament and behaviour problem symptoms in an existing longitudinal sample, and comparisons were made between the original and

purified scales. Mothers completed the CBQ (Rothbart et al., 1994) for assessment of temperament in the children at 3.5 and 4.5 years, and mothers, fathers and caregivers completed the PBQ (Behar & Stringfield, 1974), a measure of problem behaviours, also at 3.5 and 4.5 years. At 5.5 years, mothers and fathers completed the HBQ (Ablow et al., 1999) to measure behaviour problem symptoms. When the study examined the correlation between behaviour problems and both the original and purified temperament scales the magnitude of the associations using the purified scales was equivalent to those with the original scales. This suggests that measurement confounding does not account for the observed relation between temperament and behaviour problem symptoms (Lemery et al., 2002).

The previous studies have all tested the measurement confounding hypothesis using separate questionnaires to measure temperament and behaviour problems. The result of different methods used to test the measurement confounding hypothesis in these studies indicates that some of the possible confounding between temperament and behaviour problems may lie in the conceptual nature of the items used in the respective questionnaires. Looking specifically at irritability, Lemery and colleagues (2002) identified through expert judgement that two CBQ temperament questionnaire items on the anger dimension ('has temper tantrums when s/he doesn't get what s/he wants' and 'gets mad when provoked by other children') were confounded. In the same study the experts judged one item from the hostile-aggressive dimension of the PBQ behaviour problem symptom questionnaire item 'irritable, quick to fly off the handle' as confounding. Similarly, irritability was identified by psychologists and through subsequent analysis as a confounding variable in the ATP study (Sanson et al., 1990).

Lengua and colleagues also found that both psychologists' judgement and CFA identified some items on the negative emotionality scale to be potential confound items, but no significant difference was found when comparing the relationships between the temperament and behaviour problem symptoms using either the original or purified scales (Lengua et al., 1998). Sanson and colleagues concluded that the relationship between irritability and behaviour problems was largely artefactual, but the reliability of the decontaminated temperament scale in their study may have been affected by the removal of too many items. In contrast Lemery and colleagues tested reliability of their irritability scales and the decontaminated irritability scale reliability was good with $\acute{\alpha} = .73$ (Lemery et al., 2002). The relations between irritability and symptoms remained significant after eliminating potentially confounding items.

The critical observation here is that the behaviour problem scales are screening instruments for potential disorders and are not measures that can be used in isolation to make clinical diagnoses. Crucially, the argument about measurement confounding as a potential explanation of the relationship between temperament and psychopathology may itself be flawed as in some cases the relationship being referred to is a relationship between temperament and potential symptoms of disorder as opposed to the relationship between temperament and clinically diagnosed psychopathology. The conceptual difference is important because symptoms referred to in behavioural disorders may also be temperament characteristics but will only be clinically valid if these characteristics cause distress to the child and/or interfere with the child's everyday life. This is the point at which research on the relationship between temperament and disorder would benefit from the combined efforts from psychology and psychiatry, as suggested by Frick

(2004).The expert raters used in previous studies on measurement confounding were all psychologists or psychology researchers and therefore the judgements made on the relevance of scale items to temperament or behaviour problems may have been subjectively influenced by their field of work. It is likely that the raters will have experienced the different items in the questionnaires presented within the context of their psychology work. This was partly improved in the use of clear definitions provided to the experts in both the Lemery and Lengua studies, but did not go far enough to bring together the two research fields of psychology and psychiatry to examine a criticism of the important relationship between temperament and psychopathology.

Lemery and colleagues came closest to resolving this issue by using a clinically based parent questionnaire to assess the prediction from temperament constructs to clinically relevant symptoms in the children at 5.5 years (HBQ; Ablow et al., 1999). The HBQ assesses symptoms from DSM-IV (1994) and was used to assess the symptoms for three behaviour problem composites: 1) internalising problems (depression; separation anxiety, and overanxious subscales), (2) externalising problems (oppositional defiant, conduct, and overt aggression subscales), and (3) attention deficits (inattention and impulsivity subscales). In addition the HBQ assesses distress and interference caused by the symptoms, although these reports were not used by Lemery and colleagues. An additional step in understanding the relationship between temperament and psychopathology would have involved Lemery and colleagues testing the measurement confounding hypothesis between the temperament measure (CBQ) and the clinically based psychiatric measure (HBQ). Egger and Angold (2006), in a review of the presentation, nosology, and epidemiology of common emotional and behavioural

disorders in preschool children suggested that it would be informative to assess measurement overlap in a concurrent study of temperament dimensions and symptoms of disorder in a preschool population.

3.1.2 *Testing Measurement Confounding within Clinical Instruments*

Within the present study I aim to further our understanding of the nature of the relationship between irritable temperament and psychopathology by testing the measurement confounding hypothesis using a conceptual approach that bridges the gap between psychology and psychiatry. To my knowledge no study has investigated the measurement confounding hypothesis using a clinically-based psychiatric interview as a measure of symptoms of disorder. This has been largely due to the lack of suitable preschool diagnostic interview measures prior to the development of the Preschool Age Psychiatric Assessment (PAPA) developed by Egger and colleagues (Egger, Ascher, & Angold, 2002; Egger, Erkanli, Keeler, Potts, Walter, & Angold, 2006), which is a parent psychiatric interview to assess clusters of symptoms for DSM-IV (1994) disorders in preschool children. I will further the existing research on measurement confounding by examining the relationship between irritability as a specific dimension of negative emotionality and its relationship with clinical symptoms of DSM-IV (1994) disorders using the PAPA in a demonstration project in a UK community.

Rather than examining correlations across temperament and behavioural problem questionnaires, I used a theoretically based operational definition of irritability to develop an irritability scale from individual PAPA items. I then examined any items in the composite irritability scale that overlapped with problem scales in a screening questionnaire measure for problem behaviours, the Strengths and Difficulties

Questionnaire (SDQ; Goodman,1997) and internalising and externalising symptom scales in the PAPA itself. The derivation of a theoretically informed irritability scale based on clinical symptoms and the removal of any of those specific symptoms from both the SDQ and PAPA respectively will allow the test of the measurement confounding hypothesis, through the comparison of relationships between irritability and the original and purified behaviour scales and scales of disorder symptoms.

Measurement confounding may be a valid explanation for the relationship between irritability and particular psychiatric disorders of children. As discussed in Chapters 1 and 2, it is only in recent years that research has been carried out on the prevalence of psychiatric disorders in preschoolers, and on its presentation, nosology, and epidemiology (Egger & Angold, 2006). This has largely been due to concerns that such efforts would lead to inappropriate labelling of young children at a point when there are rapid changes occurring within the behavioural, emotional, and cognitive development of these young children. It has also been believed that criteria for disorders had not reflected developmentally appropriate symptoms (Luby et al., 2003). However, recent research has identified developmentally specific criteria for psychopathology in preschoolers and modifications have been proposed to the DSM criteria (Task Force on Research Diagnostic Criteria: Infancy and Preschool, 2003). DSM-IV-TR (2000) does include some developmentally appropriate symptom classifications for preschoolers, with irritability identified as a key symptom of childhood depressive disorder, but there is still some way to go to incorporate the recommendations from the Task Force.

The debate about symptom classification for preschool psychopathology is important in understanding pathways from temperament to disorder. Following children

from infancy to early childhood will enable researchers to highlight the factors that influence how, when and why infants with certain temperament characteristics develop symptoms that cause them distress and interference. Whilst temperament characteristics have been considered as risk factors for psychiatric disorders across the lifespan, Egger and Angold (2006) suggest that it is also possible that early-measured temperament characteristics could represent the early presence of disorders themselves. These two perspectives could reflect the vulnerability and the spectrum models respectively (Nigg, 2006).

3.1.3. *Measurement Confounding and Comorbidity*

Examination of the types of disorders themselves suggests that the spectrum model may be the case for some psychiatric disorders but by no means all (Egger & Angold, 2006). In particular, ODD is of interest in relation to the temperament characteristic of irritability. ODD has 9 symptoms, of which 3 are conceptually similar to irritability, i.e. 'loses temper,' 'is touchy/easily annoyed,' and 'is angry/resentful' (DSM-IV-TR; 2000). ODD is also of interest in view of its role in the high levels of comorbidity that exist in preschool psychopathology. In studying preschool co morbidity, Egger and colleagues identified ODD as a central mediator in the relationship between anxiety disorders and depression, depression and Conduct Disorder (CD), and the emotional disorders and ADHD (Egger & Angold, 2006). This is of relevance to the measurement confounding hypothesis because preschool depression includes irritability as a main symptom, and two studies of preschool psychopathology have identified high rates of comorbidity between depression and ODD, much higher than expected from studies of older children and adolescents (Angold, Costello, & Erkanli, 1999; Luby et al, 2003).

3.1.4. *Aims of Study 1*

In the present study I aim to further the understanding of the relationship between irritability and childhood disorder by firstly testing the measurement confounding hypothesis through the comparison of relationships between irritability and the original and purified behaviour scales and scales of disorder symptoms.

The second aim of Study 1 is to specifically examine the potential measurement confounding between irritability and symptoms of ODD. I anticipate that there will be significant confounding that may indicate that ODD is a clinical manifestation of later childhood irritability as suggested by the spectrum model.

The third aim of Study 1 is to test the hypothesis that irritability may be a mediating influence in the relationship between internalising and externalising symptoms. Given the high levels of co morbidity between internalising and externalising problems in childhood (Angold, Costello, & Erkanli,1999), and the evidence that irritability has been identified as a common symptom between internalising and externalising problems in childhood (Luby et al., 2003), I anticipate that irritability will mediate the relationship between preschoolers' internalising and externalising symptoms.

The final aim of Study 1 is to test the role of irritability as a potential mediator between ODD and both internalising and externalising symptoms. If both irritability and ODD independently relate to internalising and externalising disorders, it is unlikely that ODD would be an extreme clinical manifestation of irritability.

3.2. *Method*

3.2.1. *Participants*

The children attended nursery or reception classes (in all but one case attached to a local primary school) in Cardiff, Barry, or Newport, Wales, UK, and were aged between 3 and 5 years (mean age = 3.9 years). Of 234 parents who were informed about the study, 94 (40%) gave active consent to take part in Phase 1 of the study, which entailed collection of teachers' reports. From Phase 1 of the study, 93 reports were received from teachers. Sixty-five parents (69%) agreed to participate in Phase 2, and complete a screening questionnaire and an interview. Two parents did not provide complete interview data from Phase 2 and two teacher reports from Phase 1 were incomplete.

Analyses of the teachers' reports of SDQ problem behaviours indicated that there was no significant difference between the children for whom the teacher reports were available from Phase 1 only and the children for whom both teacher and parent reports were available from both phases. The results of the comparative analyses between those children in Phase 1 and those in Phase 2 are shown in Table 3.1. Taking these results into account, those children for whom we had complete questionnaire and interview data from both phases were used as the study sample (N=61; girls = 33; boys = 28).

Table 3.1

Differences in the Teacher SDQ scale scores for samples in Phases 1 and 2

	Phase 1: With no PAPA interview (N= 31)		Phase 2: With PAPA interview (N= 62)			
TSDQ Scales	M	SD	M	SD	*t*	*p* <
Emotional Problems	2.03	2.36	1.87	2.17	0.328	0.743
Conduct Problems	1.26	1.39	1.90	2.58	-1.297	0.198
Hyperactivity Problems	3.60	2.39	3.89	2.86	-0.475	0.636

From the study sample, 11 children (18%) were living in families in receipt of income support benefit, which is slightly lower than the population data for the local authority areas covered (percentage of children in families on key benefits = 23.5%; Department of Work and Pensions, 2006). In addition, 14.8% of the children (N = 9) were living with a lone parent, which is lower than the 24% in UK population (Office of National Statistics; ONS, 2007). Thus this volunteer sample was slightly under-representative of the broader UK population.

3.2.2. *Procedure*

The design and all procedures were approved by the Cardiff University School of Psychology Research Ethics Committee. The Starting School study was established as a pilot study of the Preschool Aged Psychiatric Assessment (PAPA; Egger et al., 2002; 2006) in a British sample of children between ages 3-to5-years. My role in the Starting School study included the planning of the study, finding and negotiating schools willing to participate in the pilot study, meeting with teachers and parents (to inform them about the nature of the study), planning the interview sessions with parents, interviewing

parents, collating the data from teachers and parents, inputting the data, and subsequently analysing the data for this present thesis.

After consultation with head teachers at the schools and nurseries the children attended, families were invited to participate in the Starting School Study, which was described as a study of children's adjustment to the challenges of formal education. Classroom teachers sent information letters and consent forms to the parents of all children attending the nursery and reception classes. The researchers visited the school at the beginning and end of the school day to remind parents to respond to the letters and to answer any queries.

In Phase 1 of the study, classroom teachers completed the SDQ forms for all children whose parents had given permission to participate in the study. The second phase of the study took place within two months of the teachers' assessments.

In Phase 2 of the study one of the child's parents (in 90% of cases, the mother) was given an extended interview about the child and any difficulties he or she might be experiencing. As part of the interview parents also provided information about family structure, any recent life events the children had experienced, and whether or not they were in receipt of state benefits. They also completed the parent's version of the SDQ.

Parents were given £10 gift vouchers for participating in the study. Each school was given a £25 book voucher for participating in the study.

3.2.3. *Measures*

3.2.3.1. *Internalising and Externalising Behaviour Problems: The Strengths and Difficulties Questionnaire (SDQ).* The SDQ (Goodman, 1997) was developed as a screening questionnaire. The SDQ exists in different versions, depending on the

informant (teacher, parent, or self), and age of the child (3 to 4 years, 4 to 16 years). Teachers completed the teacher version SDQ 3 - 4 for nursery children and the teacher version SDQ 4 - 16 for reception class children. Parents completed the parent version SDQ 3-4 for nursery children and the parent version SDQ 4-16 for reception class children.

The SDQ measures behaviour across 5 scales, with 5 items per scale: conduct problems, hyperactivity, emotional problems, peer problems, and prosocial behaviour. The 25 SDQ items are underpinned by DSM-IV (1994) and ICD-10 (World Health Organisation, 1978) classifications of child psychopathology (Goodman & Scott, 1999). The questionnaire is a 2-page questionnaire with the 5 domain items mixed together on the first page. Respondents are asked on the first page to check a box for each item on a three-point scale for Not True, Somewhat True or Certainly True. The second page asks respondents to check the relevant box on items of impairment and interference in relation to any difficulties identified on page 1 of the questionnaire. Each of the 5 scales has a score range of 0-10 if all items are completed. A total difficulties score can also be yielded by summing the scores for all scales except the prosocial scale. The SDQ scores are used as continuous variables in previous studies (Goodman, 2001; Goodman & Scott, 1999). For the purposes of this study two dimensional scales were created from the SDQ scales, an internalising scale which included the emotional problems items, and an externalising scale which included the conduct and hyperactivity problems. Table 3.2 outlines the items used in each SDQ scale for this study.

Table 3.2

SDQ Internalising and Externalising Scales and Items

SDQ Scale	Scale Items
Internalising Problems	
Emotional problems	Often complains of headaches, stomach-aches or sickness (Somatisation)
	Many worries, often seems worried
	Nervous or clingy in new situations, easily loses confidence
	Often unhappy, downhearted or tearful
	Many fears, easily scared
Externalising Problems	
Conduct problems	Often has temper tantrums or hot tempers[1]
	Generally obedient, usually does what adults request[2]
	Often fights with other children or bullies them
	Often argumentative with adults
	Can be spiteful to others
Hyperactivity problems	Overactive/restless cannot stay still for long
	Constantly fidgeting or squirming
	Thinks things out before acting[2]
	Easily distracted, concentration wanders
	Sees tasks through to the end, good attention span[2]

[1]Items conceptually related to irritability and removed for purified scale
[2]Items reverse scored

The SDQ has been used in a nationwide British epidemiological study of psychopathology in children aged 5 to 15 years (Goodman, 2001), and demonstrates good reliability and validity. Reliability was reported for internal consistency (Cronbach's alpha = .73); cross-informant correlation (mean = .34) and retest stability after 4 - 6 months (mean = .62). Goodman tested the validity of the SDQ scales against clinical review of participant interview reports using DSM-IV diagnoses. SDQ scores above the 90th percentile predicted a raised probability of independently diagnosed psychiatric disorders (mean odds ratio 15.7 for parent scales and 15.2 for teacher scales). The parent SDQ has also been used in a large epidemiological study (N> 10,000) of 4-year-olds in southwest England (Dunn, Deater-Deckard, Pickering, O'Connor, & Golding, 1998). Dunn and colleagues reported the scale alphas in their study as follows: Conduct Problems .59; Emotional Symptoms .68; Hyperactivity .75; Peer Problems .54; and Prosocial .72.

In the present study internal consistency of the SDQ internalising (emotional problems) and externalising (conduct and hyperactivity) scales were tested using Cronbach's alpha for both the original and the pure scales and the results are given in Table 3.3. Inter-rater reliability was also tested using correlations between teachers' and parents ratings. The results are shown in Table 3.4.

Table 3.3

Scale Reliability Alphas for Original and Pure Teacher and Parent SDQ scales

	No. of scale items	Cronbach Alpha
PSDQ internalising	5	.72
PSDQ original externalising	10	.89
PSDQ pure externalising	9	.89
TSDQ internalising	5	.81
TSDQ original externalising	10	.94
TSDQ pure externalising	9	.94

Table 3.4

Inter-correlations among Informants on the SDQ Original Internalising and Externalising Scales

	1	2	3	4
1.Parent internalising	------			
2. Parent externalising	.11	------		
3.Teacher internalising	.38*	-.15	--------	
4.Teacher externalising	.2	.67*	-.05	--------

N = 61. * $p<.005$

3.2.3.2. *Clinical Symptoms of Childhood Psychopathology: The Preschool Age Psychiatric Assessment (PAPA).* The PAPA (Egger et al., 2002; 2006) is an interviewer-based interview schedule derived from the Child and Adolescent Psychiatric Assessment (CAPA; Angold & Costello, 2000). The PAPA includes some significant changes to the

CAPA to make it relevant for the assessment of 2- to 5-year-olds. PAPA items include all DSM-IV (1994) criteria as they are relevant to this age group, plus the items in the Diagnostic Classification: 0-3 (DC: 0-3; Zero to Three, 1994). The PAPA uses operational definitions of each symptom provided in the glossary. Like the CAPA, the PAPA is designed to focus on the previous 3 months of the child's life, as this recall period has been associated with more accurate recall (Angold, Erkanli, Costello, & Rutter, 1996). The PAPA has been examined in a test-retest reliability study on 114 completed interviews of parents with children aged 2- to 5-years (Egger et al., 2006). Egger and colleagues reported the reliability for symptom scales, measured by intraclass correlation, ranged from .56 to .89.

3.2.3.3. *Construction of the PAPA internalising and externalising symptom scales.* In the present study, which was the first to use the PAPA in a British sample, symptom scales were constructed that were relevant to the same domains as the internalising (Separation Anxiety Disorder, SAD; Generalised Anxiety Disorder, GAD; & Major Depressive Episode, MDE), and externalising (Conduct Disorder, CD; & Attention Deficit Hyperactivity Disorder, ADHD) problem scales of the SDQ. Because of its special relevance to the construct of irritability, a separate symptom scale was constructed for Oppositional Defiant Disorder (ODD). Symptom scales were derived from the PAPA interview using DSM-IV-TR (2000), and the interview items used are shown in Table 3.5. Purified symptoms scales were then developed for the PAPA internalising symptom scale and the ODD symptom scales by removing any items that were conceptually linked to irritability. These items are noted in Table 3.5. The internal consistency for each scale was calculated using Cronbach's alpha and the results for both

the original and the pure PAPA symptom scales are given in Table 3.6. All of the scale reliability alphas are within the range of those recommended for psychological constructs (Kline, 1999). The scales with the larger number of items are more likely to have larger scale reliabilities as illustrated by the results of the PAPA symptom scales whether testing the original or purified scale (Cortina, 1993). Overall the scale alphas indicate that the scales should consistently reflect the constructs being measured.

Table 3.5

PAPA Symptom Scales according to DSM-IV-TR Diagnostic Categories and the Relevant PAPA Items

PAPA Symptom Scale	DSM-IV-TR Symptoms	Relevant PAPA Items
PAPA Internalising symptoms:	General Anxiety Disorder (GAD)	Excess worry/anxiety (pcaoi01) Restlessness/keyed up (pcdoi21; pca3i01) Easily fatigued (pcc4i01) Difficulty concentrating (pcc3i01) Irritability (pda8i01)[1] Muscle tension (pcdoi14)
	Separation Anxiety Disorder (SAD)	Recurrent excessive distress (pbf5i01; pbf6i01; pbf7i01) Fear of losing or harm (pbe8i01) Calamitous separation (pbe9i01) Fear of school/day care (pbd8i01; pbd9i01; pbf9i01; pbg9i01; pbg3i01;pbg6i01) Fear of being alone (pbf4i01) Physical symptoms (pbgli01)
	Major Depressive Episode (MDE)	Depressed mood (pdaoi01) Looks unhappy (pdgoi01) Tearfulness (pda4i01) Loss of interest (pdbli01) Anhedonia (pdb2i01) Motion slowing (pdb4i01) Fatigue/loss of energy (pdb3i01) Worthlessness/ guilt – unloved (pdcoi01) Self-deprecation (pdcli01) Pathological guilt (pdc3i01) Depressive thoughts – sorry for oneself (pdc2i01) Helplessness (pdc6i01) Hopelessness (pdc7i01) Loneliness (pdb9i01)
PAPA Externalising symptoms	Conduct Disorder (CD)	Aggression – Initiates physical fights (pge5i01), uses weapon (pge8e01) cruel to animals (pgh2i01) cruel to people (pgf3e01) Destruction of property- Vandalism (pge2i01) Deceitfulness/theft- Often lies (pgc3i01) Stolen (pgl8e01) Violations of rules – Run away from home (pgc8e01)

[1]PAPA items that are conceptually related to irritability

PAPA Symptom Scale	DSM-IV-TR Symptoms	Relevant PAPA Items
	Attention Deficit Hyperactivity Disorder	Inattention- Inattention general (prb5i01) Fails close attention to detail (prb4i01) Sustained attention in tasks (pra7i01) Follow through instructions (pra8i01) Difficulty organising tasks (prc7i01) Avoids tasks that require sustained attention (prc5i01) Loses things (prb2i01) Forgetful (prc8i01) Difficulty concentrating on adult directed tasks (pra7i02) Easily distracted (pra9i01) Doesn't listen (prb3i01) Hyperactivity- Fidgets (praoi01) Uncontrollable fidget across situations (pra1i01) Always on the go (prc4i01) restlessness (pra4i01) difficulty remaining seated (pra2i01) Runs and climbs excessively (pra3i01) Difficulty doing things quietly (pra6i01) Talks excessively (pra5i01) Impulsivity- Acts before thinking (prc2i01) impulsivity (prc3i01) Difficulty waiting turns (prb7i01) Often interrupts (prc1i01) Blurts out answers (prb8101) Accident prone
PAPA Oppositional Defiant Disorder (ODD) symptoms	ODD	Loses temper (pgeoi01)[1] Temper tantrums (pge1i01)[1] Touchy/easily annoyed (pda6i01)[1] Angry or resentful (pda7i01)[1] Spiteful or vindictive(pga3i01) Argues with adults (pgl5i01) Deliberately annoys people (pga2i01) Blames others (pgj3i01) Teasing (pgl6i01) Defies/Refuses to comply (pgl2i01)

[1]PAPA items that are conceptually related to irritability

Table 3.6

Scale Reliability Alphas for Original and Pure PAPA Symptom Scales

	No. of scale items	Cronbach Alpha
PAPA ODD original	10	.78
PAPA ODD pure	6	.68
PAPA externalising	33	.95
PAPA original internalizing	30	.85
PAPA pure internalising	28	.80

Parents' reports on the questionnaire were consistent with their responses to the PAPA interview. Significant cross-instrument agreement was found between parents' SDQ reports of emotional problems and the PAPA internalising symptom scale, r (61) = .53, $p < 0.05$; and parents' SDQ reports of externalising problems and the PAPA externalising symptom scale and PAPA ODD symptoms respectively, r (61) = .67, $p < 0.05$, and r (61) = .61, $p < 0.05$. Significant cross-informant, cross-instrument agreement was found between teachers' SDQ reports for externalising problems and PAPA externalising symptom scales, r (61) = .57, $p <$ 0.05, and PAPA ODD symptoms, r (61) = .36, $p < 0.05$. Agreement between teachers' SDQ reports of emotional problems and the parent PAPA reports of internalising symptoms approached significance, r (61) = .23, p= 0.07.

3.2.3.4. Construction of the irritability scale. An irritability composite scale was derived from 7 PAPA interview items, in line with the operational definition of irritability used in the temperament literature (Goldsmith, 1996; Rothbart, 1981; 1996; Buss & Plomin, 1984) for

children in this age group. Rothbart has two measures that straddle this age group the Toddler Behaviour Assessment Questionnaire (TBAQ; Goldsmith 1996) and the CBQ (Rothbart, 1996). The anger/frustration dimension on both the TBAQ and the CBQ are the dimensions that Rothbart has mapped onto the 'distress to limitations' dimension of the IBQ. The anger/frustration dimension is described as 'negative affect related to interruption of ongoing tasks' and assessed on questionnaire with items such as, gets angry, has temper tantrums, gets frustrated, easily frustrated, easily irritated, gets irritated. The irritability scale from the PAPA interview questions measured the parents' reports on the child's behaviours across different settings for irritable mood, touchy or easily annoyed, angry or resentful, easily frustrated, loses temper, and temper tantrums. There was good scale reliability for the irritability scale with Cronbach's alpha = .78.

3.2.4. *Data Analysis*

Firstly, scales were derived from the questionnaire for internalising and externalising problems, as described above. Secondly, scales were derived from the interview for irritability and DSM-IV symptom scales that conceptually matched the internalising and externalising scales of the SDQ. The distribution of scores on all the scales used in the present study was examined for normality and heterogeneity of variance through examination of graphical distributions and computation of z-scores to test the significance of any skewness and kurtosis. The SDQ composite scales are usually treated in the literature as continuous measures (Goodman 1997; 2001), and, subject to normality tests, were therefore treated as such in this study. SDQ data are therefore analysed using parametric analyses. The irritability scale and PAPA symptom scales have been derived from PAPA interview items that are measured on an ordinal scale (0, 1, 2, 3) for presence of symptoms. It is anticipated that many children in a non-clinical sample will get

zero scores; therefore the irritability and PAPA symptom scales were examined for normality and heterogeneity, and transformed as necessary for parametric analyses.

Relationships between irritability and internalising or externalising problems on the SDQ were tested using parametric correlations for both the original and pure scales (i.e., with items relevant to the construct of irritability taken out). The comparison of relationships between irritability and the pure and original SDQ scales were tested using a measure of differences between non-independent correlations, known as the Williams formula, T_2 (Williams, 1959), as recommended in the psychology and statistics literature (Steiger, 1980). The same formula was used to test differences in correlations between irritability and the pure and original symptom scales derived from the PAPA. To explore further the role of irritability and ODD as potential mediators between the relationship of internalising and externalising symptoms, multiple regression analysis was carried out on the data.

3.2.4.1. *Irritability Composite Scale.* The irritability scale scores were examined for distributional properties. The range of possible irritability scores was 0 – 21, based on the PAPA symptom scoring procedure of 0 = no symptom present, 1 = partial evidence of symptom present, 2 = definitive evidence of symptom present, 3 = many examples of symptom present. The mean irritability score was 4.38 (SD = 3.90; N = 61), with a range 0 – 17. The distribution of irritability scores indicated some positive skew (1.38) and some concentration of scores at the lower range (0 - 4; kurtosis = 1.79). Examination of frequencies and the histogram indicated that the irritability scale could be treated as normal and therefore it was analysed as a continuous variable using parametric analyses.

3.2.4.2. *SDQ Internalising and Externalising Scale Scores.* The SDQ internalising and externalising scale scores were examined for normality and heterogeneity; the skewness,

kurtosis and *z*-scores are shown in Table 3.7. Whilst there is some positive skewness on both the internalising and externalising scales, examination of the frequency distribution on the scales is comparable to those reported for the SDQ for UK population norms for 5-10 year-olds (Meltzer, Gatman, Goodman, & Ford, 2000). The SDQ scales were therefore treated as normal and analysed using parametric analyses in this study.

Table 3.7

Normality tests for the Parent and Teacher SDQ Internalising and Externalising Scores

SDQ Scale	Skewness	Skewness z-score	Kurtosis	Kurtosis z-score
TSDQ Internalising	1.01	4.05	-0.11	-0.21
TSDQ Externalising	0.99	2.48	0.75	0.22
PSDQ Internalising	1.54	5.22	2.1	3.61
PSDQ Externalising	0.73	2.48	0.13	0.22

3.2.4.3. PAPA Symptom Scales. The PAPA symptom scales were also examined for normality and heterogeneity. All the symptoms scales showed a significant skew and were therefore transformed using a square root transformation consistent with other studies using symptom scales (Hudziak et al., 2004). Examination of the Q-Q plots for the transformed variables indicated that the transformed data for the PAPA symptom scales were approximately normal and therefore parametric analyses were used with the transformed data.

3.3. Results

3.3.1 *Preliminary Tests for Gender Differences*

Descriptive data are presented separately for girls and boys in Tables 3.8 to 3.13. Significant gender differences were found in teachers' reports on the SDQ but not for parents' reports on the SDQ or PAPA. Teachers reported boys as having significantly more conduct and hyperactivity problems than girls, but did not report any significant difference between boys and girls for emotional problems measured using the SDQ.

The pattern of association between the SDQ scale scores were similar for boys and girls, regardless of the informant reporting on the SDQ. Both teachers and parents reported a significant association between conduct and hyperactivity problems, and no significant association between conduct and emotional problems, and no significant association between hyperactivity and emotional problems. There was no significant difference found between girls and boys irritability.

The difference between informants was not due to statistical power; teachers' reports on the subsample of children who participated in Phase 2 also revealed gender differences not found in parents' reports on the same subsample (for more details, see Hay, Hudson, & Liang, in press). In view of the sample size and inconsistency across informants in detecting gender differences, subsequent analyses are conducted on the whole sample.

Table 3.8

Phase 1. Teachers' ratings of girls' and boys' behaviour on the Strengths and Difficulties Questionnaire (SDQ)

		Conduct Problems*	Hyperactivity Scale*	Externalising Scale* (conduct & hyperactivity)	Internalising Scale (Emotional problems)
Girls	*M*	1.23	3.07	4.30	2.02
	SD	1.74	2.40	3.67	2.04
Boys	*M*	2.49	4.97	7.46	1.80
	SD	2.80	2.78	5.18	2.04
Total	*M*	1.71	3.79	5.50	1.94
	SD	2.28	2.70	4.55	2.23

Note. *$p < .05$ by univariate tests

Table 3.9

Phase 1. Parents' ratings of girls' and boys' behaviour on the Strengths and Difficulties Questionnaire (SDQ)

		Conduct Problems	Hyperactivity Scale	Externalising Scale (conduct & hyperactivity)	Internalising Scale (emotional problems)
Girls	*M*	2.24	3.45	5.68	1.87
	SD	2.21	2.66	4.55	2.24
Boys	*M*	3.25	4.64	7.90	1.54
	SD	2.58	2.75	4.85	1.79
Total	*M*	2.67	4.00	6.62	1.73
	SD	2.41	2.74	4.77	2.06

Note. *$p < .05$ by univariate tests

Table 3.10

Phase 1. Comparison of pattern of association of teacher-reported SDQ scale scores for girls and boys

	Conduct Problems	Hyperactivity Problems	Internalising Scale (emotional problems)	Externalising Scale (conduct & hyperactivity)
Conduct Problems	---	.56*	-.15	.84*
Hyperactivity Problems	.72*	----	-.04	.92*
Emotional Problems	-.17	-.03	---	-.10
Externalising Scale (conduct & hyperactivity)	.93*	.47*	-.11	---

Note: Pearson *r* correlation coefficients are presented above the diagonal for girls, below the diagonal for boys.

+ $p < .10$, * $p < .05$

Table 3.11

Phase 1. Comparison of pattern of association of parent-reported SDQ scale scores for girls and boys

	Conduct Problems	Hyperactivity Problems	Internalising Scale (emotional problems)	Externalising Scale (conduct & hyperactivity)
Conduct Problems	---	.75*	.24	.92*
Hyperactivity Problems	.66*	---	.04	.95*
Emotional Problems (Internalising scale)	-.20	.05	---	.14
Externalising Scale (Conduct & Hyperactivity)	.90*	.92*	-.07	---

Note: Pearson *r* correlation coefficients are presented above the diagonal for girls, below the diagonal for boys.

$+ p < .10$, $* p < .05$

Table 3.12

Phase 2. Parents' reports of girls' and boys' symptoms of clinical disorders on the PAPA interview

		Conduct Symptoms	ADHD Symptoms	PAPA Externalising Symptoms	ODD Symptoms	PAPA Internalising Symptoms
Girls	*M*	2.39	13.27	15.67	5.97	5.15
	SD	2.34	14.11	15.60	4.24	6.23
Boys	*M*	3.61	13.79	17.40	7.71	5.00
	SD	3.88	16.22	18.50	5.10	7.36
Total	*M*	2.95	13.51	16.50	6.77	5.08
	SD	3.18	14.99	16.86	4.70	6.71
	Range	0 to 16	0 to 56	0 to 72	0 to 21	0 to 32

Note: Descriptive statistics are based on DSM-IV symptom scales, in which symptoms definitely present are given a score of 2 and symptoms sometimes present a score of 1. Scales differ in number of items, due to differing operational criteria for the disorders.

No significant sex differences were found.

Table 3.13

Phase 2. Mean irritability scores for girls and boys

	N	M	SD
Girls	33	3.58	3.53
Boys	28	5.32	4.16

No significant sex differences were found.

3.3.2. *Testing the Measurement Confounding Hypothesis for Irritability and the SDQ Externalising Problem Scales*

When both parents' and teachers' original externalising SDQ scales were examined in relation to the irritability scale, irritability was significantly associated with the externalising problem scale for both informants. The results are shown in Table 3.14.

Table 3.14

Inter-correlations among Original and Purified Teacher SDQ Scales, Parent SDQ Scales and PAPA Symptom Scales

	1	2	3	4	5	6	7	8	9	10	11	12
1. Irritability Composite	---											
2.TSDQ internalising	-.03	---										
3.TSDQ externalising original	.47***	-.11	---									
4.TSDQ externalising pure	.40 **	-.12	.99***	---								
5.PSDQ internalising	.11	.46***	.02	-0.03	---							
6.PSDQ externalising original	.65***	-.20	.63***	.63***	.11	---						
7.PSDQ externalising pure	.62***	-.19	.65***	.60***	.10	.94***	---					
8.PAPA ODD original	.88***	-.01	.36**	.30*	.11	.61***	.57***	---				
9.PAPA ODD pure	.58***	-.02	.41**	.38**	.11	.62***	.60***	.65***	---			
10.PAPA internalising original	.49***	.23	.13	.08	.50***	.18	.20	.43**	.33*	--		
11.PAPA internalising pure	.41**	.26*	.11	.07	.56***	.13	.16	.37**	.30*	.99***	---	
12.PAPA externalising	.65***	-.004	.57***	.53***	.31*	.67***	.75***	.61***	.56***	.33*	.32*	---

N=61.
*p<0.05
**p < 0.005
***p < 0.0005

An analysis of the difference between the correlations obtained from the original and the pure SDQ externalising scales was carried out using the Williams (1959) formula. There was no significant difference in the correlations between the original and the purified parent SDQ externalising scales, t (58) = -0.46, n.s., but a significant difference between the original and the purified teacher SDQ externalising scale, t (58) = -3.10, $p<.05$.

3.3.3. *Testing the Measurement Confounding Hypothesis for Irritability and the Internalising PAPA Symptom Scales*

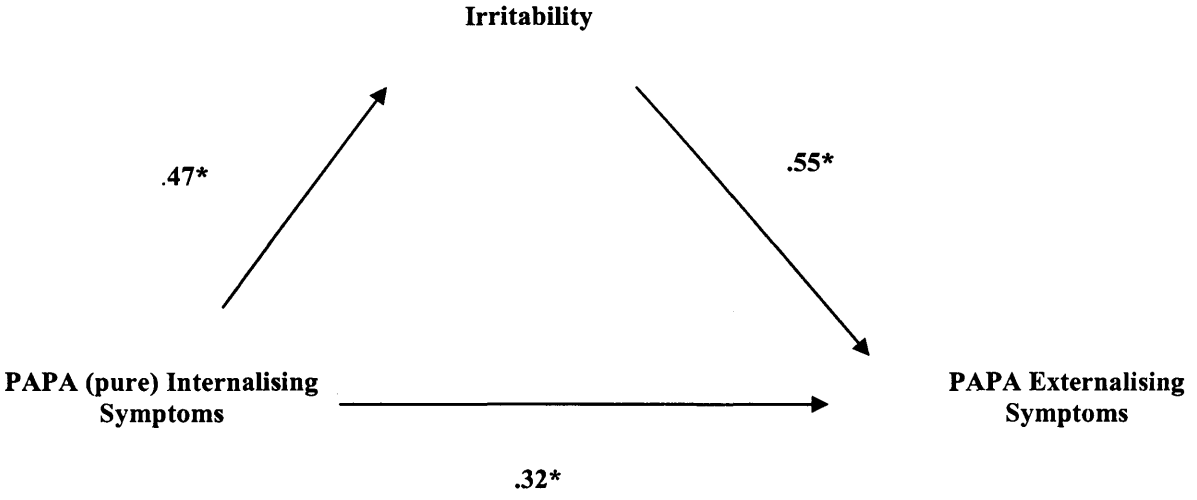
The correlations between the irritability scale and the internalising PAPA scales are shown in Table 3.14. There were significant correlations between the irritability scale and both the original and the purified internalising PAPA symptom scales. Analysis of the difference between the significant correlations between irritability and the original and the purified internalising PAPA symptom scales was carried out using the Williams T_2 formula. A significant difference was found between the two relationships t (58) = -2.26, $p<.05$.

3.3.4. *Testing the Measurement Confounding Hypothesis for Irritability and the ODD PAPA Symptom Scales*

The relationship between the irritability scale and the original and purified ODD PAPA symptom scales are shown in Table 3.14. A significant relationship was found for both the original and purified ODD symptom scale. Using the Williams T_2 formula, the difference between the correlations for irritability with the original and purified ODD symptom scales was significant, t (58) = -7.86, p <.001.

3.3.5. *Examination of the Role of Irritability as a Potential Mediator in the Relationship between Internalising and Externalising Symptoms*

The Baron and Kenny (1986) model was used for testing the mediating relationship of irritability between the presence of PAPA internalising symptoms and the presence of PAPA externalising symptoms. The mediational path model is illustrated in Figure 3.1.



* significant at $p < 0.05$

Figure 3.1. A mediating path model: Influence of PAPA (pure) internalising symptoms and irritability on PAPA externalising symptoms.

To show a mediating relationship Baron and Kenny claim that three conditions have to be met. The three paths connecting the three variables under investigation should demonstrate significant relationships between the variables. Figure 3.1 sets out how the proposed mediating relationship model meets the criteria. The path between the PAPA (pure) internalising symptoms and PAPA externalising symptoms shows a significant relationship, the path connecting the PAPA (pure) internalising symptoms and irritability shows a significant relationship, and the path between irritability and the PAPA externalising symptoms shows a significant relationship. The final step in testing a mediating path consists of demonstrating that when irritability as the mediator is used simultaneously with the PAPA (pure) internalising symptoms (independent variable) to predict the PAPA externalising symptoms (dependent variable), the previously significant path between the PAPA (pure) internalising symptoms and the PAPA externalising symptoms should no longer be significant. An hierarchical regression was used to test the mediating path, using irritability and PAPA (pure) internalising symptoms as the predictors of PAPA externalising symptoms. The results are shown in Table 3.15.

Table 3.15

Summary of Hierarchical Regression Analysis for Variables Predicting PAPA Externalising Symptoms (N = 61)

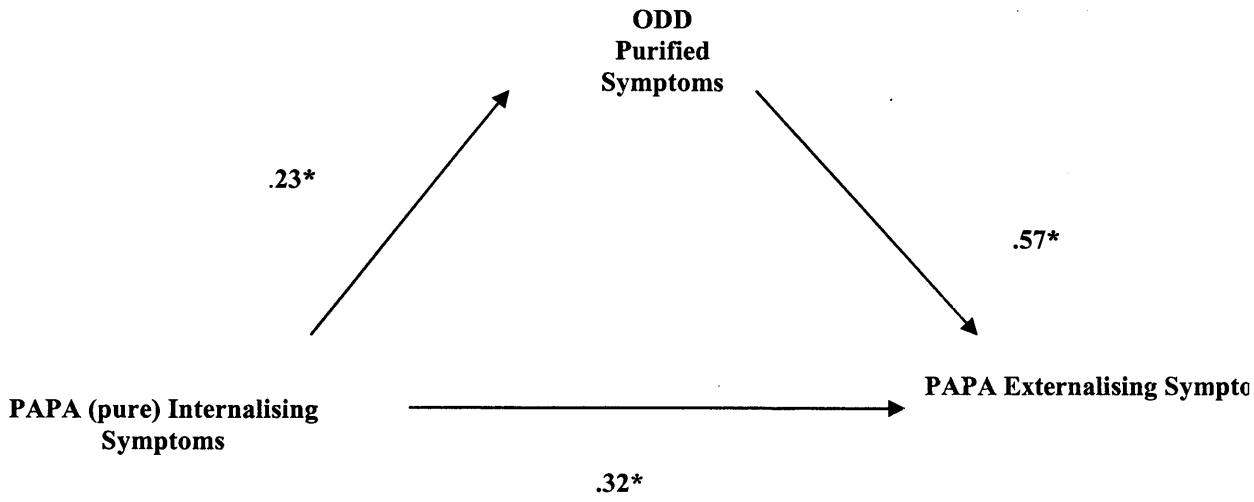
	B	*SE B*	*β*
Step 1			
PAPA internalising symptoms	.44	.17	.32*
Step 2			
PAPA (pure) internalising symptoms	.10	.17	.07
Irritability	.26	.06	.52*

Note. $R^2 = .10$ for Step 1; $\Delta R^2 = .21$ for Step 2 ($ps < .05$).
*$p < .05$.

The first model in Table 3.15 uses PAPA (pure) internalising symptoms as the sole predictor. The second model adds irritability as a predictor. When irritability is added to PAPA (pure) internalising symptoms, the PAPA internalising symptoms are no longer significant predictors of the PAPA externalising symptoms, $t = 0.57$, $p = .60$. Irritability appears to act as a mediator between PAPA (pure) internalising symptoms and PAPA externalising symptoms. As these are all contemporaneous variables further mediating paths were considered but did not meet the Baron and Kenny criteria to remove the significant relationship between the independent and dependent variables.

3.3.6. *Examination of the Role of ODD as a Potential Mediator in the Relationship between PAPA Internalising and PAPA Externalising Symptoms*

The potential role of ODD purified symptoms as a mediator in the relationship between PAPA (pure) internalising and PAPA externalising symptoms was also tested using the Baron and Kenny (1986) model. Figure 3.2 illustrates that the three conditions are met for the three relationships in the model, with all relationships significant.



* significant at $p < 0.05$

Figure 3.2. A mediating path model: Influence of PAPA (pure) internalising symptoms and ODD purified symptoms on PAPA externalising symptoms.

For the next stage in the model to test the role of ODD (pure) symptoms as a mediator in the relationship between the PAPA (pure) internalising and PAPA externalising symptoms hierarchical regression was used with ODD purified symptoms and PAPA (pure) internalising symptoms used as predictors of PAPA externalising symptoms. The results are shown in Table 3.16.

Table 3.16

Summary of Hierarchical Regression Analysis for Variables Predicting PAPA Externalising Symptoms (N = 61)

	B	*SE B*	*β*
Step 1			
PAPA internalising symptoms	.44	.17	.32*
Step 2			
PAPA internalising symptoms	.27	.15	.19
ODD purified symptoms	1.26	.26	.52*

Note. $R^2 = .10$ for Step 1; $\Delta R^2 = .26$ for Step 2 ($ps < .05$).
*$p < .05$.

The first model in table 3.16 uses PAPA (pure) internalising symptoms as the sole predictor. The second model adds ODD (pure) symptoms as a predictor. When ODD (pure) symptoms is added to PAPA (pure) internalising symptoms, the PAPA internalising symptoms are no longer significant predictors of the PAPA externalising symptoms, $t = 1.79$, $p = .08$. ODD purified symptoms appear to serve as a mediator between PAPA (pure) internalising symptoms and PAPA externalising symptoms.

As irritability and ODD (pure) symptoms were contemporaneous measures, these two variables were also tested together as predictors of PAPA externalising symptoms. Whilst the first 2 conditions of the Baron and Kenny mediator model were met, the crucial stage of removing the significance between the independent and the dependent variable when ODD or irritability were used as the potential mediators was not successful. ODD purified symptoms did not appear to act as a mediator in the relationship between irritability and the PAPA externalising symptoms and irritability did not appear

to act as a mediator in the relationship between ODD purified symptoms and the PAPA externalising symptoms.

3.4. Discussion

This present study adds to the current literature on the important relationship between temperament and psychopathology by testing the measurement confounding hypothesis for a specific temperament dimension, 'irritability', and its relationship with both behavioural problems and symptoms of disorder, using a psychiatric based interview. Previous questionnaire-based studies have examined the measurement confounding hypothesis between irritability and problem behaviours (Lemery et al., 2002), but to my knowledge no study has examined the measurement confounding hypothesis between irritability and symptoms of disorder recorded as part of a psychiatric interview.

Previous studies on measurement confounding have used expert judgements and empirical methods to identify and remove any potential confounding items between temperament and behaviour measures (Lemery et al., 2002; Lengua et al., 1998; Sanson et al. 1990). Lemery and colleagues concluded that there was little correspondence between the two methods of identifying confounded items, and emphasis should be put on developing measures on a conceptual basis through clarification of constructs. Within the present study irritability was measured using a scale derived according to an operational definition from Rothbart's psychologically-based theory of temperament. Any potential items and symptoms fitting this definition were removed from the SDQ behaviour problems questionnaire and the PAPA psychiatric based parent interview. This approach allowed the construct of irritability measured on a reliable scale ($\acute{\alpha} = .78$) to be

tested in relation to both questionnaire measures of behavioural and emotional problems and symptoms of childhood disorders.

Scales for internalising and externalising problem behaviours and symptoms were derived from the SDQ and the PAPA respectively. The SDQ internalising scale included behaviours that were then reflected in the PAPA internalising symptom scale, and the SDQ externalising scale included behaviours that were then reflected in the PAPA externalising scale. ODD symptoms were examined separately using the PAPA interview because of the number of similar items included in the PAPA symptom scale of ODD and the operational definition of irritability.

This study used a simple empirical test, the Williams formula (1959), to test the magnitude of differences between the correlations for irritability with the original scales and the purified scales (decontaminated of possible irritability confounding items). A significant difference would suggest that measurement confounding contributes to the relationship between irritability and the relevant scale tested. The externalising scale of the SDQ had one item that reflected irritability, *temper tantrums*, and thus the relationships between irritability and the original and pure externalising SDQ scales were compared. The SDQ internalising scale did not contain any items that were considered to be potential confounds with irritability. The internalising PAPA symptom scale had items that reflected irritability as detailed in Table 3.5, and the relationships between irritability and the original and pure internalising PAPA symptom scales were thus compared. A symptom scale for ODD was derived separately from the PAPA externalising symptom scales to allow specific analyses of the relationship between irritability and ODD. Three out of 10 symptoms for ODD were considered to reflect irritability and thus the

relationship between irritability and an original and pure symptoms scale for ODD were compared. The PAPA externalising symptom scale reflected symptoms of CD and ADHD that did not contain any items reflecting irritability. Decontamination of the SDQ externalising problems scale, the PAPA internalising scale and the PAPA ODD did not result in a loss of scale reliability, with the range of scale reliabilities given as α =.68 to .94, which compared favourably to those within other studies (Lemery et al., 1999; Lengua et al., 1998).

Testing the measurement confounding hypothesis for the relationship between irritability and the PAPA internalising symptom scales revealed a significant difference in the magnitude of the relationship after the potential confound items were removed. These results suggest that there was some measurement confounding occurring between the measurement of irritability and the measurement of items on the PAPA internalising symptom scales. The PAPA original internalising symptom scale included irritability as a symptom of anxiety disorders and items that fit the irritability operational definition as part of depressive symptoms, but even when these items were removed from the PAPA internalising scale, irritability continued to be significantly associated with the PAPA internalising symptoms scale. This result is particularly important in relation to the results from the study by Lengua and colleagues. When Lengua and colleagues examined the relationship between negative emotionality and depression symptoms, the irritability items were removed from the temperament measure and thus the relationship that remained between negative emotionality and depression symptoms was explained in relation to the fear and sadness dimensions of negative emotionality. Within the present study the theoretically derived measure of irritability continued to relate to the PAPA

internalising symptoms scale when any potential confounding items were removed from the internalising symptom scales suggesting that irritability was related to symptoms such as fear and sadness that remained as part of the PAPA (pure) internalising symptom scale. This result suggests that the vulnerability model may be more relevant to explain the relationship between irritability and internalising symptoms. Further longitudinal studies would be required to test this hypothesis.

When the measurement confounding hypothesis was tested for the relationship between irritability and the SDQ externalising scale different results were revealed for the parent and teacher reports. A significant difference was revealed for the parents' reports between the magnitude of association with irritability and the original and purified SDQ externalising scales, but no significant difference was found in the magnitude of correlations for the teacher SDQ externalising reports. This reflects similar differences between informants found by Lengua and colleagues (1998). In Lengua's study decontamination of the negative emotionality scale for mother reports resulted in a significant but decreased relation between negative emotionality and conduct symptoms; in contrast, for self-reports from the child, decontamination of the negative emotionality scale resulted in a reduced and non-significant relation between negative emotionality and conduct problems (Lengua et al., 1998). This result emphasises the importance of measuring behaviour problems across contexts using different informants.

From the results of this present study and from the study by Lengua and colleagues measurement confounding appears to be influential in the relationship between irritability and externalising problems, although this appears to be dependent upon informant. The relationships between irritability and the purified externalising

problems remained significant within this present study for both teachers' and parents' reports using the SDQ, and suggest that the measurement confounding hypothesis does not provide the complete answer to the relationship. Lemery and colleagues reported no significant difference when comparing the relationships between the temperament and behaviour problem symptoms using either original or purified scales (Lemery et al., 2002). Psychologists had judged some irritability items to be potential confounds between the CBQ temperament scale and the PBQ behavioural problems scale but this was not borne out in the subsequent analyses (Lemery et al.). From the present results it appears that there may be some conceptual overlap between irritability and items used to describe externalising problems on the SDQ but there remains a relationship between irritability and externalising problems that warrants further investigation.

Examination of the relationship between irritability and the PAPA externalising symptom scale revealed a significant correlation, $r = 0.55$, $p < 0.05$, but the PAPA externalising symptom scale did not contain any potentially confounded items and therefore further analysis in relation to measurement confounding was not required. The PAPA externalising symptom scales included symptoms for Conduct Disorder and Attention Deficit Hyperactivity Disorder, and the results here suggest that there is a significant relationship between irritability and symptoms of these disorders. This relationship between irritability and externalising symptoms of conduct disorders and ADHD suggests that the vulnerability model may provide an explanation, but further longitudinal studies would be required to confirm this.

Looking specifically at the PAPA ODD symptom scales, there were three symptom items that were conceptually similar to irritability. When these items were

removed from the ODD symptom scale the association between irritability and the purified ODD symptom scale remained, but there was a significant difference in the magnitude of the association compared with the original ODD symptom scale. These results suggest that measurement confounding may be a factor between items used to measure irritability and items used to measure symptoms of ODD. This does not mean that irritability and its relationship with symptoms of ODD should not be subject to further investigation, because despite the decrease in the magnitude of the association between the original and the purified measures, the associations between irritability and significant ODD symptoms remained. The SDQ externalising scale included items relevant to ODD (e.g., often argumentative) and thus the finding in this present study of potential confounding between irritability and the parent SDQ externalising scale may be due to the apparent relationship between irritability and symptoms of ODD. It would imply that ODD may be a clinical manifestation of irritability. The relationship between irritability and ODD was tested further in the present study using mediational models to understand the role that irritability plays in the relationship between internalising and externalising symptoms and the role that ODD (pure) symptoms play in the same relationship.

The issue of measurement confounding is not unique to studies of temperament and psychopathology, but is also an issue grappled with in many other areas of research, including comorbidity of child psychopathology (Angold, Costello, & Erkanli, 1999). Similar issues will therefore arise, such as the need to further understand the underlying mechanisms for the remaining associations between temperament and symptoms of comorbid disorders. Rather than dismiss results that indicate potential measurement

confounding we should learn from the work on comorbidity and consider the possibility of ODD perhaps as an extreme manifestation of irritability with the relationship representing a hierarchical model of dimensions of temperament and psychopathology, such as the spectrum model (Nigg, 2006).

It is through examination of co morbidity that Egger and Angold (2006) recently reported ODD as central in mediating the relationships between depression and CD, and the emotional disorders and ADHD. Within this present study ODD symptoms have been examined further to understand the potential mediating role that ODD may hold in the relationship between internalising and externalising symptoms. Whilst the present analyses used symptom scales rather than categorical diagnoses, the results here indicate that ODD (pure) symptoms may indeed be a mediator in the relationship between internalising and externalising symptoms.

Co morbidity studies have also revealed the importance of irritability as a symptom of internalising and externalising problems (Luby et al., 2003). Taken together with the reported relations between irritability and both internalising and externalising problems, it was essential that the present study tested the role of irritability as a potential mediator in the relationship between internalising (pure) and externalising symptoms. The results from this present study indicate that irritability does play a mediator role in the relationship between internalising and externalising symptoms.

After testing the measurement confounding hypothesis for the relationship between irritability and questionnaire and interview measures for symptoms of internalising and externalising problems, the results from this study suggest some degree of measurement confounding between measures of irritability when defined within the

temperament framework and measures of internalising symptoms, and externalising problems, but particularly with ODD symptoms. Whilst this measurement confounding does appear to be an issue that future studies of temperament and disorder should consider as part of the methodology, the significant relationships that remain between irritability and decontaminated measures of internalising symptoms, externalising symptoms, and ODD suggest that studying irritability at an early stage of development would be worthwhile if we are to understand better the pathway between irritable temperament and psychopathology in children.

The strong relationship between irritability and ODD symptoms demonstrated in this study were tested further by examining the independent roles that irritability and ODD (pure) symptoms play in the relationship between internalising (pure) and externalising symptoms. Both irritability and ODD (pure) independently mediated the relationship between the decontaminated internalising symptoms and the externalising symptoms. These results suggest that whilst ODD may have irritability symptoms that would suggest that it may be a clinical manifestation of irritability (the spectrum model), the independent influence that both ODD pure and irritability play in the relationship between internalising and externalising symptoms, suggest that irritability is more likely to be a vulnerability factor in the development of externalising problems such as conduct disorder. This finding fits with the results of Egger and Angold (2006) who found that ODD mediated the relationship between depression and conduct disorder in preschoolers. In Egger and Angold's study ODD contained the symptoms that reflect irritability and may have masked irritability's role (as opposed to the broader profile of symptoms

comprised in the diagnostic category of ODD) in mediating the relationship between sadness in depression and conduct disorder. This hypothesis needs further investigation.

Having identified the important role that irritability appears to play in the relationship with childhood internalising and externalising symptoms, the next stage of this thesis is to consider the manifestation of irritability in infancy in relation to other early infant behaviours. Investigation of potential shared risk factors may also help to reveal further the relationship between irritability and symptoms of disorder in childhood. Both intergenerational transmission and the social context have been identified as influential in the development of disorders in children (Egger & Angold, 2006), and are two processes that also influence the development of temperament (Plomin, 2006). As a further step in understanding the pathway from irritability to disorder, the next study detailed in Chapter 4 will examine the relationship between irritability and disorder at an adult stage in the lifespan, and examine the nature of irritability in infants, and the potential intergenerational transmission of irritability between mothers and infants.

CHAPTER 4

STUDY 2

Understanding the Role of Irritability in Adult Disorders, and the Potential Intergenerational Transmission of Irritability from Mother to Infant

---

4.1. Aims of Study 2

Study 1 (described in Chapter 3) provided evidence that irritability was a contributing factor to the positive association between emotional and behavioural symptoms in preschool children, suggesting the importance of irritability as a factor in the pathway to both internalising and externalising disorders in childhood. This in turn suggests that the co morbidity between internalising[1] and externalising[2] disorders may derive from irritable temperament.

Studies about irritable temperament usually focus on the early stage of development in children (mostly in infancy) and longitudinal studies about irritability as a potential symptom of disorder predict from early to later childhood. The relationship between irritability and disorder across the child's development suggests that irritability may be an enduring factor that influences development throughout the lifespan. This is a concern, because studies have shown that early problems in children have stability through to young and later adulthood (Harrington, Fudge, Rutter, Pickles, & Hill, 1990; Loeber & Hay, 1997; Overton, 2004). There is clinical evidence of irritability as an important symptom in adult psychopathology (Snaith & Taylor, 1978; Nigg, 2006), such as Generalised Anxiety Disorder

[1] Emotional is used in place of internalising for adults in Study 2

[2] Behavioural is used in place of externalising for adults in Study 2

and Bipolar Depression (DSM-IV-TR; 2000). The stability of psychopathology across the lifespan, and the importance of irritability as a symptom of psychopathology in both children and adults suggest that irritability itself may be an enduring characteristic across the lifespan.

Within the adult personality literature there is evidence of the stability of angry temperament over the life course (Costa & McCrae, 2001; Caspi, Elder, & Bem, 1987), and a strong relationship between negative affect and neuroticism demonstrated across age groups (Rothbart, Ahadi, & Evans, 2000). From these studies, it is possible to propose that irritability may continue to play a role in the development of psychopathology throughout the lifespan. Genetic studies have highlighted the heredity of irritability (Henderson, 1982; Plomin et al., 1988) and temperament studies have identified similar temperament between mothers and their infants (Vaughn et al. 1987). These observations lead to two hypotheses examined within this present study. Firstly, irritability may be a temperamental liability to psychopathology in adults (see Figure 4.1), and secondly, the relationship between mother and infant irritability may demonstrate an intergenerational transmission of irritability (Figure 4.3), that may indicate the potential risk of later disorder.

The focus of Study 2 is the mother-infant micro-system (see Chapter 1 for Bronfenbrenner's 1977 model of developmental systems), and the developmental transition from pregnant woman to being the mother of a first-born infant. Study 2 begins with the examination of mothers' irritability and the relationship between mothers' irritability and the mothers' own history of disorders. Within this study, we examine mothers' history of emotional and behavioural disorders, namely depression and anxiety disorders and past symptoms of conduct disorder. Then the potential intergenerational transmission of

irritability from mother to infant is explored through the examination of maternal predictors of the infants' irritable temperament. The three specific aims of Study 2 are detailed below:

4.1.1 ***Aim 1: To examine the role of maternal irritability in relation to mothers' own emotional and behavioural problems***

The potential for irritability to be a risk factor in the development of psychopathology over the lifespan has implications for adults' transition to parenthood and the transmission of irritability across generations. Therefore, in Study 2, the mothers' own irritability is examined before and after the birth of her first child, within the context of their past and present depression and anxiety disorders and past history of conduct symptoms.

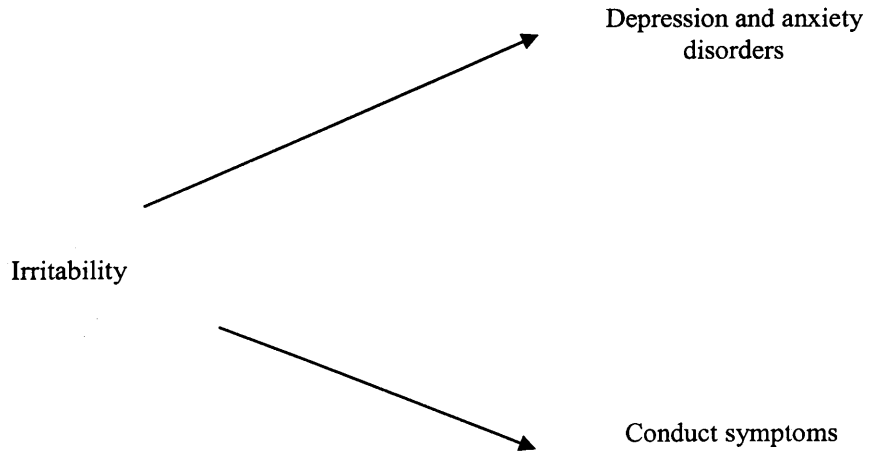


Figure 4.1. The potential temperamental liability of irritability to emotional and behavioural disorders in adults

4.1.2 *Aim 2: To examine infant irritability in relation to other temperament dimensions and behaviours that may indicate early pathways to internalising and externalising problems*

The infant temperament literature is full of studies that use a variety of descriptors for infant irritability, such as emotional negativity, difficultness and negative emotionality, and a variety of measures to assess the same or slightly varied construct. To extend the literature on irritable temperament, I aim to describe infant irritability using a measure that can be traced to a theoretical explanation of the irritability construct (distress to limitations of the IBQ; Rothbart, 1981). Previous temperament studies using the 'distress to limitations' scale have identified a relationship between 'distress to limitations' and other temperament factors, namely, 'distress and latency to approach novel or sudden stimuli' which is an operational dimension to tap fear, and with the activity level dimension (Hane et al., 2006; Rothbart; 1981; Rothbart & Bates, 1998). If irritability is an enduring characteristic that influences disorder throughout the lifespan, it may also be the case that early combinations of irritability with other temperament factors may lead to particular disorders. This specificity argument has been discussed at some length, and particular temperament profiles have been suggested as placing a child at predominant risk for different psychopathologies (Nigg, 2006). The combination of negative affectivity and frequent activity has been predicted to lead to *emotional problems,* i.e., later fearful and anxious behaviours (Kagan & Snidman, 1991). High activity levels in infancy may also predict to *behavioural problems,* i.e., symptoms of ADHD. Within this study, I aim to assess infant irritability in relation to infant fear and infant activity to understand the potential temperament combinations in infancy that may

predispose children to later emotional and behavioural disorders. Figure 4.2 illustrates the potential pathways suggested by previous studies.

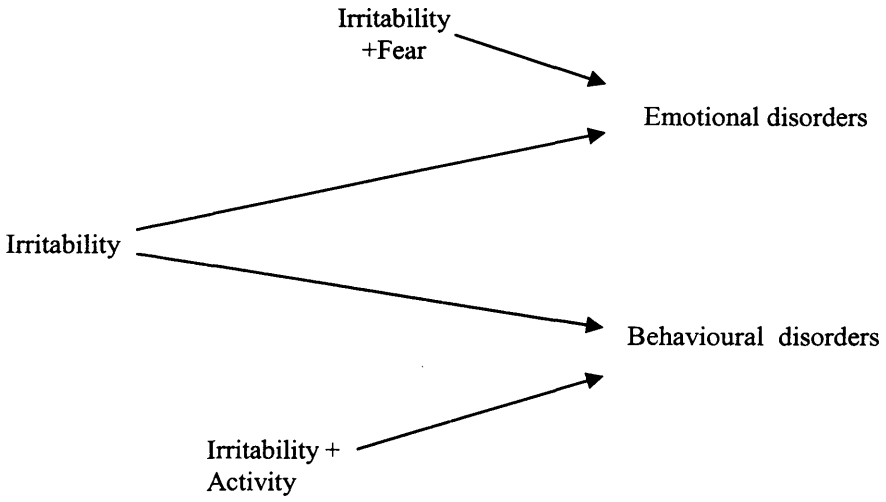


Figure 4.2

The potential early influences of temperament factors on internalising and externalising problems

4.1.3 Aim *3: To examine the intergenerational transmission of irritability from mother to infant*

There is already evidence of maternal anxiety and depression predicting infants' difficult temperament at 4 or 6 months (Austin, Hadzi-Pavlovic, Leader, Saint, & Parker, 2005), but

does this relationship actually reflect an intergenerational transmission of irritability between mother and infant? The continuity or discontinuity of infant irritability has been shown to be dependent upon family factors, such as maternal ones (e.g., Belsky, Fish, & Isabella, 1991). The third aim of this study is therefore to examine the predictive relationship between maternal irritability and infant irritability, within the context of other potential influences such as mothers' mental health and socio-economic circumstances.

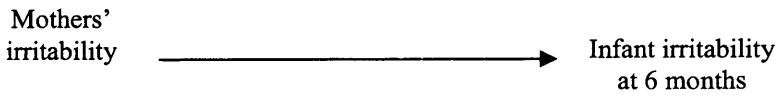


Figure 4.3

The pathway for the potential intergenerational transmission of irritability between mother and infant

Having described the three aims of Study 2, I now discuss the conceptual and methodological issues used to inform the design of Study 2, followed by presentation of the methods and analyses undertaken to meet the three aims of the study.

4.2 Conceptual and Methodological Issues in Studying Mothers' Assessments of Infant Temperament

The methods adopted in Study 2 derive from a consideration of the literature on infant temperament and, in particular, critiques of the use of mothers' reports as primary sources of information about infant temperament, and on the potential maternal predictors of infant irritability.

4.2.1. *Links between Maternal Characteristics and their Reports of Infant Temperament*

When seeking evidence for the intergenerational transmission of irritability from mother to child, it is important to acknowledge that most research on infant temperament has been carried out using reports from mothers (see Chapter 2). Mothers provide information about their own irritability and symptoms of disorder and about their infants' irritability and other dimensions of temperament.

Historically, the choice of mothers as the primary informants about infant temperament makes sense, as the majority of mothers are the main carers during infancy and are therefore likely to know the baby's behaviour better than other reporters. It has also made economic sense as it is usually cheaper to ask the mother as main carer of the infant to complete a questionnaire on the infant's temperament than it is to set up either home or laboratory based observations. Despite the benefits of this common-sense approach, there has been considerable debate in the temperament literature about the reliability and validity of mothers' reports of infant temperament (Kagan, 1994; Rothbart & Goldsmith, 1985; Vaughn, Toraldson, Cuchton, & Egeland, 2002; Wachs & Bates, 2004). There is concern about shared methods variance and also about the possibility that additional measurement error could be introduced to studies using mothers' reports, due to the social desirability of responses, the

limited accuracy of the caregiver's memories, lack of comparison with other children of the same age, the mother's limited knowledge of infant behaviour and its meaning, and variance attributable to the mother's own characteristics before and at the time of completion (Rothbart & Bates, 1998; Crockenberg & Acredolo, 1983).

Important methodological concerns were raised about the use of mother reports on infant temperament during an extensive programme of work carried out by Vaughn and colleagues, who examined the reliability and validity of the ITQ (Carey 1970; Carey, & McDevitt, 1978; Vaughn, Deinard, & Egeland, 1980; Vaughn, Taraldson, Crichton, & Egeland, 1981; Vaughn et al., 2002; Vaughn, Bradley, Joffe, Seifer, & Barglow, 1987). The ITQ is a questionnaire derived from the Thomas and Chess (1968) parent interview of infant temperament, which has been widely used and revised following reports of some psychometric problems with the measure (Rothbart & Hwang, 2002). The ITQ measures nine temperament dimensions and has an algorithm to organise the results on the dimensions into Thomas and Chess temperament diagnostic categories (i.e. easy, intermediate low, intermediate high and difficult). In their first two studies using the original ITQ, Vaughn and colleagues (Vaughn et al.,1980; Vaughn et al., 1981) found that mothers' psychological characteristics measured before the infants were born were related to the mothers' reports of infant difficulty, and that mothers' reports of infant temperament did not relate to observers' reports of the infants during feeding. Subsequent studies using the revised ITQ also revealed that mothers of difficult infants were more anxious, suspicious, and impulsive before birth than mothers of easy infants (Vaughn et al., 1987), and that these same maternal prenatal characteristics did not relate to observed infant behaviours during mother-infant interactions

(Vaughn et al., 2002). Vaughn and colleagues concluded from this programme of work that the ITQ was a better measure of maternal characteristics than of infant temperament.

The work of Vaughn and colleagues highlights important methodological issues to be considered when studying infant temperament. The critique of the ITQ led to much work on developing instruments to measure reports of infant temperament, and there are now measures of infant temperament that have good validity and reliability across measures and between informants (Rothbart & Bates, 1998; Rothbart, Chew, & Gartstein, 2001).The ITQ is only one of a variety of questionnaires available to measure infant temperament. Chapter 2 outlined the measures available and, within the present study, the IBQ (Rothbart, 1981) has been chosen as the questionnaire to measure infants' irritability.

Vaughn and colleagues reached their conclusions about the reliability of the ITQ without using different informant reports on the ITQ. To extend the literature on the reliability of mothers as informants about their infants' temperament, within Study 2, multiple informants are asked to report on infant temperament using the IBQ.

4.2.2. *Maternal Mental Health and its Relationship with Infant Irritability*

The work of Vaughn and colleagues also focused attention on maternal characteristics that may influence the development of infant irritability. These include the mother's own symptoms of psychopathology. Evidence that maternal anxiety prior to birth was related to mothers' reports of infant difficulty was taken by Vaughn and colleagues as a poor indicator of the ITQ's validity in measuring infant temperament. Whilst there may be methodological concerns about the ITQ, this important finding by Vaughn and colleagues may also indicate important intergenerational issues. The construct of temperamental difficulty measured on the ITQ includes fear and irritability. Maternal anxiety as defined by DSM-IV-TR (2000)

includes irritability as a main symptom of GAD. The relationship between maternal anxiety and infant difficulty could therefore reflect intergenerational transmission of irritability, not just measurement error, and is worthy of further investigation using a more reliable infant temperament questionnaire and method. I aim to build upon the work of Vaughn and colleagues to explore the role of maternal psychopathology in relation to the mothers' own irritability, and further to consider whether both mothers' history of psychopathology and mothers' temperamental irritability influence the development of infant irritability.

There is a large body of evidence that indicates maternal mental health has an influence on the development of psychopathology in children (e.g., Beck, 1999; Pawlby, Hay, O'Keane, Waters, & Sharp, 2009), and on the development of infant temperament (e.g., Austin, Hadzi-Pavlovic, Leader, Saint, & Parker, 2005). Study 1 indicated the importance of children's irritability as a mediator in the co-morbidity of internalising and externalising problems in childhood. In adulthood, irritability is a prominent symptom in anxiety and mood disorders and in antisocial personality disorder (DSM-IV-TR; 2000). Understanding the potential intergenerational transmission of irritability between mother and infant in this present study will therefore require a close examination of maternal irritability in the context of the mothers' mental health. Study 2 will expand on present knowledge through consideration of the role that mother's irritability has in the relationship between adult emotional and behavioural disorders, such as depression and anxiety, and conduct symptoms respectively.

4.2.3 *Convergence between Maternal Reports of Infant /Temperament and Observations of Infant Behaviour*

Finally, Vaughn and colleagues measured the validity of the mothers' reports on the ITQ by comparing the ITQ reports with observed infant behaviours during mother-infant interaction tasks. The mothers completed the Personality Research Form (Jackson, 1967) which includes a social desirability scale. The mothers who reported themselves highly on the social desirability scale were significantly more likely to demonstrate good caretaking skills and positive affect during the observed feeding task and the correlations were approaching significance for the relationship between mothers' social desirability ratings and their play skills and attitude towards their infant during the observed interactive play task (Vaughn et al., 2002). These results may indicate that mothers were adapting their behaviours during the observation tasks to demonstrate social desirability in front of the researchers, and thus the observations may not give a realistic impression of the everyday mother-infant interactions that may serve to inform mothers' ratings of their infant's temperament. Within this study, I aim to use an observation task that is focused on the infant behaviour rather than the mother-infant interaction to avoid this social desirability risk.

The IBQ has been designed to reduce the likelihood of error in relation to criticisms about the use of mothers as informants of infant temperament. The IBQ asks about concrete behaviours rather than abstract behaviours, such as, *during feeding how often did the baby fuss or cry when s/he had enough to eat?*, which limits bias and removes the need to make comparative judgements with other infants. The IBQ is designed to ask about recent events (i.e. within the last week) to minimise any recall problems, and focus questions on a particular situation, such as bathing or feeding, to enable the informant to recall specific

recent examples (Gartstein & Rothbart, 2003). Questions asked in this way reduce the likelihood of socially desirable answers.

Using convergence across methods to test the reliability and validity of mothers' reports of infant temperament has raised methodological and theoretical questions about research on infant temperament. The factors derived from scales used on a questionnaire of infant temperament may differ from the operational definitions of behaviours observed either in the laboratory or in home settings. Different studies of convergence across measures suggest this may be the case (Pauli-Pott, Mertesacker, & Beckmann; 2004; Pauli-Pott; Mertesacker, Bade, Bauer, & Beckmann, 2000; Stifter, Willoughby, & Towe-Goodman, 2008).

In a recent study, Stifter and colleagues examined the convergence on infant temperament, in 955 infants at age 6-7 months as part of the Family Life Project, between mothers and observers using three methods: the IBQ for mothers' reports of infant temperament, an adaptation of challenges from the Lab-TAB (Goldsmith & Rothbart, 1996), for independent observation of infant temperament, and observers' global ratings of infant temperament following the observer's visit to the infant's home using the Infant Behaviour Record (IBR; Bayley, 1969; Stifter et al., 2008). Stifter and colleagues combined the 'distress to limitations' and 'distress and latency to approach novel stimuli' scales of the IBQ to create a negativity dimension from the mothers' reports of infant temperament. Using Structured Equation Modelling (SEM) to test the variation of infant temperament by method or trait (i.e. positivity or negativity), Stifter and colleagues found that, regardless of which observation method was tested, mothers' IBQ reports and observers agreed only on the extent to which the infant was positive.

In comparison, a study of 4-month-olds by Pauli-Pott and colleagues found good convergence, using infant facial expressions and vocalisations to measure observed infant negativity with mothers' reports of negative emotionality on the IBQ, $r(101) = .38, p < .001$ (Pauli-Pott et al., 2000). The follow-up study by Pauli-Pott and colleagues when the children were 8- and 12-month-olds revealed inconsistent results, with convergence across observed and IBQ measures of infant negativity not significant at 8 months but significant at 12 months (Pauli-Pott et al., 2004). Discontinuity in negative emotionality across the first year of life has been well documented, with the 6-month stage reflecting developmental changes in the infant control of emotions (Belsky et al., 1991; Rothbart et al., 2000). This does provide a particular challenge for research on the validity of measures of temperament during this development period.

Further evidence of this developmental plasticity is provided by Hane and colleagues (Hane et al., 2006), who sought to understand the situations in which mothers are most likely to be influenced by their infant's expression of emotions. Hane and colleagues obtained mothers' IBQ reports of infant temperament in 59 infants at 9 months, two home-based observations of infant negativity using Kochanska's (1997, 1998) scales, plus Lab-TAB observational ratings of infants' anger and fear in the laboratory. The IBQ and Lab-TAB are derived from the same theoretical tradition (Goldsmith & Rothbart, 1996; Rothbart, 1981), and temperament ratings on the IBQ have been shown to converge with observed behaviours on the Lab-TAB (Bridges, Palmer, Morales, Hurtado, & Tsai, 1993), but when the two distress scales were used as a composite score by Hane and colleagues, convergence with the Lab-TAB was not apparent. The Kochanska scale of infant negativity is derived from event counts for frowning, fussing, and crying in 30-second segments, a similar observational

coding method to the method used by Pauli-Pott and colleagues. Hane and colleagues examined the zero-order correlations between three IBQ scores, the IBQ distress to limitations scale, the IBQ distress and latency of approach to novel stimuli scale, and a composite IBQ distress score (distress to limitations + distress and latency), in relation to the 3 Lab-TAB scores for anger, fear and a combined negativity score (anger + fear). The zero-order correlations revealed no significant convergence across these scales and measures. When examining convergence of the IBQ and Lab-TAB with the Kochanska scale of infant negativity on the home-based observation tasks, Hane and colleagues examined the composite IBQ distress score and the composite Lab-TAB negativity score. Excluding the single IBQ dimension scales from a comparison with the Kochanska scale may have lost important information about the relationships between similar constructs on the different measures.

I would suggest that lack of convergence between mothers' reports of infant temperament using questionnaires and independent observation measures of infant temperament is not helped by the switching of dimensions used to measure negativity. I would argue for a need to focus on a construct measured by questionnaire and then observe infant behaviour using descriptive counts to understand better the nature of the infant behaviour occurring at each stage of development, much in the same way that Shirley (1933) in her classic studies of infants provided a descriptive method to study infants over the first 12 months. The reports of infant temperament collated by researchers on infants during the 4 to 12 month development period may be reports based on the informant's experience prior to the changes taking place in the infant, e.g. a mother may be influenced by her experiences since birth to the present day with the infant, whilst observations may be starting to track

behaviour adaptation and emotional development. Within this study, I aim to test the convergence between the mothers' reports of infant irritability by examining the relationship between the single dimension of irritability as defined by the IBQ 'distress to limitations' scale and an independent descriptive observation of the infant's behaviour at 6 months. I aim to use an everyday home-based activity suggested by Hane and colleagues as a measure that better reflects the situations in which mothers' perceptions of infant temperament are more salient.

4.2.4 Measuring *Irritability in the Context of Other Dimensions of Infant Temperament*

From Study 1 it was clear that irritability may occur in the context of both emotional and behavioural problems in childhood. Study 1 demonstrated the relationship between irritability in pre-school children and symptoms of emotional problems, which included excessive fear and anxiety, and behavioural problems, which included symptoms of conduct disorder and hyperactivity (see Table 3.4 in Chapter 3 for details of symptoms measured). When studying irritability in infancy in a developmental psychopathology framework, it is therefore important to understand what other infant behaviours and dimensions of temperament relate to infant irritability, as this may provide a greater insight in to the developmental pathways to co morbid emotional and behavioural disorders. As we have seen, there is evidence that infant irritability correlates with fear and activity (Janson & Mathiesen, 2008; Komsi et al., 2006). Within this study I aim to further understanding of the early pathway from irritability to both emotional and behavioural disorders by examining the relationship between irritability and other temperament factors using the IBQ (Rothbart, 1981), and describing the wider behavioural context within which infant irritability exists.

This review of the conceptual and methodological issues has informed the study design for Study 2. The aims and factors examined in Study 2 are set out in relation to the two hypotheses for Study 2 in Figure 4.4. The method for the study of the intergenerational transmission of irritability is then outlined, followed by the results of the analyses and a discussion of these results in relation to the literature reviewed.

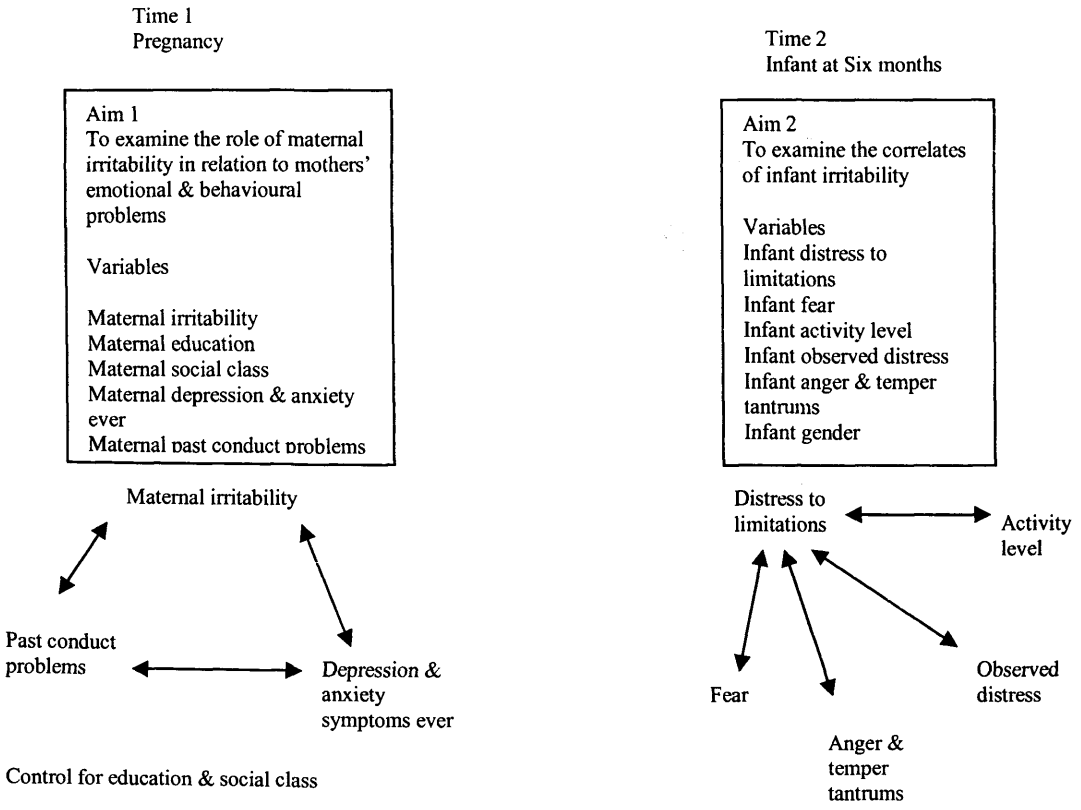


Figure 4.4. The model of analyses for Study 2: Testing the relationship of irritability with adult emotional and behavioural problems, the description of infant irritability and its correlates, and the intergenerational transmission of irritability from mother to infant.

4.3. Method

4.3.1 *The Participants*

4.3.1.1. *Sampling and recruitment.* The mothers and their infants in this study participated in the Cardiff Child Development Study (CCDS), a prospective longitudinal study in which N = 332 primiparous women were recruited from hospital and general practice antenatal clinics across the South Wales area, UK. The antenatal clinics visited were selected in consultation with the midwifery teams in the local NHS Trusts to recruit a broad spectrum of families across the South Wales area, including specialist antenatal clinics for families with medical problems, and specialist outreach antenatal services for vulnerable first-time parents. The recruitment strategy resulted in a sample for the CCDS that represents approximately 50% of the available population of primiparous women presenting in the clinics attended. The main reason articulated by women who declined to participate in the study was the time commitment required for a 7-wave longitudinal study.

4.3.1.2 *Measurement of socio-demographic characteristics.* Within the present study, socioeconomic status is embodied in the idea of capital (Bradley & Corwyn, 2002). Capital, in terms of resources and assets, embodies 'access to financial capital' as measured by occupational status, and 'access to human capital' as measured by nonmaterial resources such as education. Socioeconomic status of the mothers in this study is assessed by measuring the women's access to these resources, the women's occupational status as a proxy for evidence of financial capital, and human capital as measured using the women's educational qualifications. It is suggested that the modes of capital can be treated as either a composite measure of socioeconomic status or individually (Bradley & Corwyn, 2002), and within this study the items are measured individually.

As part of the antenatal assessment mothers were asked to report on their educational qualifications. Using the information provided a dichotomous variable was created, and women were classified as having either 5 GCSE's +, or less than 5 GCSE's. These categories are based on the basic educational qualification criteria to be achieved by age 16, currently used in the present UK educational system as a minimum standard for A 'level study and access to University. The mothers' educational variable is therefore reported as having received education up to 16 years (less than 5 GCSE's) and post 16 years (5 GCSE's +).

Maternal social class was determined using the Standard Occupational Classification 2000 (SOC 2000; Elias, McKnight, & Kinshott, 1999). The highest ranked employment ever held by the mother was used to classify the mother's socioeconomic status. A dichotomous variable was created classifying mothers as either middle or working class. The middle class group included the SOC major groups 1 to 3:

1 Managers and Senior Officials

2 Professional Occupations

3 Associate Professional and Technical Occupations

Working class included those never employed plus the SOC major groups 4 to 9, as follows:

4 Administrative and Secretarial Occupations

5 Skilled Trades Occupations

6 Personal Service Occupations

7 Sales and Customer Service Occupations

8 Process, Plant and Machine Operatives

9 Elementary Occupations

Maternal childbearing age was determined using measured mother's date of birth and the infant's date of birth (Waters, 2008).

4.3.1.3 Representativeness *of the sample.* The CCDS study assessed families at five time points, and the first two time points are used in the present study: Wave 1 assessments in pregnancy and Wave 2 assessments six months after the birth of the first infant. The CCDS sample (N=332) is not significantly different from first-time parents of the nationally representative Millennium Cohort Study with respect to socio-demographic characteristics (MCS; Kiernan, personal communication 2008).

In the CCDS, in 86% of cases both biological parents were assessed (92% of fathers in stable relationships with the mothers of their children). All women gave birth to live infants. The average age of first-time mothers was 28.1 years (range 16 to 43), not significantly different from the mean age (27.5 years) of first-time mothers in the MCS sample. The proportion of married parents at Wave 1 in the CCDS was 50.3% compared with 53% in the MCS sample, and 50.9% of mothers in the CCDS were classified as middle to upper class compared with 55% classified as the same in the MCS.

The present analyses focus on women with infants born before 31st December 2007 to ensure that the infants would be eligible for a six-month assessment of temperament and all analyses would be within the time constraints of this thesis. From the total N=332 CCDS sample, N=267 women and their infants were eligible for inclusion in the present study. Six of the 267 eligible women had twins and were excluded from the present analyses. From the remaining N=261 eligible women available for this present study, 8 women refused to participate at Wave 2, 5 could not be traced, and 4 advised that they had difficult family or medical circumstances which precluded their assessment at Wave 2. The remaining N=244

women were eligible to participate in this present study, but not all women had provided complete data at W1. This present study sample of mothers for whom we have complete data at W1 is N=201 (82% of potential N=244 sample). The demographic characteristics of the mothers in the full CCDS sample (N=332), the eligible sample of CCDS for the present study (N=244), the defined sample for the present study (N=201), and the millennium cohort sample are compared in Table 4.1. Mothers in the present study sample were slightly better qualified in terms of education and more likely to be in a married relationship than the full CCDS sample, but the mothers' characteristics in the present study sample were not significantly different from the sample characteristics of the Millennium Cohort Sample (MCS; Kiernan, personal communication).

Table 4.1

The Maternal Characteristics of the Present Study Sample compared with the full CCDS sample (N=332), the potential sample for the present study (N=244), the actual sample of women at Wave 1(N=201), and the Millennium Cohort Sample.

Maternal characteristics	*N = 201*	*N= 244*	*N=332*	MCS
Social Class (%) Middle/upper class (SOC 1-3)	58.2%	52%	50.3%	55%
Maternal age at first birth (years)	29.03 (range 17.01 to 41.81)	28.43 (range 16.1 to 43)	28.2 (range 16.1 to 43)	27.5
Education up to 16 years	18.9%	20.7%	21.7%	---
Marital status at Wave 1 (%)				
Married	55.7%	51.6%	50.3%	53%

At Wave 2, 180 families provided questionnaire data for the infant assessment. Student t-tests were used to measure differences between families who provided questionnaire data at Wave 2 (N=180) and those families who did not provide questionnaire data at Wave 2 (N=21). Families who did not provide questionnaire data at Wave 2 were significantly more likely to be in the working class group, $t(199) = -3.65, p<.005$, and the mothers were significantly more likely to have left education at 16 years, $t(199) = 3.46$, $p<.005$. Using an independent t-test to examine the differences in mothers' irritability at Wave 1 for the mothers who did and did not provide Wave 2 infant temperament data, there was a significant difference, $t(199) = 2.62, p< .05$. Mothers who did not provide infant temperament data on time at Wave 2 were significantly more likely to be irritable at Wave 1.

4.3.2 *Procedure*

All of the study procedures were approved by the Ethics committee of the School of Psychology Cardiff University and the Multi-Centre Research Ethics Committee for Wales which serves the NHS. Assessments were conducted during the third trimester of pregnancy and 6 months after the infant was born. The assessments made at each time point are shown in Table 4.2. My role as a researcher within the Cardiff Child Development Study was substantial and included (a) the preparation of material for submission to the ethics committees; (b) the negotiation with the local NHS Trusts to establish recruitment of participants; (c) a contribution towards the measures and design of the different waves of assessment; (d) completion of assessments at Waves 1, 2 and 4; (e) the establishment of databases for Waves 1 and 2; (f) data coding; (g) the input of data; and (h)subsequent analyses of data for this present thesis. Two time point assessments were used for analyses within this present thesis, Waves 1 and 2.

Table 4.2

Assessments at each Time Point of Study 2 for Mother and Infant.

	Wave 1 Third Trimester of Pregnancy	Wave 2 Early Infancy Mean 6 months
Maternal Measures	Home Visit Maternal interview Maternal questionnaire	Home Visit Maternal questionnaire
Infant Measures		Home Visit Questionnaire (Mother, Father, & third person who knows infant well) Observation

4.3.2.1 *Recruitment procedures.* Primiparous women and their partners were recruited to the CCDS from hospital and general practice antenatal clinics in the Counties of Newport and Cardiff and the Vale, Wales, UK. Trained researchers approached women at the clinics who were expecting their first child and gave a verbal description of the study. The researchers were also able to show a short DVD describing the study to any women and their partners who had literacy difficulties. Following the verbal and/or DVD description of the study, the researchers gave the women a leaflet about the study and asked that they could come back to discuss any queries after 10 minutes. Following further discussion with each woman in the antenatal clinic, the researcher asked the women if they would be willing to be contacted in 1-2 weeks to find out if they would be willing to participate in the study. If a woman agreed to further contact, the researcher wrote down the contact information and passed these details to the project administrator to contact the potential participants. The researchers also offered the women a chance for a further discussion about the study at home if required. The project administrator contacted all prospective participants and if they agreed

to participate, an appointment was made for an antenatal interview. Translators were employed for interviews if required.

4.3.2.2 *Wave 1: Antenatal assessment.* Research assistants visited the women and their partners at home and gave a complete verbal and written description of the study, following which written informed consent was obtained from the participants. A second written consent was obtained at this visit for an audio-recording of the interview to be made. This recording of the interview was required to ensure accuracy and consistency of the interviews and to enable standardisation of coding of interview data. Women and their partners were interviewed separately by different researchers, in separate rooms of the family home. For the purpose of this study only mothers' antenatal interview data are analysed.

All researchers who interviewed the participants received training in the use of the Schedules for Clinical Assessment in Neuropsychiatry (SCAN: Wing et al., 1990), a set of instruments validated in assessing, measuring, and classifying the symptoms of major psychiatric disorders. For the purposes of this study all participants were asked to respond to all the questions relating to depression and anxiety disorders, regardless of whether or not they screened into these disorders within the SCAN, to examine the normal variation with respect to emotional and physical health in pregnancy. During the interview participants were asked to report on their mood state from the time of conception until the day of the interview. Participants were also asked about any past history of major psychiatric disorders. Participants were also interviewed about their education and work history, antisocial behaviour, family medical and psychiatric history, and what social support was available to them during their pregnancy. The interviews lasted approximately 2 hours, and all participants were offered the chance for adequate breaks throughout.

In addition to the interview the participants were asked to complete a questionnaire and return it to the university in a large stamped addressed envelope provided with each questionnaire. A date for return was written on the top of the questionnaire and a telephone and e-mail contact was made available for participants to use if they required assistance in completing the questionnaire. The questionnaire asked about the participants' general health, lifestyle, life events that had taken place in the respondents' last 12 months (adapted from Barnett, Hanna, & Parker, 1983; & Brugha, Bebbington, Tennant, & Hurry, 1985), personality (Loranger et al., 1997; Rushton, Chrisjohn, & Fekken, 1981) and irritability (Snaith, Constantopoulos, Jardine & McGuffin, 1978), demographics (education, work, partner's work, family income & finances; Harold, Aitken, & Shelton, 2000), and relationship issues with partner (Locke et al.,1967). Participants received a gift voucher for their participation in this part of the study.

4.3.2.3. *Wave 2: Six-month postnatal procedure.* Researchers were allocated a family caseload to build a lasting relationship with the families as a means to improve the retention rate in this longitudinal study. Researchers would contact the respective families when the infant was approaching 6 months (age range for Wave 2 assessment being 5- to 8-months), initially by telephone, but in some cases, researchers would contact the families by post or in person. During the initial contact the researcher provided more information about this stage of the research and would ask the family if they were prepared to make an appointment for a home visit. Mothers were asked to nominate a time of day for the infant observation session that would fit with the infant's usual routine at a point when the infant was unlikely to be drowsy. Each family was sent a detailed information sheet about the 6-month visit. At the home visit the lead researcher would again provide a full verbal and written description of

this stage of the study, and written informed consent was obtained from the infant's mother to participate in Wave 2 of the study.

The 6-month postnatal home visit was made up of three sections: the postnatal interview of the mother, the infant observation assessment, and the questionnaire assessment of up to three informants. For the purposes of the present study the questionnaire and infant observation assessment were used for analyses.

A packet of three questionnaires was given to the mother. The first questionnaire was for the mother to complete, the second for the father[3] to complete, and the third for a person who knew the infant well. The father's and third family member's consent forms were included as part of their questionnaires. In situations where the biological mother and father were no longer in a relationship, the researcher would contact the biological father separately in order to obtain consent and completion of the questionnaire. All questionnaires were placed in a stamped, addressed envelope, with instructions for completion and contact details for assistance if required. Each questionnaire had a date for completion written on the top of the questionnaire. At this stage families with reading difficulties or families whose first language was not English, were already known to the researchers and in those circumstances, researchers would either read out the questions to the participants or organise a translator to assist with completion of the questionnaires when appropriate.

The infant observation procedures were explained to the mothers prior to each stage of the observation. The observation was organised in three sections, the first section assessed the infant's attention and social learning; the second began with an assessment of the infant's response to an everyday form of restraint in a car seat, followed by an assessment of responses to the experimenter's social intrusions (including collection of salivary DNA), and

[3] Father is used to describe the second parent but in two cases the parents were same-sex

the third section assessed mother-infant interaction during a game, a joint attention task, and a feeding task. The observation of the infant response to the everyday restraint task was the observation task analysed within this study. All observation assessments were video-recorded and coded later by independent observers.

The everyday restraint task used at Wave 2 was based on the Lab-TAB (Goldsmith & Rothbart, 1996) restraint in car seat task designed as a task to measure anger/frustration. Whilst the task is designed for use in the laboratory, it has been used within a naturalistic home environment with 8-month old infants (Clark et al., 2000). The researcher brought a standardised infant car seat suitable for the infant's age and size for use in the task at the infant's home. The mother was asked to pick up the infant and place him or her into the car seat. The mother was asked to stand to the side of the car seat whilst buckling in the infant to prevent the camera being blocked. The mother was asked to do this without talking to the infant and asked to stand or sit slightly behind the car seat. The infant was left in the seat for 30 seconds. The researcher video-recorded the infant from the point at which the mother secured the infant into the car seat for 30 seconds. The view recorded was a close-up, frontal shot of the infant's face and entire body.

Participants received a gift voucher plus a Polaroid photo of the infant at the visit as a thank you for participating in this stage of the study.

4.3.3. *Measures of Mothers' Irritability, Internalising and Externalising Problems*

4.3.3.1 *Maternal irritability.* Information about the mothers' irritability was obtained by questionnaire, before and after the birth of the infant, using the irritability scales from the Irritability, Depression, Anxiety Scale (IDA; Snaith et al., 1978; Snaith & Taylor, 1985). The IDA was developed as a clinical self-assessment scale for measuring irritability as a present

state of mind (Snaith et al., 1978), and has been widely used by the UK Department of Health (DoH, 2000) as the Adult Wellbeing Scale in the Family Assessment Framework throughout the UK (Cox & Bentovim, 2001). The IDA has also been used in other studies to measure mothers' irritability (Dunn, Slokomwski, Beardsall, & Rende, 1994). The IDA consists of four scales, including two irritability scales that were combined in this study to give a measure of irritability at each wave of assessment (i.e. 3rd trimester of pregnancy and 6 months post-partum). Irritability is defined for this scale as, 'a temporary psychological state characterised by impatience, intolerance and poorly controlled anger.... Expressed outwardly towards others or directed inwardly towards oneself' (Snaith et al., p.164). Snaith and colleagues tested the validation of the IDA on both clinical and non-clinical samples comparing self-assessments with interviews, the resulting correlations ranged between .70 and .84.

The IDA has 18 statements to which the respondent is asked to circle the response out of four choices to indicate how they are feeling or have been feeling in the last few days, e.g. 'I feel cheerful' the respondents are asked to choose from 'yes, definitely', 'yes, sometimes', 'no, not much', or 'no, not at all'[4]. The scale measures four subscales: irritability (inward-directed and outward-directed), depression and anxiety. For the purpose of this study irritability was taken as a continuous measure using a composite variable of the sum of the inward and outward expressed maternal irritability.

The mothers' irritability was measured twice using the IDA, firstly the women were asked about their irritability in the third trimester of pregnancy and again when the infants were 6 months old. N = 168 mothers completed the IDA at both time points. The descriptive statistics for the results of both irritability reports are shown in Table 4.3. Reliability of the

[4] A copy of the IDA is provided in the appendix 1

scales was examined using the Cronbach scale alpha, with reliability assumed if $\alpha \sim .80$ (Cohen & Cohen, 1983). The scale alphas indicated that the irritability scales were reliable. The skewness z-scores indicate a significant positive skew; therefore the mothers' irritability scores were transformed using a $\log_{10}$ transformation ($\log_{10} +1$). Following the log transformation the skewness z-scores were no longer significant, antenatal irritability z-score = -2.10 and 6-month postnatal irritability z-score = -1.18 (Field, 2005). The transformed scores were analysed using parametric analyses.

Table 4.3

Descriptive Statistics of the Self-reported Mothers' Irritability measured using the Irritability, Depression, Anxiety (IDA) scale (Snaith et al., 1978)

Mothers' Irritability	*N*	*Mean*	*S.D.*	*Range*	*Skewness*	*Skewness z-score*	*Kurtosis*	Kurtosis z-score
Antenatal	201	4.14	3.03	0-15	1.10	6.28	1.15	3.37
Six-month postnatal	168	4.08	3.24	0-19	1.50	8.01	3.07	8.24

4.3.3.2 *The continuity of irritability in mothers from pregnancy to 6 months after the birth of the first child.* There was a significant relationship in mothers' irritability across childbirth, r (168) = .62, $p < .0005$. As there was a significant relationship between mothers' self-reported irritability at both time points, a composite variable, taken as the mean scores for irritability across Waves 1 and 2, was computed as a measure of mothers' dispositional irritability. The descriptive statistics for the composite irritability disposition scale are shown in Table 4.4. The skewness and kurtosis were within acceptable levels therefore parametric analyses were used.

Table 4.4

Descriptive Statistics of the Composite Dispositional Irritability Scale

Mothers' Irritability	*N*	*Mean*	*S.D.*	*Range*	*Skewness*	Kurtosis
	201	4.21	2.85	0-15	1.20	1.84

4.3.3.3 Validation of *mothers' self-reports of irritability in a sub-sample, using fathers' reports of mothers' conflictual behaviour.* A subsequent analysis was undertaken to see if mothers' self-reported irritability was related to observed irritable behaviour perceived by a second informant. To test the validity of mothers' self-reports of irritability, the mothers' self-reports of irritability at Wave 2 were examined in relation to reports from those fathers who completed questionnaires at Wave 2. As part of the questionnaire at Wave 2, fathers were asked about the mothers' conflictual behaviours towards the father. The fathers were each asked four questions about their perception of their partner's conflictual behaviour towards them. Fathers were asked how often the mothers' (1)'got angry with them', (2) 'criticised you or your ideas', (3) 'shout at you because she was upset with you', and (4) 'argue with you whenever you disagreed about something'.

The questions on the questionnaire were organised on a reversed Likert scale (1 to 7, with 1 representing always and 7 representing never). The scores were added together to create a total score with range 0 – 28, with lower scores reflecting a higher level of conflictual behaviour. This scoring method would give a negative relationship when examined with other variables that use increasing value scales. The descriptive statistics for

the fathers' reports of mothers' conflictual behaviour towards him at Wave 2 are given in Table 4.5. The scale scores were normally distributed and the scale reliability was acceptable.

Table 4.5.

Descriptive Statistics of the Fathers' Reports of Mothers' Conflictual Behaviour at Wave 2

Mothers' Conflictual Behaviour	*N*	*Mean*	*S.D.*	*Range*	*Skewness*	*Kurtosis*	Scale Reliability ά
W2 6-month postnatal	129	19.70	5.27	4-28	-.83	.09	.91

There was a significant negative relationship between the mothers' self-reports of irritability at Wave 2 and the fathers' concurrent reports of the mothers' conflictual behaviours, r (126)= -.34, p< .0005. At Wave 2, mothers' self-reported irritability was related to fathers' reports of higher levels of conflictual behaviour in the mothers. In addition, mothers' dispositional irritability measured as a composite mean across Waves 1 and 2 was also examined in relation to father's reports of mothers' conflictual behaviour at Wave 2. There was a significant negative relationship between the measure of mothers' dispositional irritability and fathers' reports of the mothers' conflictual behaviour, r (129) = -.42, p< .0005.

4.3.3.4 *Maternal mental health.* At Wave 1, maternal depression and anxiety was assessed using the SCAN psychiatric interview (Wing et al. 1990). The SCAN is a comprehensive assessment measure of major psychiatric disorders that uses the two classification systems, DSM-IV and the ICD-10 to classify major psychiatric disorders (WHO, 1994). For the purpose of this study DSM-IV-TR (2000) criteria was used to

diagnose Major Depressive Episode and Disorder (MDE/MDD) and anxiety disorders. Researchers trained in the use of the SCAN asked questions as directed through the SCAN to gain information on DSM-IV-TR (2000) symptoms for MDE/MDD, and anxiety disorders. Researchers were trained to assess the presence and severity of each symptom. Following the completion of each interview researchers transcribed the interview and coded the presence or absence of symptoms according to the SCAN glossary. If a participant had evidence of symptoms for MDE/MDD and/or anxiety disorders, the case was discussed at a monthly case conference with a consultant psychiatrist who would assess the information and make a judgement on the presence or absence of a DSM-IV diagnosis. Inter-rater reliability was checked between two psychiatrists using a 10% random sample of interviews (N = 22). There was significant overall agreement for pregnancy diagnoses, kappa = 0.78, $p < 0.001$. Based on the psychiatrist's decision, two dichotomous variables were created that measured whether or not the mother had met DSM-IV-TR criteria for MDE/MDD or anxiety disorders during pregnancy. In addition a variable was created to measure co morbidity of depression and anxiety in pregnancy, co morbidity was identified when diagnostic evidence indicated that both depression and anxiety were present concurrently in pregnancy. A single dichotomous variable was created from the combination of the depression, anxiety and co morbid variables to indicate the presence or absence of depression and/or anxiety in mothers in pregnancy.

Past history of depression and anxiety disorders were recorded as part of the SCAN using DSM-IV-TR (2000) criteria. The past history was taken as the worst past episode of either depression or anxiety and subjected to the same rigour for diagnosis as the present diagnoses. Two dichotomous variables were created that measured the presence or absence of past depression, and presence or absence of anxiety disorders. The past depression and

anxiety disorders were combined to create a single dichotomous variable indicating the presence or absence of past depression and/or anxiety in mothers. The prevalence rates are detailed in Table 4.6.

Table 4.6

Prevalence Rates of Depression & Anxiety Disorders in Pregnancy and the Past (N=201)

	Depression in pregnancy	*Anxiety in pregnancy*	*Co morbid depression and anxiety in pregnancy*	*Depression or anxiety disorders in pregnancy*	*Past depression*	*Past anxiety*	**Past depression & anxiety disorders**
Frequency	22	9	6	37	44	15	53
%	10.9	4.5	3	18.4	21.9	7.5	26.4

To assess the relationship between irritability and mothers' experiences of internalising disorders within this present study, a dichotomous variable of lifetime experience of mothers' depression and anxiety disorders was created from the diagnostic information on mothers' past and present depression and/or anxiety disorders. Of 201 mothers interviewed at Wave 1, 76 (36.3%) were diagnosed as having experienced depression and/or anxiety up to and including the pregnancy. A categorical variable for mothers' lifetime experience of depression and/or anxiety disorders was computed with 4 categories. The categories included: no past or present caseness (those who had never experienced depression and/or anxiety disorders); present caseness (those experiencing current depression and/or anxiety disorders); past caseness (those who had experienced depression and/or anxiety disorders prior to pregnancy); and past and present caseness (those who had experienced depression and/or anxiety disorders both before and during pregnancy).

The number of mothers in each category is shown in Figure 4.5. Out of the 201 mothers in the present study, 36 (18%) mothers had experienced depression and/or anxiety disorders in the past; 20 (10%) mothers were experiencing depression and/or anxiety disorders in pregnancy, and 17 (8%) mothers had experienced depression and/or anxiety before and during pregnancy.

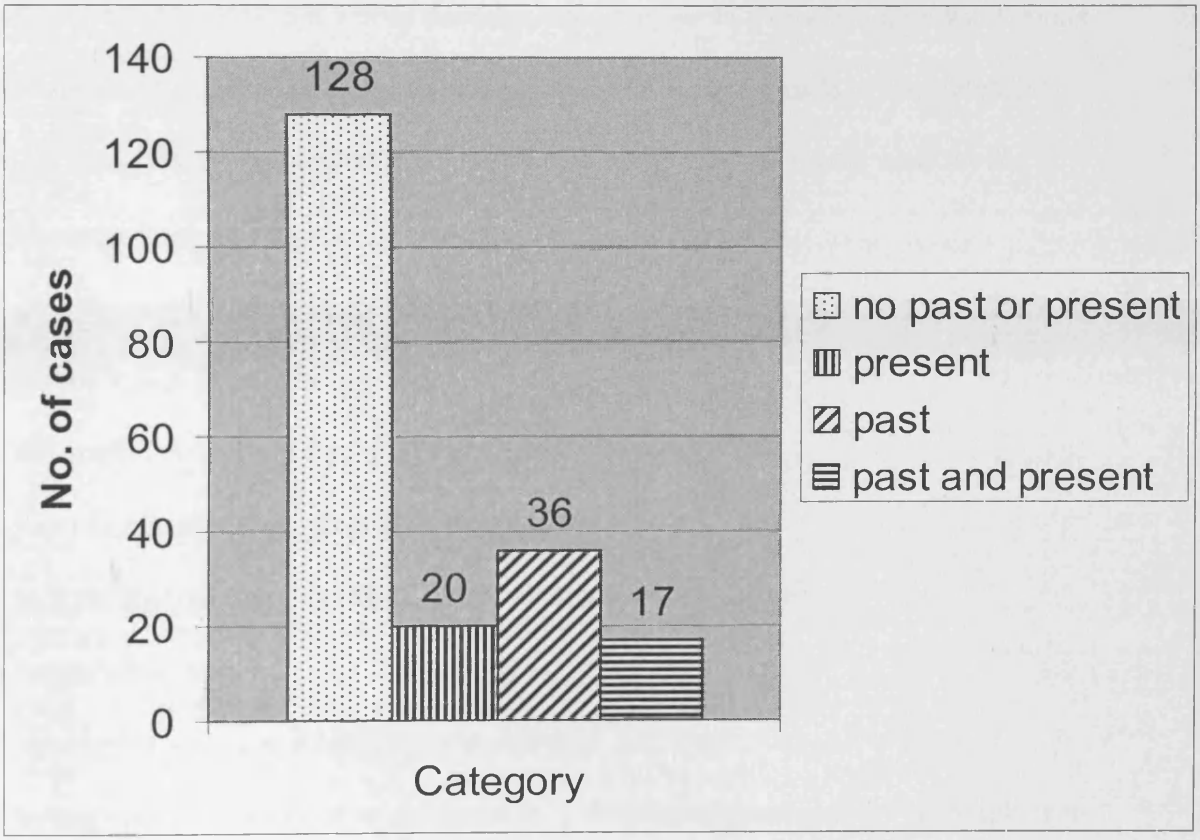


Figure 4.5.

Number of mothers' who have experienced depression and/or anxiety disorders up to and including pregnancy

4.3.3.3 *Maternal History of Conduct Problems.* As part of the antenatal assessment questionnaire, women were asked retrospectively about their history of behavioural problems during adolescence. In particular, the women were asked about antisocial behaviours to match the core DSM-IV symptoms of conduct disorder: aggression, serious violations of rules, deceitfulness or theft, and destruction of property. The women were asked to consider statements about behaviours that would describe what they were like when they were young (before the age of 16), and were asked to tick on a 3-point scale for each of the statements whether this was not true, somewhat true or certainly true. The statements used for the conduct problems were as follows:

I got very angry and often lost my temper

I played truant from school

I did graffiti or damaged property in other ways

I was often accused of lying or cheating

I took things that were not mine from home, school, or shops

I fought a lot

I usually did what I was told (reverse scored)

The responses were scored as 0 = not true, 1 = somewhat true, and 2 = certainly true. A past conduct problem variable was created by adding together the scores from the responses as a continuous measure. The composite conduct problems scale had acceptable levels of internal consistency, $\acute{\alpha} = .76$, and was significantly associated with mothers' reports of having been arrested, $r = .46, p < .001$.

For the present analyses, to avoid any potential measurement confounding the anger/temper item was removed from the scale. The items remaining reflected the DSM-IV symptoms of conduct disorder: aggression (fought), destruction of property (graffiti or damage); deceitfulness or theft (lie or stole), and serious violations of rules (truancy). The descriptive statistics for the past conduct problems scale is shown in Table 4.7. The scale had acceptable skewness and kurtosis scores. The pure conduct problem scale without the anger/temper item showed an acceptable reliability, ά = .72, within the guidelines for adequate reliability (Cohen & Cohen, 1983).

Table 4.7

Descriptive Statistics of the Mothers' Past History of Conduct Problems (DSM-IV-TR; 2000)

	N	*Mean*	*S.D.*	*Range*	*Skewness*	Kurtosis
Past Conduct Problems	201	1.46	1.72	0-8	1.55	2.32

4.3.4 *Infant Measures*

4.3.4.1 *IBQ reports of infant irritability.* For the present study, I used the IBQ Distress to Limitations scale as a measure of infant irritability, defined as:

'Child's fussing, crying or showing distress while (a) waiting for food, (b) refusing a food, (c) being in a confining place or position, (d) being dressed or undressed, (e) being prevented access to an object toward which the child is directing her/his attention' (Rothbart, 1981, p573).

The IBQ (Rothbart, 1981) is one of the most commonly used questionnaires in current infant temperament research (Putnam & Stifter, 2008). The IBQ has 87 items, and requires

the parents or carers to assess the frequency of the occurrence of temperamentally salient infant behaviours along a 7-point Likert-type scale across a number of temperamental dimensions, including *activity level, soothability, distress to limitations, distress and latency to approach novel or sudden stimuli (fear), duration of orienting,* and *smiling and laughter.* Reliabilities for the IBQ scales based on coefficient alphas range from .67-.80 for 6-month-old infants and .72-.84 for 12-month-old infants (Rothbart, 1981). Specifically for the *distress to limitations scale* (also referred to as irritable distress) the reliabilities are reported as .80 at 6 months and .78 at 12 months. The IBQ has good internal consistency and construct validity when compared with other infant temperament measures, such as the Revised Infant Temperament Questionnaire, (RITQ; Carey & McDevitt, 1978) and the Infant Characteristics Questionnaire, (ICQ; Bates, Freeland & Lounsbury; Goldsmith & Rieser-Danner, 1986).

The distress to limitations dimension was chosen as a measure of irritability because, in examination of the literature, it became apparent that the use of this dimension was derived from attempts to measure irritability (Rothbart, 1981; and Rothbart & Bates, 1998; see Chapter 2). In many studies the 'distress to limitations' and 'distress and latency to approach sudden or novel stimuli' dimensions of the IBQ are used as a composite measure of negative emotionality. For this study 'distress and latency to approach sudden or novel stimuli' was not used as a measure of irritability because examination of the origins of this dimension indicated that it was designed to tap fear rather than irritability (Rothbart, 1981; Rothbart & Bates, 1998).

The IBQ 'distress to limitations' score was taken as the mean score of the distress to limitations scale, achieved by adding together numerical scores on 20 items from the questionnaire. The IBQ has a Likert-type scale for respondents to answer in relation to the

infant. If a respondent does not answer an item or answers as 'does not apply' then that item received no numerical score and the total sum for the scale is calculated by the total number of scale items receiving a numerical response. For example, given a sum of 40 for a scale of 20 items, with one item receiving no response, and one item marked 'does not apply', and 18 items receiving a numerical response, the sum of 40 would be divided by 18 to yield a mean of 2.22 for irritability. There are a number of items on the IBQ which are reverse score items, which have to be entered as reverse scores to achieve the required scale means. Three IBQ 'distress to limitations' scores were computed, based on the reports from the three possible informants: mothers, fathers and other informants who knew the infants well.

Using the IBQ, two further infant temperament dimensions were measured, 'distress and latency to approach sudden or novel stimuli' (designed to tap fear) and 'activity level'. The 'distress and latency to approach sudden or novel stimuli' dimension, referred to as the 'fear' scale, measures the child's distress to sudden changes in stimulation and the child's distress and latency of movement toward a novel, social, or physical object (Rothbart, 1981). The fear score was taken as the mean score on the scale, achieved by adding together numerical scores on 16 items from the questionnaire. The scoring system was the same as detailed in the IBQ 'distress to limitations' paragraph above. Three IBQ fear scores were computed for mother, father and third person reports on each infant. The activity level dimension of the IBQ measures the child's gross motor activity, including movement of arms and legs, squirming and locomotor activity, and is assessed using 17 items on the IBQ. Three IBQ activity level scores were derived for mother, father and third person reports on each infant.

Within the present study, the distributions of the IBQ scale scores reported by mother, father and a third person (another family member or family friend who knew the infant well) were examined for normality and linearity. Whilst there was some evidence of skewness and kurtosis on the distress and latency of approach dimension of the IBQ (see table 4.8), the activity and distress to limitations dimensions were normally distributed. Previous studies using the IBQ dimensions have treated the scales as normally distributed and used parametric analyses without transforming IBQ scale scores (Crockenberg & Acredelo, 1983), and therefore it was deemed appropriate to treat the IBQ scales overall as normally distributed and parametric analyses was subsequently used on the data.

Table 4.8.

Descriptive Statistics for the IBQ scales by Each Informant

Informant	*Scale*	*N*	*Mean*	*S.D.*	*Range*	*Skewness*	*Skewness z-score*	*Kurtosis*	Kurtosis z-score
Mother IBQ	distress to limitations	166	3.08	.82	1.42-5.50	0.38	2.04	-0.18	-0.48
	distress & latency	166	2.27	.75	.33-4.71	0.82	4.34*	0.86	2.30
	activity	166	3.97	.87	2.00-5.60	-0.04	-0.19	-0.18	-0.47
Father IBQ	distress to limitations	131	3.05	.75	1.24-5.17	0.25	1.18	0.03	0.08
	distress & latency	131	2.27	.66	1.00-5.06	0.89	4.20*	2.03	5.36*
	activity	131	4.08	.79	2.00-5.60	-0.25	-1.17	-0.46	-1.10
3rd Person IBQ	distress to limitations	126	2.92	.81	1.00-5.25	0.64	2.98	0.25	0.58
	distress & latency	126	2.38	.84	1.00-5.40	1.01	4.68*	1.22	2.85
	activity	126	3.94	1.00	1.40-7.92	0.41	1.91	1.21	2.82

* Z-scores significant at $p<.01$

4.3.4.2 *Reliability of mothers as informants on infant temperament dimensions using the IBQ.* The reliability of mothers as informants of their infants' temperament at 6 months was examined by measuring agreement on the three scales used (distress to limitations, fear, and activity level) between the mothers and fathers, and between the mothers and a third person who knew the infant well. The resultant correlations are shown in Table 4.9. As can be seen, significant correlations were found between mother, father, and third person reports of distress to limitations and infant activity. There were significant correlations between

mother and father, and mother and third person reports of fear, but no significant relationship between father and third person on the same scale. Item analyses were performed for the IBQ scales and compared against the published reliabilities for this scale (Rothbart, 1981). The results are shown in Table 4.10.

Table 4.9

Pearson correlations between IBQ scales (distress to limitations, fear, and activity level) across all informants (N^a)

IBQ Scale	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1. Mother: distress to limitations	---								
2. Father: distress to limitations	.57**	---							
	(127)								
3. 3rd Person: distress to limitations	.55**	.33**	---						
	(115)	(98)							
4. Mother: activity	.40**	.26**	.19*	---					
	(166)	(127)	(115)						
5. Father: activity	.25**	.42**	.26**	.52**	---				
	(127)	(131)	(98)	(127)					
6. 3rd person: activity	.28**	.15	.29**	.31**	.27**	---			
	(115)	(98)	(126)	(115)	(98)				
7. Mother: distress & latency	.51**	.26**	.38**	.23**	.12	.03	---		
	(166)	(127)	(115)	(166)	(127)	(115)			
8. Father: distress & latency	.29**	.52**	.23*	.13	.23**	.02	.44**	---	
	(127)	(131)	(98)	(127)	(131)	(98)	(127)		
9. 3rd person: distress & latency	.30**	.03	.59**	.12	.02	.12	.44**	.17	---
	(115)	(98)	(126)	(115)	(98)	(126)	(115)	(98)	

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

[a] N varies according to attrition between informants

Table 4.10

Intra-class Correlations on IBQ scales for Study 2 Compared with Published Correlations (Rothbart, 1981; & Gartstein & Rothbart, 2003).

IBQ Scale	*Mother Reports* á	*Father Reports* á	*Third person reports* á	*Published Studies* á
Distress to limitations	.86	.83	.84	.80
activity level	.82	.73	.81	.77
Distress & latency	.82	.76	.78	.81

The IBQ scale alphas in Study 2 compared well with those published by Rothbart (1981), and are within the guidelines for adequate reliability (Cohen & Cohen, 1983).

4.3.4.3. *Construction of three composite IBQ variables for infant distress to limitations, fear, and activity level.* As a result of the adequate reliabilities found on the IBQ for each of the scales between informants, and to include all participants for whom at least one informant had completed the IBQ (N = 180), a composite variable was made for each IBQ scale by taking the mean across all informants. Although comparison of different informants' reports is often very helpful (Goodman et al., 2009), a sample mean is theoretically the best estimate of the true population mean, and, in this case, serves as our global estimate of the infant's temperament with respect to these three dimensions, as

reported on by all available informants for each family. The composite IBQ reports are therefore used throughout analyses for Study 2. The descriptive statistics for the composite IBQ dimension scales are shown in Table 4.11. Overall, the skewness and kurtosis scores were within acceptable limits for normality.

Table 4.11.

Descriptive Statistics for the Composite IBQ scales (N=180)

Scale	*N*	*Mean*	*S.D.*	*Range*	*Skewness*	*Skewness z-score*	*Kurtosis*	Kurtosis z-score
distress to limitations	180	3.04	0.70	1.42-4.90	0.326	1.80	-0.20	-0.54
activity level	180	4.00	0.73	2.07-5.75	-0.028	-0.15	0.02	0.05
fear	180	2.31	0.62	1.13-4.83	1.039	5.74*	1.87	5.20*

* z-scores significant at $p<.01$

4.3.4.4 *Informants' reports of infant anger and temper tantrums.* As part of the 6-month assessment questionnaire, the three informants (mothers, fathers, and a third person), were asked questions about normative developmental attainments of the infant, using the Cardiff Child Development Study Milestones Questionnaire (CCDSMSQ; Hay et al., under review). Two items in the questionnaire were reports of angry moods and temper tantrums. Informants rated each item on a scale from 0 to 2, the scores signifying 'not yet' in the infant's repertoire, possibly present, or definitely present. The scores for angry moods and temper tantrums were combined as a measure of infants' expressions of anger. The mean scores and standard deviations for the reports of anger were examined for each informant and

are given in Table 4.12. There was significant intercorrelations between the milestones anger and temper tantrums variable across informants.

Table 4.12

Descriptive Statistics for the Anger & Temper Tantrums Variable by Informant (N=201)

Anger & Temper Tantrums variable	*M*	*SD*	spearman *rho*
Mother			
	1.02	1.13	0.57**
Father			
	0.74	1.04	0.55**
Third Person			
	0.62	0.99	0.47**

**$p < .0005$.

The reliability of mothers' reports of infant anger and temper tantrums was tested against reports from the father and third person. There was a significant relationship between mothers' reports and fathers' reports for infant anger and temper tantrums, r (201) = .43, $p<.0005$, and between mothers' reports and 3rd person reports' of infant anger and temper, r (201) = .30, $p<.0005$. To ensure the best use of data available for the reports of infant anger and temper tantrums, a composite, continuous variable was created using the mean scores from all available informants per family. This composite anger variable was used for subsequent parametric analyses.

4.3.4.5 *Observed infant distress.* Infant distress was assessed through observation of the 'restraint in car seat' anger/frustration-evoking procedure adapted from the Lab-TAB (Goldsmith & Rothbart, 1996). The task has been used previously at home with infants at 8 to 10 months (Clark et al., 2000), and was considered as being suitable for home assessment because of the limited use of props required. Within Lab-TAB the car seat restraint task is

intended to elicit mild anger in responses in some children, according to the rationale that being physically restrained or compelled to do something against one's wishes can elicit anger (Goldsmith & Rothbart, 1996). The Lab-TAB car seat restraint task has been used previously with infants aged 6 months and has shown good agreement with IBQ distress to limitations, r (70) = .46, p < .0001 (Bridges et al., 1993). Car seats are required by law for the transportation of children in the UK and it is therefore considered that being restrained in a car seat would be a common experience for infants and an ecologically valid task.

All coding was conducted from videotapes using the Distress Observation System (DOS), adapted from a coding system used to record vocal distress in older toddlers (Demetriou & Hay, 2004). The aim of DOS is to code vocal and behavioural signs of distress in infants and toddlers (ages 0 to 36 months), using a clear set of operational definitions. The DOS is based on a continuous time sampling framework, with intervals timed using INTERACT software (Mangold, 2007). Each DOS observation produces a report of the nature and frequency of distress that was observed. The DOS observations used in this study were as follows:

Vocal Distress Categories:

Fusses/whimpers: Isolated instances of low-pitched complaining sounds, of short duration (an event, of 1 second or less in duration)

Whining/whinging: Low-pitched complaining noises that are sustained, or repeated at short intervals, throughout at least one 5-second interval (a state, that is 2 sec or more in duration)

Cry/weep/sob: High-pitched, loud, inarticulate utterance, uneven breath, or distinct sobbing; tears are often present

Scream: Shrill, long, loud, piercing cry expressive of pain, alarm, surprise, or other sudden emotion; uttered in a screaming tone

The coding rules for DOS are that within each time interval, score 0 if the behaviour is not present, 2 if it is definitely present, and 1 if the behaviour is possibly present, being shown in a mild way. Observers used Interact observational software (Mangold, 2007) to record categories of vocal distress in 5-second intervals. Within the 30 seconds observation there are (6 x 5second) intervals. Independent observers recorded infants' distressed vocalisations with good agreement, ICC = 0.96 across observers. For the purpose of the present study, a dichotomous variable was created for each vocal distress category, indicating 0 = no distress present and 1 = distress present. The frequencies for each vocal distress category are shown in Table 4.13. The skewness and kurtosis scale scores were not normally distributed therefore non-parametric statistics were used for subsequent analyses.

Table 4.13

Frequencies for the Presence of Infant Vocal Distress Categories in the Car Seat Restraint Task (N=180)

DOS Scales	*N*	%
Fuss	98	54
Whine	33	18
Cry	12	7
Scream	4	2

4.3.5 *Data Analyses*

The data analyses for the present study were organised according to the three aims of Study 2. Firstly, to address the question about the role of mothers' irritability in relation to mothers' own emotional and behavioural problems, mothers' irritability, mothers' lifetime caseness depression and/or anxiety and mothers' past conduct symptoms were examined, in relation to mothers' socio-economic circumstances, using correlational analyses. Variables that showed a significant association with mothers' irritability at the univariate level were further examined using logistical and linear regression, depending upon the nature of the dependent variable used in the analysis. Any potential mediating relationships were tested using the Baron and Kenny (1986) model for testing a mediating relationship.

Secondly, to examine the correlates of infant irritability, correlational analyses were used to examine the relationships between the distress to limitations scale and other dimensions of temperament (fear and activity level), the informants' reports of infant anger and temper tantrums on the CCDSMSQ, and the independent observations of infant distress in the adapted Lab-TAB car-seat task.

Finally, to examine the intergenerational transmission of irritability between mother and infant, correlational analyses were used to measure the association between mothers' IDA irritability scores and the composite score on the distress to limitations scale. The relationship between mother and infant irritability was examined, in reference to mothers' emotional symptoms, past behavioural problems (diagnoses of depression and/or anxiety disorder and past conduct symptoms) and mothers' socio-economic circumstances. Variables that showed a significant association with infant irritability at the univariate level were further examined with hierarchical linear regression. Where potential mediators of infant

irritability were identified, the Baron and Kenny (1986) model was used to test the mediating relationships. In all analyses, significance was judged at the $p<.05$ level.

4.4 Results

4.4.1 *Results for Aim 1: The Relationship between Mothers' Irritability and Mothers' History of Emotional and Behavioural Problems*

The variables that are examined as part of the analyses for aim 1 are detailed in Figure 4.6. The diagram illustrates the possible links between mothers' irritability in pregnancy and mothers' experiences of depression and/or anxiety disorders and past conduct symptoms. Mothers' social class and education are used as control variables within this analysis. Of particular interest is whether any co morbidity in mothers' emotional and behavioural problems is attributable to her irritable temperament.

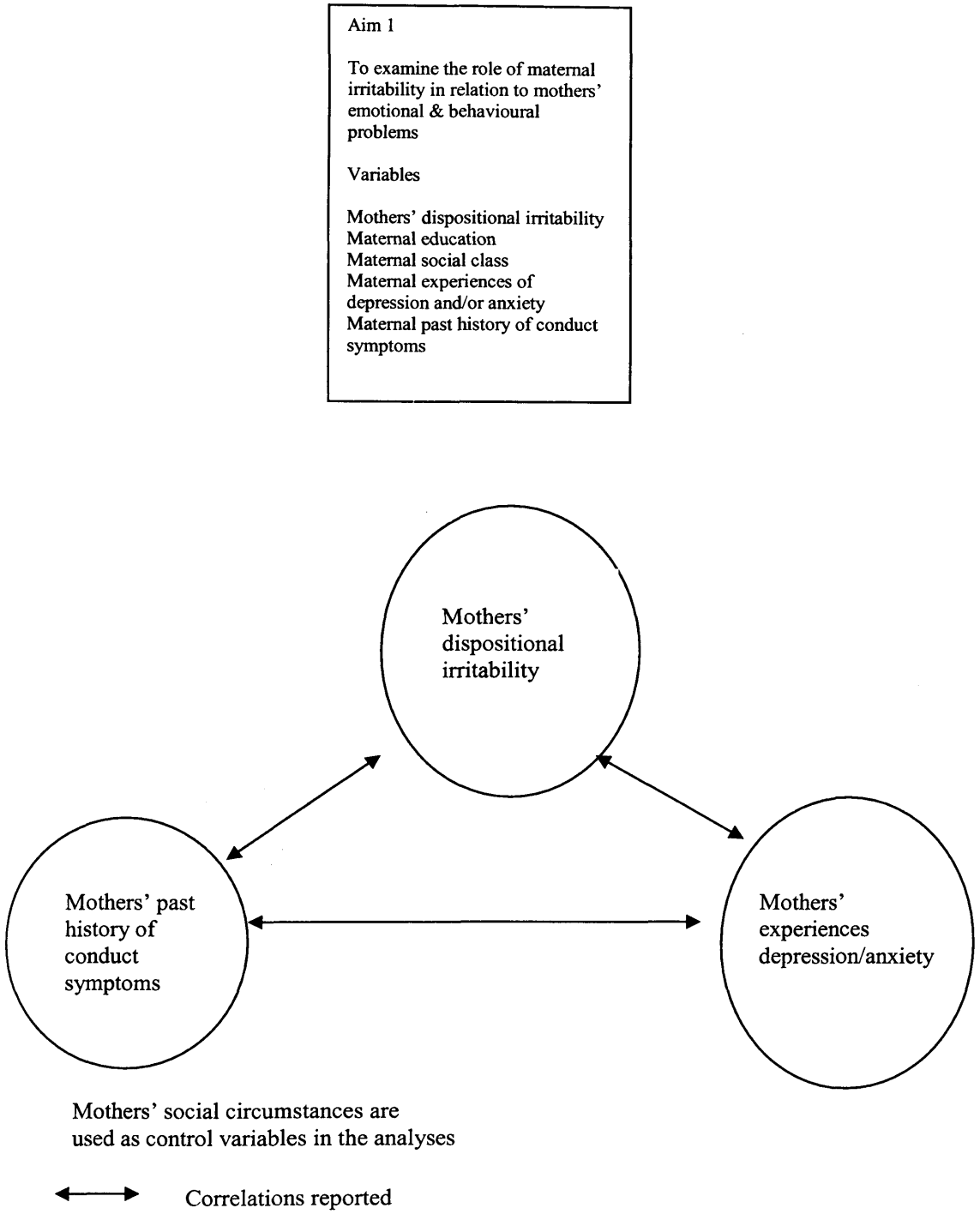


Figure 4.6. Diagram to illustrate the potential role of irritability in the relationship between mothers' experiences of depression and/or anxiety and past history of conduct symptoms.

4.4.1.1 *Examining the inter-relations between mothers' social class, mothers' education, mothers' irritability, mothers' experience of depression and/or anxiety disorders, and mothers' past history of conduct symptoms.* The inter-relations among all key variables were analysed using Pearson correlations for the continuous variables and point biserial correlations for associations between categorical and continuous variables. Mothers' dispositional irritability was significantly related to all the other variables, with the strongest relationships occurring with past conduct symptoms and mothers' experience of depression and/or anxiety disorders. The results are shown in Table 4.14.

Table 4.14 *Inter-correlations between Mothers' Social Class, Mothers' Education, Mothers' Dispositional Irritability, Mothers Past Conduct Symptoms, and Mothers' Experiences of Depression and /or Anxiety Disorders (N=201).*

Mother Variables	*1*	*2*	*3*	*4*	*5*
1 Dispositional irritability	---				
2 Experiences of depression/anxiety	.28**	---			
3 Past conduct symptoms	.32**	.33**	---		
4 Education[a]	.21**	.26**	.42**	---	
5 Social class[a]	.13*	.25**	.33**	.42**	---

[a] = point-biserial correlations

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

4.4.1.2 *Examining the differences in mothers' dispositional irritability according to mothers' experience of depression/anxiety disorders.* The differences in mothers' dispositional irritability (aggregated across Waves 1 and 2) was examined according to the mothers' experience of depression/anxiety disorders, using a one-way independent ANOVA. The 4 groups compared in the analysis were: mothers who had never experienced depression/anxiety disorders; mothers who had experienced depression/anxiety disorders in the past; mothers who were currently experiencing depression/anxiety disorders; and mothers who had both experience of depression/anxiety disorders in the past and were currently experiencing depression/anxiety disorders. There was a statistically significant main effect of the mothers' experience of depression/anxiety disorders on the mothers' dispositional irritability scores, $F(3,197) = 6.17$, $p < .0005$. Employing the Bonferroni post-hoc test, significant differences were found between mothers who had current experience of depression/anxiety disorders and mothers who had no current or past experience of depression/anxiety disorders ($p<.05$). Significant differences were also found between mothers who had both current and past experience of depression/anxiety disorders and mothers who had no current or past depression/anxiety disorders ($p<.005$). The differences in irritability by the four groups are depicted in Figure 4.7.

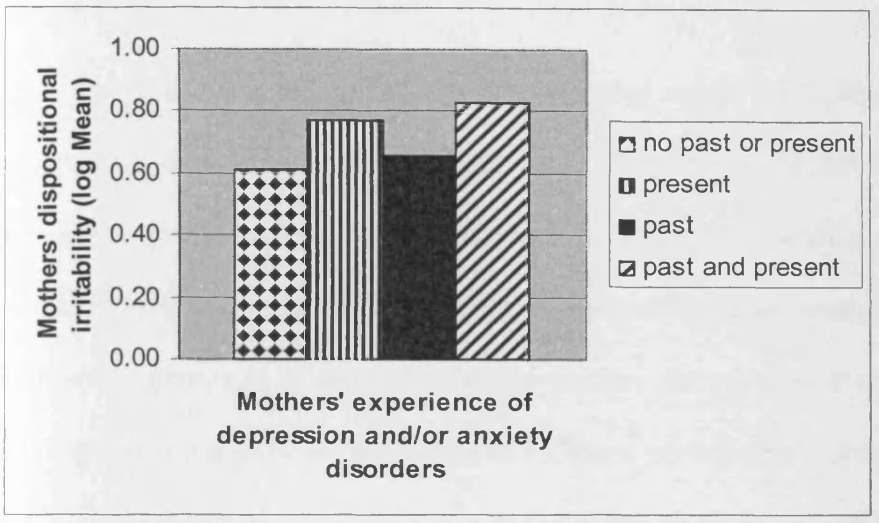


Figure 4.7 Graph to show the difference in mothers' dispositional irritability (logarithmic mean) according to mothers' experience of depression and/or anxiety disorders.

4.4.1.2 *Examining the role of mothers' dispositional irritability in the relationship between mothers' past conduct symptoms and mothers' experiences of depression and/or anxiety disorders.* The hypothesis that mothers' dispositional irritability might contribute to co morbidity between mothers' past conduct symptoms and mothers' experience of depression and/or anxiety disorders was tested using regression analyses. To control for socio-demographic factors that influence emotional and behavioural problems and temperamental irritability, mothers' social class and mothers' education were entered at step 1 in the regression models. The hypothesis was tested first using logistical regression to predict the occurrence of depression and/or anxiety disorders. Linear regression was used to predict conduct symptoms, in each case controlling for the socio-demographic factors, then testing for the influence of the other form of psychopathology (conduct problems or

depression/anxiety disorders, respectively), and finally for the mediating influence of irritability.

At the first step of the logistic regression model predicting depression and/or anxiety disorders, both mothers' social class and mothers' education significantly predicted mothers' experience of emotional disorders. Mothers' past conduct symptoms were entered next to the regression model. At this step, only mothers' past conduct symptoms predicted the mothers' experiences of depression and/or anxiety disorders, with social class and education no longer being significant predictors of mothers' emotional disorders. In the final step, mothers' dispositional irritability was added to the regression model. The results of the regression are shown in Table 4.15. Mothers' dispositional irritability significantly predicted the mothers' experience of depression and/or anxiety disorders, but did not mediate the link between the mothers' emotional and behavioural problems. Whilst the effect of the mothers' past conduct symptoms on mothers' experience of depression and/or anxiety disorders was somewhat reduced, the mothers' past conduct symptoms continued to significantly predict the mothers' diagnoses of depression and/or anxiety disorders. The model was a good fit with significant chi-squared statistics at each step, the Cook's distance values were less than 1, and there was less than 5% of cases that had absolute values above 2 on the standardized residuals, indicating that there was no influential cases having an effect on the model.

Table 4.15

Summary of Logistical Regression Analysis to Examine the Role of Maternal Dispositional Irritability in the Predictive Relationship Between Mothers' Past Conduct Symptoms and Mothers' Experiences of Depression and/or Anxiety Disorders (N = 201)

	B(SE)	*Wald*	*Lower*	*exp b*	Upper
Included					
Constant	-1.49				
	(0.18)	66.93	----	0.23	-----
Step 1					
Constant	-2.24				
	(0.31)	52.60	----	0.11	----
Social class	0.96 (0.43)	5.13*	1.14	2.62	6.03
Education	0.99 (0.44)	4.94*	1.12	2.68	6.38
Step 2					
Constant	-2.56				
	(0.34)	56.44	----	0.08	----
Social class	0.74 (0.44)	2.76	0.88	2.09	4.96
Education	0.58 (0.48)	1.45	0.70	1.78	4.55
Past conduct symptoms	0.30 (0.11)	7.40**	1.09	1.35	1.68
Step 3					
Constant	-4.12				
	(0.73)	32.03	----	0.02	----
Social class	0.72 (0.45)	2.57	0.85	2.06	4.96
Education	0.47 (0.49)	0.94	0.62	1.60	4.14
Past conduct symptoms	0.24 (0.11)	4.29**	1.01	1.27	1.58
Maternal dispositional irritability	2.41 (0.93)	32.03**	1.81	11.14	68.34

Note. For Step1 $R^2 = .08$ (Cox & Snell), .13 (Nagelkerke). Model $\chi^2(1) = 17.16, p < .005$
For Step 2 $R^2 = .12$ (Cox & Snell), .19 (Nagelkerke). Model $\chi^2(2) = 24.78, p < .005$
For Step 3 $R^2 = .15$ (Cox & Snell), .24 (Nagelkerke). Model $\chi^2(2) = 32.47, p < .005$
$^{**}p < .005$; $^{*}p < .05$.

4.4.1.3 *Examining the role of mothers' dispositional irritability in the relationship between mothers' experiences of depression and/or anxiety disorders and mothers' past conduct symptoms.* A complementary analysis was then conducted; with the mothers' past conduct symptoms as the dependent variable. The role of mothers' dispositional irritability was again tested with reference to the co morbid relationship between mothers' experience of depression and/or anxiety disorders and mothers' past conduct symptoms, taking into account socio-demographic correlates of both types of problems. Using hierarchical linear regression analysis, mothers' social class and mothers' education were entered at the first step in the regression model. At this stage, both mothers' social class and mothers' education were significant predictors of mothers' past conduct symptoms. Mothers' history of depression and/or anxiety disorders was added to the regression model at the next step. At this stage, all variables were significant predictors of mothers' past conduct symptoms. Finally, mothers' dispositional irritability was added to the model. Whilst there was a reduction in the prediction effect of social class and mothers' experience of depression and/or anxiety disorders on mothers' past conduct symptoms, all the variables continued significantly to predict mothers' past conduct problems, with both mothers' dispositional irritability and mothers' education being the strongest predictors. The results are shown in Table 4.16. The models at steps 1, 2 and 3 were a good fit for the data, and the residual plots demonstrated an accurate fit for the sample. Thus, the findings of these two complementary regression analyses showed that mothers' emotional and behavioural problems were both predicted by dispositional irritability, not accounted for by the co morbidity between the two types of problems, nor by the socio-demographic variables.

Table 4.16

Summary of Hierarchical Regression Analysis to test the role of mothers' dispositional irritability in the relationship between mothers' experience of depression and/or anxiety disorders and mothers' past conduct symptoms (N = 201)

	B	*SE B*	*β*
Step 1			
Constant	0.89	0.14	----
Mothers' social class	0.69	0.24	0.20**
Mothers' education	1.50	0.30	0.34**
Step 2			
Constant	0.80	0.14	----
Mothers' social class	0.57	0.24	0.16*
Mothers' education	1.32	0.30	0.30**
Mothers' experience of depression/anxiety	0.92	0.29	0.21**
Step 3			
Constant	-0.05	0.30	----
Mothers' social class	0.57	0.23	0.16*
Mothers' education	1.20	0.30	0.27**
Mothers' experience of depression/anxiety	0.71	0.29	0.16*
Mothers' dispositional irritability	1.40	0.44	0.20**

Note. $R^2 = .21$ for Step 1; $\Delta R^2 = .04$ for Step 2; $\Delta R^2 = .04$ for Step 3 ($p < .005$).
**$p<.005$; *$p < .05$.

4.4.2. *The Results of Aim 2: To examine the Correlates of Infant Irritability*

The variables considered in the analyses to examine the correlates of infant irritability are illustrated in Figure 4.8.

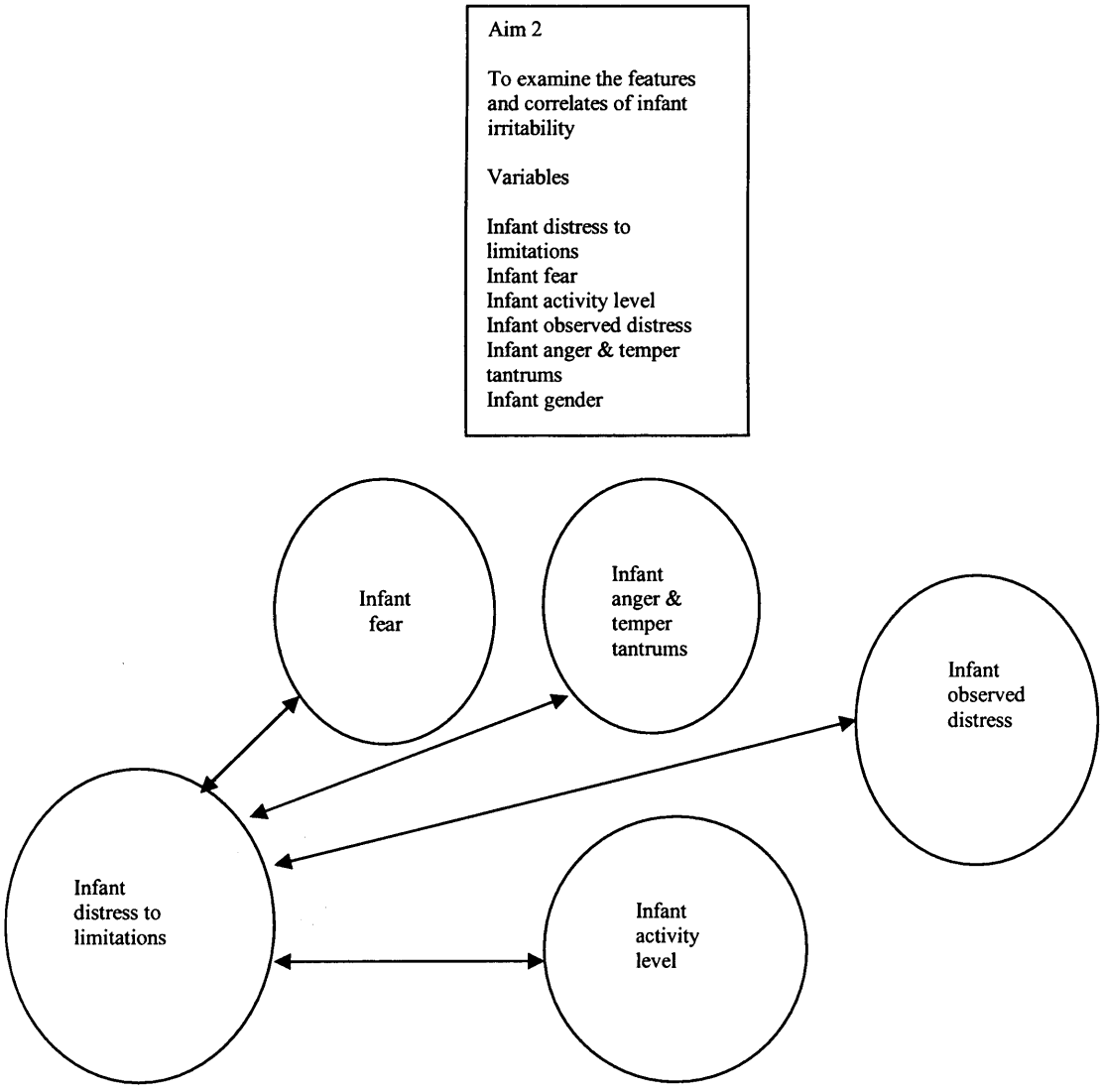


Figure 4.8. Diagram to illustrate the potential correlates of infant distress to limitations.

4.4.2.1 *Tests for gender differences in infant distress to limitations.* The means and standard deviations for infant distress to limitations for girls and boys are shown in Table 4.17. Using independent t-tests, there was a significant difference in the infant distress to limitations between girls and boys, t (178) = -3.30, p<.005.

Table 4.17

The Means and Standard Deviations for Infant Distress to Limitations by Gender (N=180)

	N	*M*	SD
Girls	87	2.87	0.60
Boys	93	3.20	0.73

4.4.2.2 *Examination of the relationship between infant distress to limitations and infant fear and infant activity levels using the composite IBQ reports.* Examination of associations between infant distress to limitations, infant fear, and infant activity level using the three IBQ composite variables revealed that infant distress to limitations was positively related to both fear and activity level dimensions of temperament (Table 4.18).

Table 4.18

Pearson Correlations between IBQ Composite scales (N=180)

Composite Scale	1	2	3
1. Distress to limitations	---		
2. Fear	.53**	---	
3. Activity level	.34**	.22**	---

** Correlation is significant at p<.01
* Correlation is significant at p<.05

4.4.2.2 *The relationship between infant irritability and reports of anger and temper tantrums.* There was a significant relationship between the composite reports of infant anger on the CCDSMSQ and the IBQ distress to limitations scale, r (180) = .44, p< .0005.

4.4.2.3 *The relationship between infant distress to limitations and observed infant distress at 6 months.* Two sets of analyses were carried out to identify the relationship between infant distress to limitations and observed infant distress in response to a Lab-Tab adapted car seat restraint task in the infant's home. Firstly, the composite distress to limitations scale was examined in relation to the Distress Observation System variables, fuss, whine, cry, and scream. Using Spearman non-parametric point biserial correlation analyses there were no significant relationships between the composite reported distress to limitations and the observed distress.

The previous studies that have investigated the relationship between reported infant distress to limitations and independent observed distress have suggested that mother reports of infant distress to limitations are more likely to converge with home-based, routine tasks, as infant behaviour is more salient to the mother within these normal everyday routines (Hane et al., 2006). Whilst there was some attrition with respect to the number of mothers who reported on their infant's temperament (N = 168 of mothers completed the IBQ), the question of the relationship between mothers' reports of infant temperament and independent observations of infant behaviour is still an important question to consider within this present study. The mothers' reports of infant 'distress to limitations' were therefore examined in relation to the four DOS variables. Using point biserial, non-parametric analyses, infant distress to limitations, as reported by the mothers, was significantly associated with observations of infant whining (i.e., according to the DOS definitions, vocalized in a

complaining tone) in response to the car seat restraint task, *rho* (148) = .18, $p<.05$, but was not significantly related to observed infant fussing, crying or screaming in the car restraint task.

4.4.2.4 *The relationship between mothers' reports of infant anger and temper tantrums and observed infant distress at 6 months.* Mothers' reports of infant anger and temper tantrums were also analysed in relation to the observed infant distress. Using point biserial, non-parametric analyses, infant anger and temper tantrums, as reported by mothers, was significantly associated with observations of infant crying, *rho* (148) = .22, $p<.005$.

4.4.3 *The Results of Aim 3: To examine the Intergenerational Transmission of Irritability Between Mothers and First Born Infants*

The variables that are examined as part of the analyses for aim 3 are detailed in Figure 4.9. The diagram illustrates the potential intergenerational transmission of mothers' irritability in pregnancy to infants at 6-months. From the analyses carried out for Aim 1 of the present study, we know that mothers' irritability at Wave 1 is related to their past conduct symptoms, their past history of depression and/or anxiety disorders, and their social class and education. To examine the predictive relationship of mothers' irritability at Wave 1 to infant distress to limitations at Wave 2, the significant correlates of mothers' irritability were all entered into a hierarchical regression analysis. Firstly, I explore the interrelationships between the potential predictors of infant irritability (mothers' irritability at Wave 1, mothers' social class, mothers' education, mothers' experience of depression and/or anxiety disorders, and mothers' past conduct symptoms). The role of mothers' irritability at Wave 2 as a potential mediator in the relationship between mothers' irritability at Wave 1 and infant distress to limitations are then explored in further regression analyses.

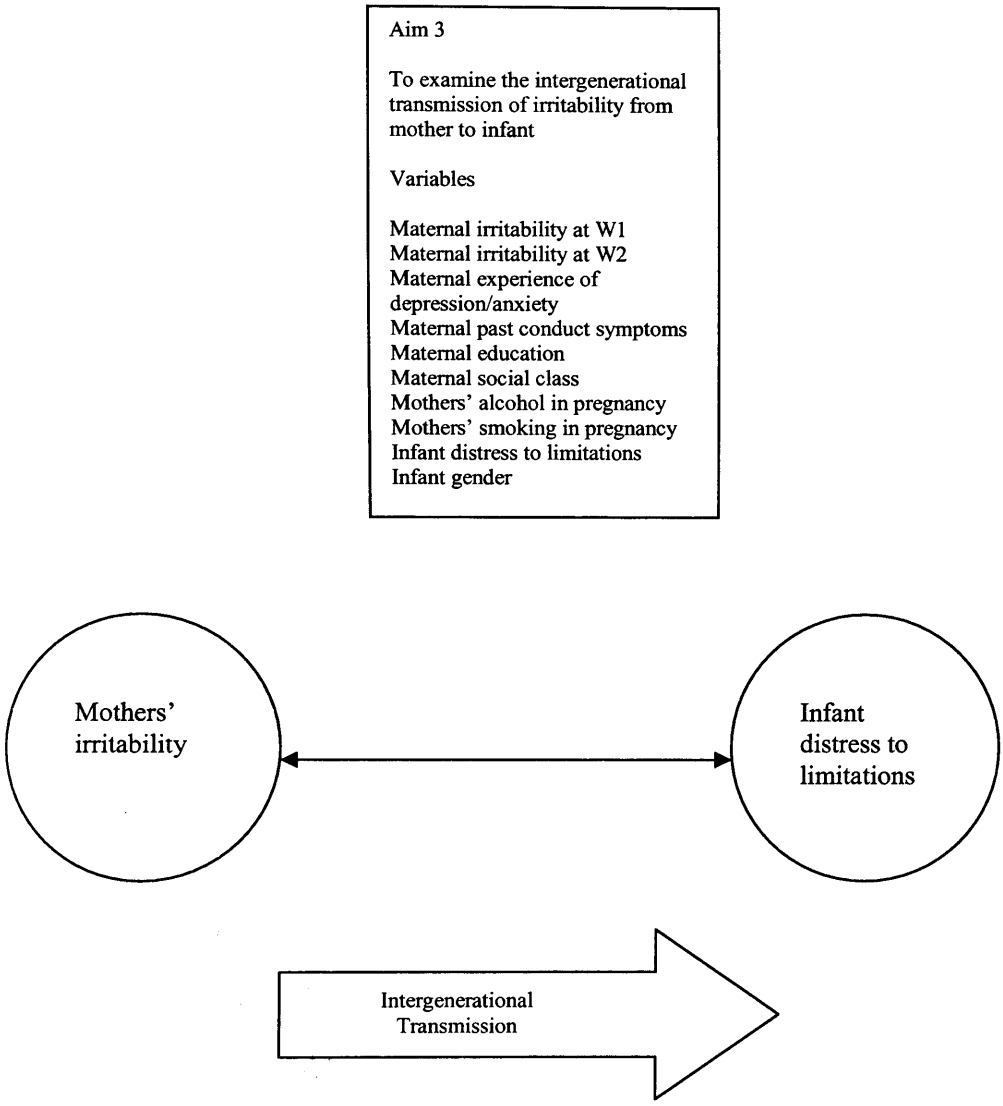


Figure 4.9 **Diagram** to illustrate the potential intergenerational transmission of irritability from mother to infant.

4.4.3.1 *The relationships between potential maternal predictors and infant distress to limitations.* The relationships between the potential maternal predictors and infant distress to limitations at 6 months are shown in Table 4.19. Maternal irritability at Wave 1 was significantly related to the composite measure of infant 'distress to limitations', r (180) = .25, $p<.005$. Mothers' history of depression and/or anxiety disorders, mothers' past conduct symptoms, mothers' social class and mothers' education were all significantly related to infant 'distress to limitations.' These predictors were entered into the hierarchical regression to explore the intergenerational transmission of irritability between mother and infant.

Table 4.19 *Inter-correlations (N[a]) between Mothers Irritability at Wave 1, other Maternal Variables and Infant Distress to Limitations.*

Variables	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1 Infant distress to limitations	----							
2 Mothers' irritability w1	.25** (180)	----						
3 Mothers' depression and anxiety disorders	.23** (180)	.33** (201)	----					
4 Mothers' past conduct symptoms	.20** (180)	.36** (201)	.29** (201)	----				
5 Mothers' education[b]	.23** (180)	.19** (201)	.24** (201)	.41** (201)	----			
6 Mothers' social class[b]	.17* (180)	.18** (201)	.20** (201)	.28** (201)	.46** (201)	----		
7 Mothers' alcohol in pregnancy	.06 (180)	.11 (201)	-.05 (201)	-.02 (201)	-.14 (201)	-.09 (201)	-	
8 Mothers' smoking in pregnancy	.07 (178)	.22** (199)	.15* (199)	.49** (199)	.29** (199)	.29** (199)	.11 (`199)	--

[a] = N varies between Wave 1 N=201 and Wave 2 N = 180

[b] = point-biserial correlations

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

4.4.3.2 *Testing the predictive relationship between maternal irritability and infant distress to limitations.* The significant predictors of infant irritability were examined, using hierarchical linear regression. The results are shown in Table 4.20.

At the first step, mothers' social class and mothers' education were entered into the model. At this stage, only mothers' education was a significant predictor of infant distress to

limitations, accounting for approximately 6% of the variance. Social class did not predict infant distress to limitations.

At the second step, mothers' past conduct symptoms and mothers' history of depression and/or anxiety disorders were entered into the model. Mothers' experience of depression and/or anxiety disorders was a significant predictor of infant distress to limitations, with the model at this stage accounting for 9% of the variance. Mothers' education was no longer a significant predictor of infant distress to limitations, and mothers' past conduct symptoms was not a significant predictor of infant distress to limitations.

In the final block of the regression, mothers' irritability at Wave 1 was added to the model. When mothers' irritability was added to the regression model, mothers' experience of depression and/or anxiety disorders no longer predicted infant distress to limitations. Mothers' irritability at Wave 1 significantly predicted infant distress to limitations, with the model accounting for 11% of the variance. The results are shown in Table 4.20. The models at steps 1, 2 and 3 were a good fit for the data, and the residual plots demonstrated an accurate fit for the sample.

Table 4.20

Summary of Hierarchical Regression Analysis for Variables Predicting Infant Distress to Limitations (N = 180)

	B	*SE B*	*β*
Step 1			
Constant	2.94	0.06	----
Mothers' education	0.38	0.16	0.20*
Mothers' social class	0.11	0.12	0.08
Step 2			
Constant	2.88	0.07	----
Mothers' education	0.26	0.16	0.14
Mothers' social class	0.07	0.12	0.05
Past conduct symptoms	0.03	0.03	0.08
Experiences of depression and/or anxiety disorders	0.31	0.14	0.16*
Step 3			
Constant	2.66	0.13	----
Mothers' education	0.27	0.16	0.14
Mothers' social class	0.05	0.12	0.04
Past conduct symptoms	0.01	0.04	0.03
Experiences of depression and/or anxiety disorders	0.23	0.15	0.12
Maternal irritability	0.43	0.21	0.17*

Note. $R^2 = .06$ for Step 1; $\Delta R^2 = .04$ for Step $2\Delta R^2 = .02$ for Step 3; ($p < .05$).
*$p < .05$.

4.4.3.3 *Testing the role of mothers' irritability at Wave 2 as a potential mediator in the relationship between mothers' irritability at Wave 1 and infants' distress to limitations.* Using hierarchical regression, the role of mothers' irritability at Wave 2 as a potential mediator in the relationship between mothers' irritability at Wave 1 and infant distress to limitations was explored. According to the Baron and Kenny (1986) model for mediation, all three variables must be significantly related to each other. Mothers' irritability at Wave 1 and 2 were significantly related, mothers' irritability at Wave 1 was significantly related to infant 'distress to limitations', and mothers' irritability at Wave 2 was significantly related to infant 'distress to limitations' (Table 4.21).

Table 4.21

The Intercorrelations between Mothers' Irritability at Wave 1 and Wave 2 and Infant Distress to Limitations (N)

	1	2	3
1. Mothers' irritability W1	---		
2. Mothers' irritability W2	.59** (168)	---	
3. Infant Distress to Limitations	.22** (180)	.34** (168)	---

**$p < .005$

To test the meditational model, a hierarchical regression analysis was performed on the 3 variables. In the first step of the hierarchical regression, mothers' irritability at Wave 1 significantly predicted infant distress to limitations at Wave 2. When mothers' irritability at Wave 2 was added to the second step of the regression model, mothers' irritability at Wave 1 was no longer a significant predictor of infant distress to limitations at Wave 2. Mothers' irritability at Wave 2 was a mediator in the relationship between mothers' irritability at Wave 1 and infant distress to limitations at Wave 2. The model accounted for 14% of the variance, and the model was a good fit for the data with the standardised residuals within the accepted levels. The results are shown in Table 4.22.

Table 4.22

Summary of Hierarchical Regression Analysis to test the potential mediating role of mothers' irritability at Wave 2 in the relationship between mothers' irritability at Wave 1 and infant distress to limitations (N = 180)

	B	*SE B*	*β*
Step 1			
Constant	2.61	0.12	----
Maternal irritability at W1	0.71	0.19	0.28**
Step 2			
Constant	2.43	0.13	----
Maternal irritability at W1	0.24	0.23	0.10
Maternal irritability at w2	0.75	0.23	0.31**

Note. R^2 = .08 for Step 1; ΔR^2 = .06 for Step 2 ($p < .05$).
**p<.005; *p < .05.

4.4.3.4 *Testing the role of infant gender as a potential moderator in the relationship between mothers' irritability at Wave 2 and infants' distress to limitations.* Using point biserial correlations infant gender was found to be significantly related to infant distress to limitations, r (180) =.24, p<.005, and mothers' concurrent irritability at Wave 2, r (168) =

.36, $p<.005$. Boys were significantly more likely than girls to score higher on the infant distress to limitations scale, and mothers of boys were significantly more likely than mothers of girls to be irritable at 6 months post-partum. The role of infant gender as a potential moderator of the relationship between mothers' concurrent irritability and infant distress to limitations was analysed using a hierarchical regression analysis. Both mothers' irritability and infant gender continued to predict infant distress to limitations indicating that infant gender was not a moderator in the relationship between mothers' concurrent irritability and infant distress to limitations. Simply put, the relationship between the mothers' concurrent irritability and the infants' distress to limitations scores was not influenced by whether the infant was a boy or a girl. The results are shown in Table 4.23.

Table 4.23

Summary of Hierarchical Regression Analysis to test the potential moderating role of Infant Gender in the Relationship between Mothers' irritability at Wave2 and Infant Distress to Limitations (N = 180)

	B	*SE B*	*β*
Step 1			
Constant	2.49	0.12	----
Maternal irritability at W2	0.89	0.18	0.36**
Step 2			
Constant	2.08	0.18	----
Maternal irritability at W2	0.83	0.17	0.34**
Infant gender	0.29	0.10	0.22**

Note. $R^2 = .13$ for Step 1; $\Delta R^2 = .05$ for Step 2 ($p < .05$).
**$p<.005$; *$p < .05$.

4.5 Discussion

The findings from Study 2 inform the research on the relationship between mothers' psychopathology and infant temperament (Edhborg, Seimyr, Lundh, & Widstrom, 2000; Vaughn et al., 1987). The first aim of the study was to explore the relationship between irritability and emotional and behavioural disorders at a later stage of the lifespan to the preschool age group examined in Study 1. The stability of both enduring characteristics of anger demonstrated in the personality literature (Nigg, 2006) and the stability of psychopathology across the lifespan demonstrated in the psychiatry literature (Loeber & Hay, 1997), suggested that irritability may be an enduring characteristic that may continue to influence disorder across the lifespan. Within this study, mothers' dispositional irritability was significantly associated with both emotional and behavioural disorders. Using an established clinical measure of mothers' history of depression and anxiety disorders up to and including pregnancy, the results indicated that mothers' irritability was significantly related to mothers' depression and anxiety disorders. This prediction continued to be significant even when other predictors were taken into consideration. Adult depression and anxiety disorders have been related to past conduct disorder symptoms (e.g., Hay, Pawlby, Waters, Perra, & Sharp, 2010; Romano, Zoccolillo, & Paquette, 2006; Zoccolillo, Pickles, Quinton, & Rutter, 1992) and to social circumstances, such as social class and education. The present study results supported those previous findings.

In Study 1, irritability had mediated the relationship between internalising and externalising symptoms in preschool children. The potential role of irritability as a mediator between mothers' past conduct symptoms and mothers' experience of depression and anxiety

disorders was examined within Study 2, but the results did not reflect the mediating role of irritability found with preschoolers. Irritability predicted mothers' experience of depression and anxiety disorders independently of the influence of mothers' social circumstances and mothers' past conduct symptoms. What was apparent from the analyses in Study 2 was that the mothers who were diagnosed with depression and anxiety disorders in pregnancy, and the mothers who had both experienced past depression and anxiety disorders and were in current episode in pregnancy, had reported more irritability than those who had only been depressed or anxious in the past, or who had been free of disorder.

The stability of irritability across two time points suggests that irritability was an enduring characteristic for some mothers. The mothers' self-reports of irritability were validated by fathers' reports on the mothers' conflictual behaviour within the partner relationship. This suggests that, as an enduring characteristic across the transition to parenthood, the mothers' irritability posed a risk to the development of infant irritability through intergenerational transmission and through a higher level of conflict in the family environment. Before discussing the analyses on the intergenerational transmission of irritability from mother to infant, however, it was important to place infants' irritability into context.

The second aim of the present study was to examine infant irritability using a measure of infant temperament that could be theoretically linked to the construct of irritability. The IBQ 'distress to limitations' scale(Rothbart, 1981) measures infant irritable distress, asking the informants to report on the infants' distress in relation to salient everyday events occurring in the past week. The IBQ 'distress to limitations' scale was shown to be an internally reliable scale within this present study, and the relationship between infant 'distress

to limitations' and the IBQ infant fear and infant activity level, reflected similar relationships found in previous temperament studies (Colder, Mott, & Berman, 2002; Rothbart & Bates, 1998). There was very good agreement on the IBQ scales between the three informants used within this study, which allowed the use of a composite temperament score to improve the availability of infant temperament data for the analyses within this present study. The results of the combined IBQ 'distress to limitations' scale did not show any convergence with the independent observation of infant distress in relation to an everyday home based challenge. The examination of mothers' reports on the IBQ 'distress to limitations' scale in relation to the independent observations of the infants' distress in response to the everyday home based challenge (car seat restraint) did show a significant relationship between the mothers' reports of infant distress to limitations and the observers reports of infant whining and whinging. Whilst some caution needs to be considered due to the non-parametric analyses used, this is an important result, as it may reflect the importance of measuring mothers' reports of infant temperament with relevant observation measures that draw on the contexts within which infants' distress is more salient to mothers' (Hane et al., 2006). It is therefore worth exploring in future studies with additional observational challenges that reflect home based everyday experiences. Unfortunately, there was some missing data from mothers' reports of infant temperament at Wave 2, and examination of mothers' who did and did not provide the temperament reports on their infants indicated that mothers who did not provide infant temperament data at Wave 2 were more likely to be irritable than mothers who did provide infant temperament data.

Mothers' reports of infant anger and temper tantrums were also examined in relation to the infants' vocal distress in the car seat restraint. Interestingly, mothers' reports of infant

anger and temper tantrums were related to infant crying in response to the car seat restraint. The different associations between reports of infant distress to limitations and reports of infant anger and temper tantrums may possibly reflect a fine differentiation within the construct of irritability. Further discussion on these results and the limitations of the present study will be discussed in the general discussion in the next chapter.

The third aim of the present study was to examine the potential intergenerational transmission of irritability from mother to infant. Firstly, longitudinal analyses were carried out to examine the effect of mothers' irritability in pregnancy on the infants' irritability at 6-months. As mothers' irritability was related to both mothers' social circumstances and mothers' mental health, these variables were taken into account in the regression model used to examine the prediction from mothers' irritability to infant irritability. When mothers' social class, education, past conduct symptoms and history of depression and anxiety disorders were all taken into account, mothers' irritability in pregnancy significantly predicted infant irritability. Previous studies have suggested that mothers' own personality will influence their perceptions of the infant's temperament; therefore within this present study a measure of infant irritability combined across informants was used to reduce potential maternal bias. As mothers' irritability was found to be stable across the transition to parenthood, it was also necessary to examine the potential mediating role of concurrent maternal irritability in the relationship between mothers' irritability in pregnancy and the infants' irritability. The analyses revealed that mothers' concurrent irritability when the infant was 6 months old mediated the relationship between mothers' irritability in pregnancy and the infants' irritability at 6-months. There was a difference in irritability between boys and girls; therefore the potential role of infant gender as a moderator in the relationship between

mothers' concurrent irritability and infant irritability was examined. Infant gender was not found to moderate the relationship between mothers' concurrent irritability and infant irritability. Both mothers' concurrent irritability and infant gender were related to infant irritability. Discussion about these results in relation to previous studies and its implications for future research will be set out in Chapter 5.

The results of Study 2 suggest that irritability is an enduring characteristic of individuals that continues to have an influence on disorder in adulthood, and appears to be transmitted between mother and infant by the time the infant is 6 months old. The final chapter will discuss the results from the present study and the results from Study 1 within the theoretical framework of developmental psychopathology, and the ecological model of development. The results from both studies have important implications for future research on the role of infant temperament in the influence of later disorder, and these implications and suggestions for further research will be discussed in Chapter 5. Inevitably, there were parts of the present research that could have been done differently, and the limitations of both empirical studies within this thesis will be discussed in Chapter 5, with suggestions about future research models to continue the exploration of irritability.

CHAPTER 5

General Discussion

5.1. Introduction

This thesis was born from the discovery that some children as young as 5 were suffering from depression (Carlson & Kashani, 1988; Earls & Jung, 1987; Luby, 2002; Luby et al., 2003a; 2003b) and other very young children were excluded from school because of behavioural problems (Gilliam, 2005). An exploration into the psychiatry literature revealed irritability as a common symptom in very young children who suffered from these debilitating conditions (Luby et al., 2003a; 2003b). Further exploration into the psychology literature revealed irritability as a temperament construct that posed a risk to the development of both emotional and behavioural problems in children (Lengua, 2003; Oldehinkel et al., 2004). It was apparent that the field of developmental psychopathology was working hard to understand the reasons for the development of disorders in children (Gilliam & Shaw, 2004; Moffitt, Caspi, Cicchetti, & Cohen, 1995; Rutter & Sroufe, 2000; Rutter, Tizard, Yule, Graham, & Whitmore, 1976), mapping pathways from early characteristics such as temperament to the diagnoses of disorders in children and adolescents.

The plight of children was also the focus of Bronfenbrenner (1972), who carried out a series of cross-cultural studies on bringing up children within the USA and USSR. From this work, Bronfenbrenner developed an ecological development model to assist society to understand the many influences and interactions that occur within the

development of a child (Bronfenbrenner, 1977). Bronfenbrenner proposed a broad ecological approach to research in human development because he believed that such research at that time had pursued a divided course that was not helpful to scientific progress.

The field of developmental psychopathology and Bronfenbrenner's ecological framework, both rooted in systems theory (von Bertalanffy, 1968), provided the broad framework for this thesis, to pursue the study of irritability. The theories and frameworks provided a guide to the examination of both literatures to explore the origins of irritability and its' potential role in the pathways to disorder. A research framework, based on systems theory, will reflect the complexities and challenges of a complex system that at times may be too difficult to examine. Researchers in personality theory grappled with the concept of systems theory for this very reason (Mayer, 1993). The challenge for researchers who are working within systems theory is the recognition that study of the whole system at one time is not required and work on one part of the system will inform other researchers working on other parts of the system.

The review of the literature on irritability within the psychology and psychiatry literature demonstrated developmental pluralism: different pathways that were not necessarily going to achieve equifinality, a common outcome. Complexities had come into the research on temperament and disorder: the conundrum of the relationship between temperament and disorder, i.e., whether this was a true relationship or a matter of measurement confounding (Lemery et al., 2002; Lengua et al., 1998; Nigg, 2006; Sanson et al., 1990); the variety of descriptors and measures used to describe similar constructs, such as irritability, within the temperament literature (Rothbart et al., 2000);

and the debate about the role of mothers as developmental influences on their infant temperament or as unreliable informants (Crockenberg & Acredolo, 1983; Vaughn et al., 1987). Unravelling these complexities to understand the origin, nature and role of irritability in the development of childhood disorders was the central aim of this thesis.

The methods used within the two empirical studies in this thesis were based on Bronfenbrenner's (1977) research approach set out in nine propositions. Bronfenbrenner suggested that not all nine propositions need to be adhered to at the same time, as they relate to different nested systems and not all the systems have to be studied simultaneously. Again, a pragmatic approach to research on systems was adopted. A review of the methods used by early researchers of temperament and disorder (Shirley, 1933; Thomas & Chess, 1977), led to the 'observe and describe' approach adopted within this thesis. Shirley's assessment of the infants' irritability aimed to record just what the babies did, 'scream, cry, fuss,' during many types of examination. Thomas and colleagues chose not to ignore their clinical observations of a lack of simple relationships between environmental circumstances and their consequences, and launched the landmark studies on children's temperament (Thomas et al., 1968). Thomas and colleagues interviewed parents at length about their children's behaviour at three-month intervals from 3- to 18-months. Informed by these two important pioneering studies on temperament, the 'observe and describe' approach has been adopted in the two studies within this thesis. Observe, used in this context, is not only about observation as a method, but also about noticing, perceiving and detecting relationships within the course of this study, using a range of methods and different informants. The findings from both

studies provide descriptions of the origins, nature and role of irritability in relation to disorders at different stages of the lifespan.

The first study, the Starting School study (reported in Chapter 3), focused on children aged 3 to 5 years in their first years in education. The second study, using a subsample of the Cardiff Child Development Study (CCDS; reported in Chapter 4), focused on adult women as they made the transition from pregnant woman to first-time mother with a 6-month-old infant.

Study 1 aimed to test the measurement confounding hypothesis in a preschool-aged population to further the psychology and psychiatry literature. Egger and Angold (2006), in a review of the presentation, nosology, and epidemiology of common emotional and behavioural disorders in preschool children, had suggested that a concurrent analysis of temperament constructs and psychiatric symptoms would allow for the examination of measurement overlap. Following the test of the measurement confounding hypothesis in Study 1, the relationship between irritability and both internalising and externalising symptoms was explored. Previous research suggested that ODD was a mediator between depression and conduct disorder as well as between other disorders in preschool-aged children (Egger & Angold, 2006). Similarities between irritability and ODD suggested the need to examine ODD separately to other externalising disorders, and to test if irritability acted similarly to ODD as a mediator between internalising and externalising symptoms in preschool children.

Study 2 had three related aims. The first aim was to test the relationship between irritability and disorder at a later stage of development, in this case, as an adult woman. Previous research had suggested that irritability may be a stable characteristic across the

lifespan (Durbin et al., 2007; Lemery et al., 1999; Pedlow et al., 1993; Putnam, Rothbart, & Gartstein, 2008), and that early disorders were predictive of later disorders across the lifespan (Costello, Egger, & Angold, 2005). These two observations led to the first hypothesis tested in Study 2 that irritability may continue to influence disorder across the lifespan. Studying women at the point of transition to new motherhood also allowed the exploration of the second aim of the study, the examination of irritability at an early stage in the development of an infant. Previous research on the stability of infant irritability suggested that it was prudent to examine infant irritability at 6 months (Lemery et al., 1999). Using a temperament measure that measured a single dimension of irritability, the nature of infant irritability was described in relation to other temperament dimensions and observed infant distress. The focus on the transition to new motherhood in Study 2 enabled the test of the final aim, the potential for the intergenerational transmission of irritability between mother and infant.

Within this chapter, the findings from the two studies are discussed within the context of developmental psychopathology and Bronfenbrenner's ecological model of development. Inevitably, both empirical studies had their limitations and these too are discussed with suggestions for future improvements in further studies. Finally, the implications of the findings are discussed; in particular, the implications for future research, for theory and policy development, and the implications for practice.

5.2 *Study 1: The Role of Irritability as a Symptom of Disorder in Preschool Children, and a Potential Factor in the Development of Disorder: Summary of Key Findings*

A review of the dual role of irritability as a factor in the development of disorders and as a symptom of disorder in childhood was set out in Chapter 2, focusing on the psychology

and psychiatry literature. The psychiatry literature provided a useful definition of irritability that was adapted for this thesis to reflect the psychology and psychiatry perspectives, 'an episode and/or enduring behaviour characterised by reduced control over temper which usually results in irascible (hot-tempered, angry) verbal or behavioural outbursts' (adapted from Snaith & Taylor, 1985).

There was concern amongst researchers that an apparent relationship between temperament and disorder was a measurement-confound rather than a true relationship (Egger & Angold, 2006; Nigg, 2006). To take the research beyond the concerns about measurement confounding, Study 1 sought to follow previous studies (Lemery et al., 2002; Lengua et al., 1998; Nigg, 2006; Sanson et al., 1990) and examine the measurement-confounding hypothesis in the relationship between irritability and symptoms of externalising and internalising disorders in preschool children.

To my knowledge, this study was the first to examine the measurement-confounding hypothesis in a preschool-aged sample, using a measure of irritability based on the Rothbart temperament definitions, but derived from two psychiatric instruments to assess symptoms of internalising and externalising disorders, a screening questionnaire (the SDQ; Goodman, 1997) and a clinically based psychiatric interview, the PAPA (Egger et al., 2002). The study took place in two phases, with teachers reporting on the children's behaviour using the SDQ in Phase 1 and, in Phase 2, the parents (90% mothers) reporting on the children's behaviour using both the SDQ and responding to detailed clinical questions in the PAPA interview.

The results from Study 1 indicated that there was some measurement confounding that accounted for the relationship between irritability as defined within the

temperament literature (Goldsmith, 1989; Rothbart, 1981; Rothbart, 1989) and a composite externalising problems measure which was derived from the SDQ conduct problems and hyperactivity problems scales. The SDQ contains one item on the conduct problems scale, 'often has temper tantrums or hot tempers', which is a screening item for symptoms of ODD. The presence of this item may be the reason for the apparent confound with the irritability scale. Irritability continued to relate significantly to the SDQ composite externalising problem measure when the temper tantrum item was removed, suggesting that measurement confounding accounted only for part of the relationship between irritability and externalising problems, as measured using the SDQ.

Measurement confounding also accounted for some of the relationship between irritability and internalising symptoms as measured using the PAPA (Egger et al., 2002). Irritability items in the PAPA internalising symptom scale were found in the GAD symptom scale, as defined using DSM-IV-TR (2000) criteria. Whilst irritability has been identified as a symptom of depression in children (Luby et al., 2003), the DSM-IV-TR (2000) criteria for MDE did not include irritability symptoms. The relationship between irritability and the pure PAPA internalising symptoms remained significant when the confound items were removed from the PAPA internalising symptom scale, suggesting that measurement confounding was the not the complete answer to the relationship.

Analyses of the ODD symptoms indicated that 3 symptoms out of ten were potential confounds with irritability. Measurement confounding was found to account for part of the relationship between irritability and the ODD symptom scale, but again the relationship between irritability and the pure ODD scale remained, suggesting that

irritability does relate to other ODD symptoms. The overall findings about measurement confounding between irritability and symptoms of disorder in preschool children suggest that when measurement confounding is taken into account the relationship between irritability and internalising symptoms, the relationship between irritability and externalising symptoms, and the relationship between irritability and ODD symptoms continues to be a significant finding.

This conclusion corroborates previous research on measurement confounding (Lengua et al., 1998). Lengua and colleagues found that decontamination of the negative emotionality scale for irritability items did not account for the total relationship between negative emotionality and symptoms of both conduct problems and depression in children. The findings in Study 1 go further than the conclusions from Lengua and colleagues in understanding the relationship between irritability and both internalising and externalising symptoms of disorder in preschool children. Lengua removed some irritability items from the higher-order negative emotionality construct that includes fear and sadness dimensions. The fear and sadness dimensions could also be confounds with depression, but the use of the higher-order factor will have possibly masked the relationships between the individual dimensions of irritability, fear, and sadness with the conduct problems and depression measures. The focus on the single dimension of irritability within this thesis informs the psychology and psychiatry literature, through the specificity of the potential confound items on two measures of symptoms of disorder (the SDQ and the PAPA), so that future researchers can be aware of and take account of when using these measures to assess the relationship between temperament and symptoms of disorder.

Lemery and colleagues tested the measurement-confounding hypothesis for the relationship between irritability and behaviour problems using the CBQ (Rothbart et al., 1994), which defines irritability in the same way as the irritability scale used in the present study. Lemery and colleagues compared irritability with original and purified measures of behaviour problem symptoms using the PBQ (Behar & Stringfield, 1974), and concluded that measurement confounding did not account for the relationship between irritability and problem behaviours. Lemery and colleagues measured behaviour problems that predicted to later symptoms of disorder measured using a clinically based psychiatric measure the HBQ (Ablow et al., 1999), but did not examine the measurement-confounding hypothesis between the CBQ and the HBQ. The findings of Study 1 add to Lemery and colleagues' work through the testing of the measurement-confounding hypothesis using a clinically-based psychiatric assessment of symptoms of disorder. The findings in this study indicate that there is a relationship between irritability and symptoms of disorder using a clinically-based psychiatric measure, the PAPA.

The relationship between irritability and ODD was an important finding within this present study. A review of the literature on ODD indicated that approximately 7% of preschoolers with some psychiatric disorder were diagnosed with ODD (Egger & Angold, 2006), and ODD was related to high levels of co morbidity (Maughan, Rowe, Messer, Goodman, & Meltzer, 2004). Egger and Angold examined the role that ODD played in the co morbidity between disorders in preschool children. ODD was found to mediate the relationship between depression and CD, and between depression and ADHD. These previous findings, coupled with the knowledge that irritability was a shared symptom of internalising and externalising disorders in children (Luby et al.,

2003a; 2003b), guided the analyses for Study 1, to assess the potential mediating role of irritability in the relationship between internalising and externalising symptoms in preschool children. These analyses showed that irritability mediated the relationship between internalising and externalising symptoms in preschool children.

The question that arose from the finding that irritability was a mediator between internalising and externalising symptoms was whether this result explained the previous research findings that ODD mediates internalising and externalising co morbidity. Using the pure ODD scale (with irritability items removed), the pure ODD scale was also found to mediate the relationship between internalising and externalising symptoms within Study 1. These findings support the results found by Egger and Angold (2006). The role of irritability as a mediator between internalising and externalising symptoms in preschool children appears to be independent of the role that other symptoms of ODD play as a mediator between internalising and externalising symptoms.

There are two conclusions about ODD that can be drawn from the results of Study 1. Firstly, whilst there is a significant relationship between irritability and ODD, the relationship also occurs between irritability and the non-irritability symptoms of ODD. Secondly, ODD appears to act independently of irritability in its relationship between internalising and externalising disorders, suggesting that ODD is not likely to be a clinical manifestation of high irritability as would be suggested by the spectrum model to explain the relationship between irritability and disorder. The importance of irritability in relation to both internalising and externalising symptoms, including ODD, suggests that the relationship between irritability and symptoms of disorder in children is better

explained by the vulnerability model (Nigg, 2006), with irritability being a vulnerability factor in the potential development of disorders in childhood.

Having examined the role of irritability in relation to childhood disorders, evidence of continuity in both irritability and psychopathologies across the lifespan (Komsi et al., 2006; Riese, 1987; Rutter, Kim-Cohen, & Maughan, 2006b; Stringaris & Goodman, 2009) led to the hypothesis examined in Study 2 that irritability would continue to be an influential factor in the development of disorders over the lifespan. The findings of Study 2 are now discussed.

5.3 *Study 2: Understanding the Role of Irritability in Adult Disorders, the Nature of Infant Irritability, and the Potential Intergenerational Transmission of Irritability from Mother to Infant: Summary of Key Findings*

Three hypotheses were tested in Study 2: (1) that irritability may continue to play a role in the relationship between disorders in adulthood; (2) that irritability in infancy may be related to other dimensions of temperament that predict later disorder; and (3) that there would be evidence for intergenerational transmission of irritability between mothers and infants. Firstly, mothers' irritability was explored at two time points, pregnancy and 6 months after the birth of the first child. Mothers' irritability was found to be stable across the two time points suggesting that irritability in these adult women was a dispositional characteristic. The mothers' self-reports of irritability across the two time points were supported by the fathers' reports of the mothers' conflictual behaviour. The finding that adult irritability was stable over time has been supported in a previous study (Stringaris & Goodman, 2009; Stringaris, Cohen, Pine, & Leibenluft, 2009).

Within Study 2, analyses were carried out to show what other problems mothers with irritability may have, because other problems may contribute towards the mothers' mental health status. Previous research on mothers' emotional and behavioural disorders has also shown that mothers' socio-demographic characteristics influence the mothers' mental health (Hay et al. 2010). The socio-demographic characteristics of the mothers' in Study 2 were considered in the subsequent analyses to understand the role those mothers' irritability plays in relation to mothers' emotional and behavioural problems. Analyses in Study 2 showed that mothers' irritability was related to mothers' social class, mothers' education and whether or not mothers smoked during pregnancy. Working class mothers and mothers with education standards below the UK standard required for A' level study were more irritable. Mothers' irritability was not related to whether or not the mothers' drank alcohol in pregnancy.

Analyses of the differences between mothers who had emotional disorders (anxiety and/or depression) in pregnancy and those who were emotionally well in pregnancy indicated that mothers with emotional disorders in pregnancy were more likely to be irritable, and mothers who had emotional disorders in the past as well as in pregnancy were also more likely to be irritable than mothers with good past and present emotional health. These results suggest that irritability is elevated during episodes of emotional disorders.

When the relationship between mothers' past behavioural symptoms and mothers' emotional disorders was examined, mothers' dispositional irritability predicted caseness for emotional disorders, when other maternal variables were taken into account (i.e. mothers' social class, mothers' education and mothers' smoking in pregnancy). The

higher levels of irritability in mothers with emotional disorders and the finding that irritability predicts emotional disorders may explain the role of irritability as both a dispositional factor influencing disorder and a symptom that is elevated during episodes of emotional disorder.

Complementary analyses on the relationship between mothers' emotional disorders and past behavioural symptoms revealed that mothers' social class, mothers' education, and mothers' emotional disorders predicted mothers' past behavioural symptoms, and these relationships remained significant but were attenuated when mothers' dispositional irritability was added to the prediction model. These results suggest a multi-factorial model of influence for the prediction of behavioural symptoms, in line with the vulnerability model (Gilliom & Shaw, 2004; Nigg, 2006). Adult depression and anxiety disorders have been related to past conduct disorder symptoms (e.g., Hay et al., 2010; Romano et al., 2006; Zoccolillo et al., 1992) and to social circumstances, such as social class and education (Jenkins, Rashbash, & O'Connor, 2003). The present study results supported those previous findings.

Irritability within adult women appears to be a dispositional characteristic that plays a part in the influence of both emotional and behavioural problems in adulthood. In comparison with the role irritability plays in the relationship between parallel childhood disorders, irritability in adult women does not appear to mediate the relationship between emotional and behavioural problems. The different relationship in adulthood may be a developmental issue in that by adulthood the pathways for emotional and behavioural problems may become more distinct (Rutter & Sroufe, 2000). The findings from this present study reflect the early thinking of Thomas and colleagues (1968), namely, that

temperament in itself does not constitute a negative versus positive adjustment, but that temperament conditions a developmental process that determines adjustment. This concept is suggested to be more fitting with a vulnerability model rather than a spectrum model (Rothbart & Bates, 1998). Irritability would thus be an early diathesis that may be influenced over time by environment and by other temperament dimensions. The early assessment of other infant behaviours that are associated with infant irritability was explored in Study 2.

From the literature review in Chapter 2, it was possible to map the variety of descriptors used by different temperament researchers for the irritability construct, and the measures derived for assessment of these descriptors (see Table 2.1). The distress to limitations dimension of the IBQ (Rothbart, 1981) was the measure that was theoretically derived from both Shirley's and Thomas and colleagues early work in this field. Whilst many researchers on temperament combine the distress to limitations dimension with the fear dimension from the IBQ to produce a global negative affectivity dimension, the distress to limitations dimension is the most precise measure of temperamental irritability and was therefore used for assessment of infants' irritability within Study 2. The combined negative affectivity dimension would have confounded irritability with fear, which as a dimension of temperament demands its own individual attention in its relation with later disorders (Kagan, Snidman, Zentner, & Petersen, 1999). Using the IBQ in Study 2 enabled the exploration of relationships between distress to limitations as the measure of infant irritability, and the fear (distress latency scale) and activity level dimensions on the IBQ. Rutter and colleagues (1976) found that a combination of risk factors can significantly increase the likelihood of childhood disorders, and specific

temperament profiles have been suggested as risk for particular childhood disorders (Colder, Mott, & Berman, 2002; Nigg, 2006). In Study 2, infants' distress to limitations was significantly related to infants' fear and activity level, supporting previous studies that have examined relationships across temperament dimensions (Rothbart & Bates, 1998). Combinations of high activity and irritability have been suggested as potential temperamental liabilities to conduct disorder and co morbid internalising and externalising disorders (Nigg, 2006). Within Study 2, infant distress to limitations was also significantly related to reports of infants' temper tantrums and angry moods, adding support to the operationalised measures of irritability derived from the PAPA in Study 1, and used to measure age appropriate irritability within the older age group measures of temperament, the TBAQ and CBQ (Goldsmith, 1989; Rothbart 1989).

The validity and reliability of mothers' reports of infant temperament continues to be an issue of debate amongst psychological researchers (Edhborg et al., 2000; Richters & Pellegrini, 1989; Stifter et al. 2008). Rothbart and colleagues (1981; 1989) have paid great attention to detail in designing measures for parents to report on infant temperament, and designed tasks that can be used for independent observation of the same constructs that are measured using the parent reports (Lab-TAB; Goldsmith & Rothbart, 1996). The low levels of convergence across measures and informants on infant temperament dimensions have been discussed as evidence for mothers' unreliability in reporting on their infants' temperament. Some studies have demonstrated that the lack of convergence may be due to the different measures assessing different constructs (Pauli-Pott et al., 2004; Pauli-Pott et al., 2000; Stifter et al., 2008), and that the experiences that mothers have with their infants are not being replicated in observational challenges (Hane

et al., 2006). When challenges are closer to the everyday experiences that mothers share with their infants, good reliability has been found between mothers' reports and independent observations of infants' behaviour (Hane et al., 2006).

Within Study 2, the car seat restraint task from the Lab-TAB was used as an everyday restraint task in the home observation of the infant at 6 months. There was very little infant distress shown across the sample in relation to this minor aversive stimulus. The car seat task had been used previously with 6-month-old infants (Bridges et al., 1993) and had shown good agreement between mothers' reports of infant distress to limitations and observed infant anger in the laboratory. The car seat task had also previously been used in a home based assessment with 8 month-old infants (Hane et al., 2006). Within Study 2, mothers' reports of infant distress to limitations were related to infant whining and whinging within the car seat. Rothbart and Bates (1998) indicated previously that there may be a fine differentiation to make within measures of irritability. It was suggested that sensitivity to minor aversive stimuli may predispose a child to whining and withdrawal, whereas irritability to frustration of reward or stimulation-seeking behaviour would be more likely to pertain to more externalising tendencies. The finding in Study 2 that reports of infant anger and temper tantrums were related to infant crying in the car seat would appear to support the fine differentiation in measures of irritability that Rothbart and Bates had previously suggested.

Infant irritability within this present study co-occurs with fear and high activity levels, and is related to both reports of infant anger and temper tantrums and observed whining and whinging in response to a minor aversive event. Infant irritability as defined in the 'distress to limitations' dimension appears to be finely differentiated from reports

of infant anger and temper tantrums as demonstrated by the relationship between infant anger and temper tantrums and infant crying in response to the car seat restraint.

The final aim of Study 2 was to examine the intergenerational transmission of irritability between mother and infant. Both the psychology and psychiatry literature indicated the importance of the influence that mothers' characteristics and mental health can have on the adaptation of her children (e.g., Hay et al., 2008). Again methodological concerns have been expressed about the true relationship between mothers' characteristics and mothers' mental health and mothers' subsequent reports about their children's' temperament (Vaughn et al., 1987; Vaughn et al., 2002). Previous studies into the predictors of irritability suggested that children born to mothers with mental health problems, such as depression and anxiety (Austin et al., 2009; Gjone & Stevenson, 1997), were at risk of being more irritable than infants born to mothers without mental health problems. Other maternal factors have also been suggested as influential in predicting infant temperament, including mothers' socioeconomic circumstances, and mothers' use of alcohol in pregnancy (Lemola, Stadlmayr, & Grob, 2009).

Within Study 2 mothers' alcohol use in pregnancy and smoking in pregnancy were not significantly correlated with infant irritability. When social indicators (mothers' social class and mothers' education) were accounted for, only mothers' dispositional irritability continued to predict infant irritability.

An important finding in Study 2 was the mediating role of mothers' dispositional irritability in the relationship between mothers' emotional disorders and infant irritability. The reports of infant irritability were combined reports from three informants. The relationship found in Study 2 between mothers' dispositional irritability and infant

irritability does not support the conclusions from Vaughn and colleagues (2002; 2003), that such a relationship indicates a reflection of mothers' characteristics and mental health rather than the infants' characteristics. Vaughn and colleagues' conclusions were formed without considering that the ITQ (Carey 1970) measure of infant difficulty includes fear and irritability, and that GAD (DSM-IV-TR; 2000) includes irritability as a main symptom. The relationship between mothers' anxiety and infant difficulty could be explained by the finding in Study 2 that mothers' irritability predicts to infant irritability.

There is evidence that irritability is heritable (Henderson, 1982), but also that genetic influence may increase over time, suggesting that family environment can augment familial resemblance when family members share environment as well as heredity (Plomin et al., 1988). It was therefore important to test the potential intergenerational transmission of irritability between mother and infant. Mothers' irritability in pregnancy and mothers' concurrent irritability at 6 months post childbirth both predicted infant irritability at 6 months. Mothers' irritability at 6 months was found to mediate the influence of mothers' irritability in pregnancy on infant irritability. These results suggest that mothers' irritability appears to be transmitted to their infants by 6 months.

Previous studies have found differing results on the relationship between gender and infant irritability, although the construct measured has differed across studies, with 'difficulty, emotionality, anger and frustration, or distress to limitations', used to describe irritability (Else-Quest et al., 2006; Hane et al., 2006; Pauli-Pott et al., 2003). No gender differences were found in preschool children's irritability in Study 1, but gender differences were found in infant distress to limitations. Boys were reported to be

significantly more irritable than girls at 6 months. To explore possible infant effects on infant irritability, infant gender was explored as a potential moderator of mothers' irritability in the relationship with infant irritability. Infant gender did not moderate the effects of mothers' irritability on infant irritability, suggesting that both mothers' concurrent irritability and infant gender were independently related to infant irritability. From the results of the present study, it is not possible to conclude the bidirectional influences among child gender and maternal and infant irritability. A further study to explore the potential interactional effects would be possible in the future, using later assessments from the Cardiff Child Development Study.

Both empirical studies within this thesis have revealed information about the nature, origin and role of irritability in disorders at three stages within the lifespan: during early infancy, early childhood, and adulthood at the transition to first time motherhood. The findings from both studies add to both the psychiatry and psychology literature, with more detail about the measurement of irritability in relation to symptoms of disorder, the role that irritability plays in relationship to both emotional and behavioural disorders at different points in the lifespan, the description of irritability and its potential for further differentiation, and the intergenerational transmission of irritability between mothers and infants. Longitudinal research involving children and families provides considerable methodological challenges. This inevitably means that there were areas for improvement within each study. The limitations of the two empirical studies are discussed with a view to informing future studies.

5.4 *Limitations of Present Studies*

Study 1 was an exploratory study with a relatively small sample size. Despite a small sample compared with other community samples, the sample allowed for in-depth interviews with parents about their children's mental health, and was of sufficient size to produce statistically meaningful results. Whilst the sample was not of sufficient size to develop full diagnosis of disorders and compare individual cases by diagnosis in relation to the irritability, the study provided information about children's symptoms of disorder from two informants, teachers and parents, allowing for analyses of reliability in reports of problems within the children.

The community sample recruited in Study 1 was slightly under-representative of the UK general population, with respect to children in lower income families and in single-parent households; and previous research has indicated that children in lower socioeconomic groups are at greater risk of disorder (Gilliom & Shaw, 2004; Rutter, 2000; 2009). It is likely that the presence of symptoms of disorder within the Study 1 sample may be an underestimate of the problems within the UK general population. The analyses of the teachers' SDQ reports between the samples of children studied at phase 1 and phase 2 of the study indicated that those children in families who participated in phase 2 had similar patterns of problems as those who participated in phase 1. Whilst this provided some confidence that attrition in the study was unlikely to affect the outcome of the study, further socio-economic data collated at phase 1 of the study would have provided a greater degree of confidence in the representative nature of the sample.

Unfortunately, there has not yet been an epidemiological study of behavioural and emotional psychiatric disorders in preschool children within the UK. A recent review

from the USA has indicated that prevalence rates for both emotional and behavioural disorders in preschool children reflect those of older children (age 5-17; Egger & Angold, 2006). Variations occur within specific emotional and behavioural disorders, with rates of depression increasing with age, the specificity on anxiety disorders moving from Separation Anxiety Disorder in the early years to more generalised anxiety in the later years and rates of ODD and CD reducing with age. This exploratory study has been able to demonstrate that the PAPA can be used reliably within a community sample within the UK to assess preschool children's mental health.

Despite the sample limitations, the results from Study 1 add to the psychology and psychiatry literature in three key ways: firstly, by demonstrating that measurement confounding is not a sufficient explanation of the many findings across the literature of a relationship between temperament and disorder; secondly, by demonstrating that ODD symptoms appear to mediate the relationship between internalising and externalising symptoms, a finding that extends earlier findings (Egger & Angold, 2006) to a British sample; and, finally, by demonstrating that irritability mediates the relationship between internalising and externalising symptoms, corroborating a previous study that linked irritability with severity of disorder and co morbidity (Oldehinkel et al., 2004),

Study 1 was not a longitudinal study and therefore it is difficult to conclude the direction of influence between irritability and symptoms of disorder in preschool-aged children. Egger and Angold (2006) had identified the need for a concurrent study that examined temperament and symptoms of disorder in preschool children. Whilst a specific temperament measure was not used to assess irritability in preschoolers within Study 1, an irritability scale was derived from an in-depth parental interview, and the irritability

construct was measured using operational definitions according to the Rothbart temperament tradition. Study 1 was therefore able to extend the literature in line with the suggestions from Egger and Angold (2006), to test the measurement-confounding hypothesis through concurrent analyses of irritability and symptoms of disorder.

Study 2 was built upon the findings from Study 1 that irritability plays a role in the relationship between internalising and externalising symptoms in early childhood. Previous examination of the continuities and discontinuities of psychopathology between childhood and adult life suggests that one disorder usually begins before the other, in a pattern of sequential co morbidity, and that early emotional disorders are related to later expressions of adult emotional disorders, and early behavioural problems are strongly associated with an increased risk for psychiatric disorder later in life (Rutter et al., 2006). The relationship between irritability and disorder was therefore examined in an adult female population in Study 2. The questionnaire measure used to gain information about the mothers past conduct problems was devised for the purpose of the CCDS to identify mothers retrospective reports of their behaviours defined according to DSM-IV conduct disorder symptoms. Whilst the scale had acceptable levels of internal consistency and validity in relation to the DSM-IV symptoms, and showed consistency in its relationship with the IPDE screening responses that measure antisocial personality disorder, the scale requires further testing for reliability and validity in future studies.

The sample for Study 2 was a subsample of the CCDS, a prospective longitudinal analysis into the early precursors to violence that began with recruitment of the sample in November 2005. Two time points from the CCDS study were used for the purposes of Study 2, the antenatal assessment of women, and the 6-months postnatal assessment of

women and their first-born infants. The sample for Study 2 included women and their infants born before end December 2007. There was some attrition within the sample at the second time point (Wave 2), and some missing data from mothers on their questionnaire reports of infant temperament at Wave 2. Fortunately, three informants were asked about the infants' temperament at Wave 2, and therefore data are available for the infants' temperament from 82% families eligible for inclusion within Study 2. This attrition rate is in line with other longitudinal studies (e.g. Stifter et al., 2008). Assessment of the sample characteristics for Study 2 indicated that the sample was comparable to first-time parents in the Millennium Cohort Study sample (MCS; Kiernan, personal communication). The mothers who did not provide data about their infants' temperament were more likely to be working class, have left education at 16 years, and were significantly more irritable than mothers who did provide infant temperament data. This attrition is likely to attenuate the number of infants with reported irritability in Study 2.

The limitations of the sample in Study 2 were alleviated to some extent through the use of multi-informants and multi-methods to assess infant irritability. There was good agreement between informants on the infant temperament dimensions of the IBQ and the correlations between informants were better than those previously published (Rothbart, 1981; Gartstein & Rothbart, 2003). The IBQ scale reliabilities for each of the temperament dimensions used were higher than those published in previous studies. The sample size used in Study 2 compared favourably with other studies of infant temperament and other longitudinal studies of relationships between mothers'

characteristics and mental health and infant temperament (Auerbach et al., 2008; Bridges et al., 1993; Clark et al., 2000; Durbin et al., 2007; Vaughn et al., 1987).

The IBQ distress to limitations was used as the reported measure of irritability from multiple informants to describe the infant irritability at 6 months. Within Study 2, this measure of infant irritability was examined in relation to other temperament dimensions and in relation to observed infant behaviour. The adapted Lab-TAB car seat task was chosen as a measure of observed infant behaviour at 6 months within Study 2 because it had been used with infants at 6 months in a previous study (Bridges et al., 1993), and within the infants' home in a further study (Clark, Kochanska, & Ready, 2000). Whilst there was no significant relationship found between the multi-informant reports of infant irritability and observed infant distress, a sub-sample of the infants using mothers' reports only of infant irritability did show a significant relationship between infant irritability and observed mild distress (whining and whinging). In contrast, strong distress in the car seat was associated with informants' explicit reports of infants' anger and use of force, and not with the IBQ distress to limitations scale (for more details see Hay et al., in press).

Further analyses of additional observational tasks would have strengthened the results of Study 2. Both previous studies (Bridges et al.; Clark et al.) included two further anger eliciting tasks, mild arm restraint of the infant and removal of a toy whilst still holding in view. The CCDS assessments at Wave 2 include further observational tasks that would provide future analyses of the relationship between the infants' reported irritability and observed infant distress.

Understanding the nature of the adult mothers' irritability in relation to the mothers' own experiences of emotional and behavioural problems was an important part of Study 2. The measure used in Study 2 to assess mothers' past behavioural problems (conduct problems) was developed for the purpose of the Cardiff Child Development Study (CCDS) based on the DSM-IV criteria for Conduct Disorder. Within the present study the conduct problems scale was validated by its association with mothers' reports of having been arrested, and is being further examined in other CCDS research papers.

The mother-infant subsystem was the focus for Study 2, but other family sub-systems may also influence the infant's temperament. The analyses of the mother-father subsystem and the father-infant subsystem would have provided more information about the potential mechanisms for intergenerational transmission of irritability. Bronfenbrenner suggested that researchers should not condemn themselves for not taking account of the whole system within their research, but pursue the exploration of the child's ecological system and its influences on the development of the child through systematic examination of parts of the system. There is a paucity of research on infant temperament that examines the influence of other family sub-systems. The important finding within Study 2, namely, that mothers' irritability predicts infant irritability when emotional disorders have been taken into account, provides evidence that should support further research into the potential mechanisms for the intergenerational transmission of irritability within different family subsystems.

Having considered the limitations of the empirical studies within this thesis, the implications of the findings from both studies are now considered in relation to both psychological and psychiatric theories.

5.5 *Implications of the Findings for Psychological and Psychiatric Theories.*

From the two empirical studies within this thesis, irritability was found to play an important role in the relationship between emotional and behavioural problems at different stages of the life span. The previous concern that measurement confounding may account for the relationship between irritability and symptoms of disorders was alleviated from the findings in Study 1. Whilst some items on particular measures of temperament and screening instruments for symptoms of disorder may be confounded, the relationships between irritability and emotional and behavioural symptoms continue when these confounds are removed from the analyses. Specific examination of the role of irritability in relation to ODD firstly indicated that ODD may be a clinical manifestation of irritability, due to the number of potential confounding symptoms with irritability. Such a result would fit well with the spectrum model for the explanation of the relationship between temperament and disorder proposed in previous theoretical accounts of this relationship (Nigg, 2006; Rothbart & Bates, 1989). Further analyses of the parallel roles that irritability and the other symptoms of ODD play in the relationship between internalising and externalising disorders in preschool children, indicated that the ODD pure symptom scale (without irritability items) continued to play a role in the mediation of the relationship between internalising and externalising symptoms, and the composite irritability scale also played a role in the mediation of the same relationship. This finding suggests that whilst ODD includes irritability as an important symptom of disorder, the other ODD symptoms also have a role to play in the relationship between internalising and externalising disorders in preschoolers. The findings from Study 2 would be better represented by the vulnerability model as an explanation of the relationship between

irritability and symptoms of disorder, with irritability being a vulnerability factor (Nigg, 2006).

The stability of irritability found in women across their transition to first-time motherhood supports previous research about the stability of irritability across different stages of the lifespan (Stringaris et al., 2008). This finding would suggest that irritability should not only be of interest to temperament researchers focusing on infancy and early childhood, but should also continue to be of interest to researchers of adult problems. The tendency of researchers of disorder in adults to focus on higher-order factors, such as neuroticism and negative affect may mask important mechanisms for the continued development of disorders in adults and in the potential intergenerational transmission of problems.

The further analyses in Study 2 of the role that irritability plays in relation to emotional and behavioural symptoms in the adult stage of the life span, revealed findings that did not reflect the exact patterns occurring in preschool children. Irritability mediated the relationship between internalising and externalising symptoms in preschool children, but with adults, irritability was related to both emotional and behavioural symptoms, but did not mediate the relationship between the two types of disorder symptoms. These findings could be understood through consideration of previous explanations within developmental psychopathology, that slightly different manifestations of disorders at different age periods probably reflect what is characteristic of those different age periods rather than of different disorders (Rutter & Sroufe, 2000). Simply those high levels of comorbidity within early childhood may through the processes of development become more distinct manifestations of disorder. Irritability, whilst influencing both types of

disorder in early childhood and predicting severity, may continue over the development pathway to interact with other factors that may lead to the development of different types of disorder in adults (an example of multi-finality).

The exploration of infant irritability using the IBQ distress to limitations scale provided a baseline to explore a description of infant irritability within this thesis. The IBQ distress to limitations scale was firmly rooted to the construct of irritability as defined by Shirley (1933) and Thomas and colleagues (1968) in the earliest studies of infant temperament. The 'observe and describe' method used by these early researchers was adopted within this thesis to describe infant irritability without combining it as a higher–order factor (e.g., global negative affectivity). The results from this present study support previous research on infant temperament that infant 'distress to limitations' is related to both fear and activity level dimensions on the IBQ. These combinations of irritability and fear, and irritability and high activity levels, reflect previous temperamental liabilities suggested to predict to particular disorders (Nigg, 2006), such as irritability and activity predicting to conduct disorder. Whilst Study 2 did not examine the specificity from the relationships found between the temperament dimensions and the future presence of disorder, the presence of relationships between particular temperament dimensions could assist further studies to identify specificity to disorders.

Infant 'distress to limitations' was also found to relate to reports of infant anger and temper tantrums, items that are used to measure irritability at later stages of child development (TBAQ; CBQ; Rothbart, 1989; Gartstein & Rothbart, 2003). Within Study 2, the comparison was made between infant irritability as defined using the IBQ 'distress to limitations' scale and reports of infant anger and temper tantrums by examining the

relationship between each of these measures with observed infant distress in response to an everyday challenge, the restraint in a car seat. Whilst there was little distress shown by the infants' overall response to the car seat restraint, a fine distinction was observed in infant distress. Infants' distress to limitations related to observed infant whining and whinging, whilst reported infant anger and temper tantrums related to crying within the car seat task. These results should be viewed with caution, but are useful in directing future research on observed infant temperament to consider fine differentiation, as previously suggested (Rothbart & Bates, 1998).

Finally, the important finding about the prediction of infant irritability from mothers' dispositional irritability when mothers' emotional disorders are taken into account has implications for theories about the intergenerational transmission of specific disorders, such as depression. Many studies have found that mothers' depression predicts subsequent depression in their children, and this has been identified in studies looking at both antenatal and postnatal depression (Hay et al., 2010). The finding in this present study that mothers' irritability measured before and after childbirth predicts infant irritability, that the effect of antenatal irritability is mediated by postnatal irritability, and the relationship found between irritability and emotional disorders at two stages of the lifespan, would suggest that irritability may be the linking factor in the intergenerational pathway to emotional disorders. To illustrate, a cyclical model of irritability in relation to disorder and generation across the lifespan is proposed and shown in Figure 5.1.

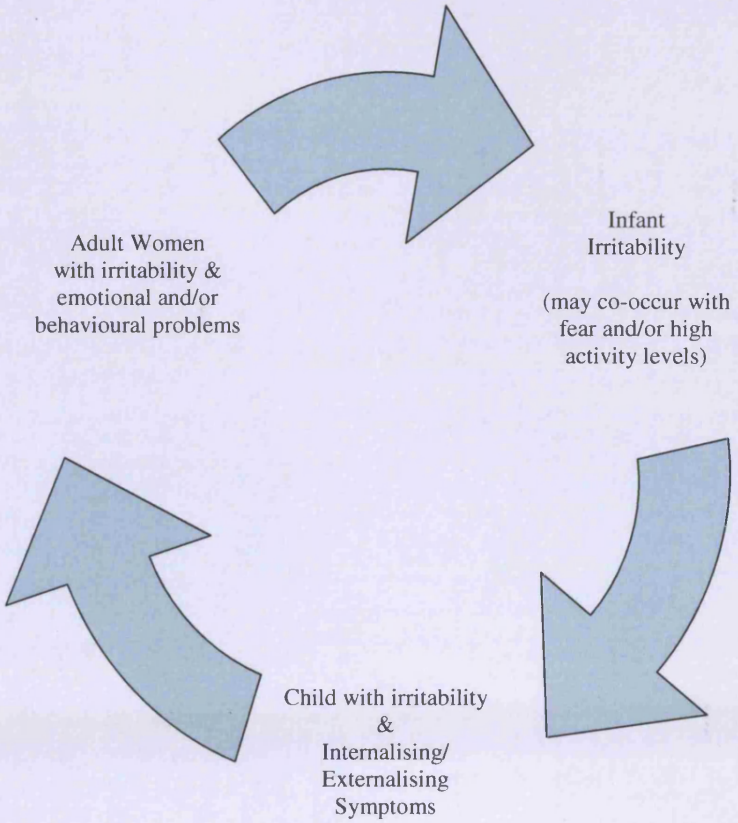


Figure 5.1

The Proposed Cycle of Irritability in Relation to Disorder and Generation across the Lifespan

The model of the cycle of irritability in relation to emotional disorders proposed in Figure 5.1 provides a rich source of ideas for future studies into this pathway from irritability to

disorder within a developmental psychopathology framework. The cycle illustrates the potential influence of one generation's irritability upon the next. The relationship between irritability and symptoms of disorder appear at different stages of the lifespan, e.g. preschoolers and women across childbirth, and has the potential to influence others within the individual's family system. Starting with the focus on the intergenerational transmission of irritability between mother and infant, future studies could explore the processes of transmission further through the consideration of other potential maternal influences and infant influences using a mother-infant interaction measure. The mother-infant subsystem will also be influenced by other family subsystems, such as mother-father and father-infant, future studies could explore these additional subsystems and their potential influence on the infants irritability.

At the adult stage of the irritability cycle, future studies may examine the role of irritability in relation to specific emotional and behavioural disorders with a clinical sample. Whilst irritability is a symptom of specific adult disorders, the measurement of irritability as a characteristic of adults and its relationship with specific disorders would reveal potential mechanisms for the development of disorders in adults.

At the preschool/early childhood stage of the cycle, further studies could focus on the potential mechanisms for the relationship between irritability and disorder. Family studies already focus on the role of conflict within families in the development of both emotional and behavioural disorders in early and middle childhood (Cox & Paley; 1997; Cox, Owen, Lewis, & Henderson, 1989; Harold et al.2007). Further studies could test the role of family conflict as a mediator in the relationship between irritability and disorder in children.

Finally, to understand better the nature of irritability, further exploration of infant distress to limitations in relation to observed infant distress across a variety of different observational challenges, may enable the further differentiation of irritability. In turn, this differentiation may allow the examination of the fine-grained differences in the pathways from temperament to specific disorders.

5.6 *Conclusions*

Within two empirical studies that have focused on different stages of the lifespan at different transition points for the individuals studied, the origin and nature of irritability, and its relationship with symptoms of disorders have been explored within a developmental psychopathology framework. In keeping with Bronfenbrenner's (1977) ecological model of development, the mother-infant subsystem was used as the focus for exploring the potential influence of mothers' characteristics and mental health on the infant's irritability.

The findings from the two empirical studies serve to advance the psychology and psychiatry literature on the theories about the relationship between temperament and disorder. Future studies on this relationship would benefit from scrutinising the measures used to assess particular temperament constructs, such as irritability, to ensure that measurement confounding is accounted for within the analyses. Additionally, future research would extend our understanding of temperament and disorder better if clear definitions were given within the studies about the constructs that are being measured, rather than assuming that higher-order constructs are a good representation of the construct that is under examination. In the pursuit to map the potential pathways from temperament to disorder the research within psychology and psychiatry has taken diverse

paths. Perhaps the drawing together of these two fields of study within the developmental psychopathology framework would be beneficial to the understanding of the processes and mechanisms that occur in the pathway from temperament to disorder, as previously proposed (Frick, 2004).

To conclude, unravelling the complexities across the psychology and psychiatry literature on the origin, nature and role of irritability in relation to emotional and behavioural disorders has focused the outcome of this thesis on the potential to prevent debilitating problems for children. The drive to understand the reasons for very early reports of both emotional and behavioural disorders in children has revealed some insight into potential mechanisms for the development of these problems. There is now sufficient evidence and reliable measures (Egger et al., 2002), to add further support for an epidemiological study into the mental health of children under 5 within the UK. There have been previous calls for screening and treating disorders in babies and infants in the UK, as waiting until adulthood is considered too late (BBC News, 2005). With the potential for 1 in 10 children aged two to five suffering from obvious signs and symptoms of psychiatric illness, such as ADHD, depression or anxiety, we need to consider simple screening methods. Within the UK, there is already a policy to focus on preventive work with families from early stages of infant development. The findings from the empirical studies within this thesis suggest that measuring mothers' irritability in pregnancy through self-report, may be a simple but useful indicator of where to target resources to support early relationships within the family system, to prevent the potential development of long term social and mental health problems.

Appendix 1

 ISBN 0 11 322426 5

Department of Health

Adult Wellbeing

THE SCALE

Name of Child:

_____

Completed by:

_____

Relationship to child:

_____

Date: _____

ADULT WELLBEING SCALE

This form has been designed so that you can show how you have been feeling in the past few days.

Read each item in turn and UNDERLINE the response which shows best how you are feeling or have been feeling in the last few days.

Please complete all of the questionnaire.

1. **I feel cheerful**

   Yes, definitely | Yes, sometimes | No, not much | No, not at all

2. **I can sit down and relax quite easily**

   Yes, definitely | Yes, sometimes | No, not much | No, not at all

3. **My appetite is**

   Very poor | Fairly poor | Quite good | Very good

4. **I lose my temper and shout and snap at others**

   Yes, definitely | Yes, sometimes | No, not much | No, not at all

5. **I can laugh and feel amused**

   Yes, definitely | Yes, sometimes | No, not much | No, not at all

---

6. **I feel I might lose control and hit or hurt someone**

   Sometimes | Occasionally | Rarely | Never

7. **I have an uncomfortable feeling like butterflies in the stomach**

   Yes, definitely | Yes, sometimes | Not very often | Not at all

8. **The though of hurting myself occurs to me**

   Sometimes | Not very often | Hardly ever | Not at all

9. **I'm awake before I need to get up**

   For 2 hours or more | For about 1 hour | For less than 1 hour | Not at all. I sleep until it is time to get up

10. **I feel tense or 'wound up'**

    Yes, definitely | Yes, sometimes | No, not much | No, not at all

---

11. **I feel like harming myself**

    Yes, definitely | Yes, sometimes | No, not much | No, not at all

12. **I've kept up my old interests**

    Yes, most of them | Yes, some of them | No, not many of them | No, none of them

13. **I am patient with other people**

    All the time | Most of the time | Some of the time | Hardly ever

14. **I get scared or panicky for no very good reason**

    Yes, definitely | Yes, sometimes | No, not much | No, not at all

15. **I get angry with myself or call myself names**

    Yes, definitely | Yes, sometimes | Not often | No, not at all

---

16. **People upset me so that I feel like slamming doors or banging about**

    Yes, often | Yes, sometimes | Only occasionally | Not at all

17. **I can go out on my own without feeling anxious**

    Yes, always | Yes, sometimes | No, not often | No, I never can

18. **Lately I have been getting annoyed with myself**

    Very much so | Rather a lot | Not much | Not at all

Scoring

19. The sheet accompanying the questionnaire indicates the method of scoring the 4 subscales.
20. Use of cut-off scores gives indicators of significant care needs with respect to depression, anxiety, and inwardly and outwardly directed irritability.
21. Inward irritability can point to the possibility of self-harm. Outward irritability raises the possibility of angry actions towards the child(ren).
22. As with any screening instrument, interpretation must be in the context of other information. Some respondents will underreport distress, others exaggerate it. A high or low score on any scale does not guarantee that a significant level of need is present.
23. Most value is obtained by using the scale as a springboard for discussion.

Reference

Snaith RP, Constantopoulos AA, Jardine MY & McGuffin P (1978) A clinical scale for the self-assessment of irritability. *British Journal of Psychiatry*. **132**: 163–71.

 ISBN 0 11 322426 5

DH Department of Health

Adult Wellbeing

GUIDANCE ON USING THE SCALE

ADULT WELLBEING SCALE

Background

1. Parent/Caregiver mental health is a fundamental component of assessment.
2. There is evidence that some people respond more openly to a questionnaire than a face to face interview, when reporting on their mental health.
3. A questionnaire gives caregivers the opportunity to express themselves without having to face another person, however sympathetic that person may be.
4. A questionnaire is no substitute for a good relationship, but it can contribute to the development of a rapport if discussed sensitively.
5. During piloting the use of the questionnaire was found to convey the social worker's concern for the parent's wellbeing. This can be particularly valuable where the parent feels their needs are not being considered.

The Scale

6. The scale is the Irritability, Depression, Anxiety (IDA) Scale developed by Snaith *et al* (1978).
7. This scale allows respondents four possible responses to each item.
8. Four aspects of wellbeing are covered: Depression, Anxiety and Inwardly and Outwardly directed Irritability.

Use

9. In principle the questionnaire can be used with any adult, who is in contact with the child whose development and context are being assessed. In practice this will usually be the main caregiver(s).
10. In piloting, social workers reported that use of the scale raised issues on more than half the occasions that it was used. Probable depression was found amongst almost half the caregivers, and significant anxiety in a third.
11. Where social workers were new to the family situation they said they learnt things they did not know. 'It helped me to be aware of the carers' needs', and 'highlighted stresses'. It helped focus on 'parents' needs and feelings'.
12. Even when parents were known to the workers it gave topics an airing and clarified areas to work on; it 'released tension'.
13. Progress can also be registered. It was 'useful to measure when things were calmer'.
14. Used flexibly it can provide openings to discuss many areas including feelings about relationships with partners and children.

Administration

15. It is vital that the respondent understands why they are being asked to complete the scale. Some will be concerned that revealing mental health needs will prejudice their chances of continuing to care for their child. For example, it can be explained that many carers of children experience considerable stress, and it is important to understand this if they are to be given appropriate support.
16. The scale is best filled out by the carer themselves in the presence of the worker, but it can be administered verbally.
17. It takes about 10 minutes to complete.
18. **Discussion is essential**. Usually this will be when the questionnaire has been completed, so the respondent has an opportunity to consider their own needs uninterrupted. However, there will be times when an important clue to how the caregiver feels may be best picked up immediately. One example occurred during piloting, when a respondent expressed distaste for questions about self-harm.

SCORING THE ADULT WELLBEING SCALE

1. **Depression** – Questions 1,3,5,9 and 12 look at depression. The possible response scores that are shown below run from the left to the right – i.e. for question 1 'I feel cheerful', the scores would be looked at from 'yes, definitely' (0), 'yes, sometimes' (1), 'no, not at all' (3), A score of 4–6 is borderline in this scale and a score above this may indicate a problem

QU1	QU3	QU5	QU9	QU12
0,1,2,3	3,2,1,0	0,1,2,3	3,2,1,0,	0,1,2,3,

2. **Anxiety** – Questions 2,7,10,14 and 17 look at anxiety. A score of 6–8 is borderline, above this level may indicate a problem in this area.

QU2	QU7	QU10	QU14	QU17
0,1,2,3	3,2,1,0	3,2,1,0	3,2,1,0,	0,1,2,3,

3. **Outward directed irritability** – Questions 4,6,13 and 16 look at outward directed irritability. A score of 5–7 is borderline for this scale, and a score above this may indicate a problem in this area.

QU4	QU6	QU13	QU16
3,2,1,0	3,2,1,0	0,1,2,3	3,2,1,0,

4. **Inward directed irritability** – Questions 8,11,15 and 18 look at inward directed irritability. A score of 4–6 is borderline, a higher score may indicate a problem.

QU8	QU11	QU15	QU18
3,2,1,0	3,2,1,0	3,2,1,0	3,2,1,0,

Use of cut-off scores gives indicators of significant care needs with respect to depression, anxiety, and inwardly and outwardly directed irritability. Inward irritability can point to the possibility of self-harm. Outward irritability raises the possibility of angry actions towards the child(ren).

As with any screening instrument, interpretation must be in the context of other information. Some respondents will underreport distress, others exaggerate. A high or low score on any scale does not guarantee that significant level of need is present.

Most value is obtained by using the scale as a springboard for discussion.

 ISBN 0 11 322426 5

DH Department of Health

Adult Wellbeing

SCORING THE SCALE

References

Ablow, J.C., Measelle, J.R., Kraemer, H.C., Harrington, R., Luby, J., Smider, N., Dierker, L., Clark, V., Dubika, B., Heffelfinger, A., Essex, M.J., Kupfer, D. J. (1999). The MacArthur three-city outcome study: Evaluating multi-informant measures of young children's symptomatology. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 1580-1590.

Achenbach, T.M., McConaughy, S.H., & Howell, C.T. (1987). Child/adolescent behavioural and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213-232.

Achenbach, T.M. (1991). *Manual for the Child Behaviour Checklist/4-18 and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1997). What is normal? What is abnormal? Developmental perspectives on behavioural and emotional problems. In S.S. Luthar, J.A. Burack, D. Cicchetti, & J.R. Weisz (Eds.), *Developmental Psychopathology: Perspectives on adjustment,risk, and disorder* (pp. 93-114). New York: Cambridge University Press.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders.* (4th ed., DSM-IV). Washington, DC: APA.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders.* (4th ed., DSM-IV-TR). Washington, DC: APA.

American Psychiatric Association. (2004). *Diagnostic and statistical manual of mental disorders.* (6th ed., DSM-IV-TR). Washington, DC: APA.

Angold, A., Costello, E.J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry, 40,* 57-87.

Angold, A., Erkanli, A., Costello, E.J., & Rutter, M. (1996). Precision, reliability and accuracy in the dating of symptom onsets in child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, 37,* 657-664.

Angold, A., & Costello, E.J. (1995). The Child and Adolescent Psychiatric Assessment (CAPA). *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 39-48.

Angold, A., Costello, E., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry, 40,* 55-87.

Angold, A., Erkanli, Costello, & Rutter, M. (1996). Precision, reliability, and accuracy in the dating of symptom onsets in child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, 37,* 657-664.

Angold, A., Prendergast, M., Cox, A., Harrington, R., Simonoff, E., & Rutter, M. (1995). The Child and Adolescent Psychiatric Assessment (CAPA). *Psychological Medicine, 25,* 739-753.

Auerbach, J.G., Berger, A., Atzaba-Poria, N., Arbelle, S., Cypin, N., Friedman, A., & Landau, R. (2008). Temperament at 7, 12, and 25 months in children at familial risk for ADHD. *Infant and Child Development, 17,* 321-338.

Austin, M.P., Hadzi-Pavlovic, D., Leader, L., Saint, K., & Parker, G. (2005). Maternal trait anxiety, depression and life event stress in pregnancy: relationships with infant temperament. *Early Human Development, 81,* 183-190.

Barnett, B.E.W., Hanna, B., Parker, G. (1983). Life event scales for obstetric groups. *Journal of Psychosomatic Research, 27,* 313-320.

Baron, R.M., & Kenny, D.A.(1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182.

Baroni, A., Lunsford, J.R., Luckenbaugh, D.A., Towbin, K.E., & Leibenluft, E. (2009). Practitioner review: The assessment of bipolar disorder in children and adolescents. *Journal of Child Psychology and Psychiatry, 50,* 203-215.

Barrera, M. (1983). A method for the assessment of social support networks in community survey research. *Connections, 3,* 8-13.

Bates, J.E. (1987). Temperament in infancy. In J.D. Osofosky (Ed.), *Handbook of infant development* (2nd ed., pp. 1101-1149). New York: Wiley.

Bates, J.E. (1990). Conceptual and empirical linkages between temperament and behaviour problems: A commentary on the Sanson, Prior, & Kyrios study. *Merrill-Palmer Quarterly, 36,* 193-199.

Bates, J.E., & Bayles, K. (1984). Objective and subjective components in mothers' perceptions of their children from age 6 months to 3 years. *Merrill-Palmer Quarterly, 30,* 111-130.

Bates, J.E., & Bayles, K. (1988). Attachment and the development of behaviour problems. In J. Belsky and T. Nesworski (Eds.). *Clinical implications of attachment* (pp.253-294). Hillsdale, NJ: Erlbaum.

Bates, J.E., Freeland, C.A., & Lounsbury, M.L. (1979). Measurement of infant difficultness. *Child Development, 50,* 794-803.

Bayley, N. (1969). *Manual for the Bayley scales of infant development.* New York: Psychological Corporation.

BBC News (2005). Infants need mental health checks. Retrieved 28th November, 2005 from *http://news.bbc.co.uk/go/pr/fr/-/1/hi/health/4478428.stm*

Beck, C.T. (1999). Maternal depression and child behaviour problems: a meta analysis. *Journal of Advanced Nursing, 29,* 623-629.

Behar, L., & Stringfield, S. (1974). A behaviour rating scale for the preschool child. *Developmental Psychology, 10,* 601-610.

Belsky, J., & Pensky, E. (1988). Marital change across the transition to parenthood. *International Review of Psychiatry, 8,* 65-71.

Belsky, J., Fish, M., & Isabella, R. (1991). Continuity and discontinuity in infant negative and positive emotionality: Family antecedents and attachment consequences. *Developmental Psychology, 27,* 421-431.

Bergman, P., & Escalona, S. (1949). Unusual sensitivities in very young children. *Psychoanalytic Study of the Child, 33,* 3-4.

Bertalanffy, Ludwig von. (1962). *General Systems Theory – A Critical Review,* in: *General Systems Yearbook,* Vol 7, pp. 1-20.

Biederman, J., Faraone, S.V., Wozniak, J., Mick, E., Kwon, A., Cayton, G.A., & Clark, S.V. (2005). Clinical correlates of bipolar disorder in a large, referred sample of children and adolescents. *Journal of Psychiatric Research, 39,* 611-622.

Bradley, R.H. & Corwyn, R.E. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53,* 371-399.

Bridges, L.J., Palmer, S.A., Morales, M., Hurtado, M., Tsai, D. (1993). Agreement between affectively based observational and parent-report measures of temperament at infant age 6 months. *Infant Behavior and Development, 16,* 501-506.

Bronfenbrenner, U. (1972). *Two Worlds of Childhood US & USSR.* New York: Russell Sage Foundation.

Bronfenbrenner, U. (1979). *The Ecology of Human Development.* Cambridge, MA: Harvard University Press.

Brugha, T.S., Bebbington, P., Tennant, C., Hurry, J. (1985). The list of life threatening experiences: a subset of 12 life event categories with considerable contextual threat. *Psychological Medicine, 15,* 194-198.

Burke, J.D., Loeber, R., Lahey, B.B., & Rathouz, P.J. (2005). Developmental transitions among affective and behavioral disorders in adolescent boys. *Journal of Child Psychology and Psychiatry, 46,* 1200-1210.

Buss, A.H., & Plomin, R. (1975). *A Temperament Theory of Personality Development.* New York: Wiley Series in Behaviour.

Buss, A.H., & Plomin, R. (1984). *Temperament: Early developing personality traits.* Hillsdale, NJ: Erlbaum.

Campbell, S. B., Cohn, J. F., Flanagan, C., Popper, S. and et al. (1992). Course and correlates of postpartum depression during the transition to parenthood. *Development & Psychopathology, 4*, 29-47.

Capaldi, D.M., & Rothbart, M.K. (1992). Development and validation of an early adolescent temperament measure. *Journal of Early Adolescence, 12,* 153-173.

Carey, W.B. (1970). A simplified method for measuring infant temperament. *Journal of Pediatrics, 77,* 188-194.

Carey, W.B. (1990). Temperament risk factors in children: A conference report. *Journal of Developmental and Behavioural Paediatrics, 11,* 28-34.

Carey, W.B., & McDevitt, S.C. (1978). Revision of the infant temperament questionnaire. *Pediatrics, 61,* 735-739.

Carlson, G.A. (1983). Bipolar affective disorders in childhood and adolescence. In Cantwell, D.P., Carlson, G.A. (Eds.), *Affective Disorders in Childhood and Adolescence: An Update* (pp. 61-83). New York: Spectrum Publications.

Carlson, G.A., & Kashani, J.H. (1988). Phenomenology of major depression from childhood through adulthood: analysis of three studies. *American Journal of Psychiatry, 145,* 1222-1225.

Carlson GA, Loney J, Salisbury H, Volpe RJ. (1998). Young referred boys with DICA-P manic symptoms vs. two comparison groups. *Journal of Affective Disorder,* 51, 113-121.

Caspi, A., Elder, G., & Bem, D. (1988). Moving away from the world: Life course patterns of shy children. *Developmental Psychopathology, 24,* 824-831.

Cicchetti, D., & Cohen, D.J. (1995). Perspectives on developmental psychopathology. In D.Ciccheti & D.J. Cohen (Eds.), *Developmental psychopathology: Vol. 1. Theory and methods* (pp. 3-20). New York:Wiley.

Clark, L.A., Kochanska, G., & Ready, R. (2000). Mothers' personality and its interaction with child temperament as predictors of parenting behaviour. *Journal of Personality and Social Psychology, 79,* 274-285.

Coffman, S., Levitt, M., Guacci-Franco, N., & Silver, M. (1992). Temperament and interactive effects: Mothers and infants in a teaching situation. *Issues in Comprehensive Paediatric Nursing, 15,* 169-182.

Cohen, J. (1962). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 10,* 37-46.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioural sciences.* Hillsdale, NJ:Erlbaum

Colder, C.R., Mott, J.A., & Berman, A.S. (2002). The interactive effects of infant activity level and fear on growth trajectories of early childhood developmental problems. *Development and Psychopathology, 14,* 1-23.

Collishaw, S. Maughan B., Goodman, R., et al (2004). Time trends in adolescent mental health. *Journal of Child Psychology and Psychiatry, 45,* 1350-1362.

Copeland,W., Shanahan, L., Costello, E.J., & Angold, A. (2009). Configurations of common childhood psychosocial risk factors. *Journal of Child Psychology and Psychiatry, 50,* 451-459.

Cortina, J. M. (1993) What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78,* 98-104.

Costa, P., & McCrae, R. (2001). A theoretical context of adult temperament. In T.D. Wachs & G. Kohnstamm (Eds.), *Temperament in context* (pp. 1-22). Hillsdale, NJ: Erlbaum.

Costello, E.J., Egger, H., & Angold, A. (2005). 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *Journal of the American Academy of Child and Adolescent Psychiatry, 44,* 972-983.

Cox, M.J. (1995). *Developmental changes in relationships: marriage across the transition to parenthood.* Presented at Intimacy Conference Institute of Soc Research University of Michigan.

Cox, A., & Bentovim, A. (2001). The framework for assessment of children in need and their families, *Child Psychology and Psychiatry Review, 6,* 4-10.

Cox, M.J., & Paley, B. (1997). Families as systems. *Annual Review of Psychology, 48,* 243-267.

Cox, M.J., Owen, M.T., Lewis, J.M., & Henderson, V.K. (1989). Marriage, adult adjustment, and early parenting. *Child Development, 60,* 1015-1024.

Crockenberg, S.B. (1986). Are temperamental differences in babies associated with predictable differences in caregiving? In J. V.Lerner & R.M. Lerner (Eds.), *Temperament and social interaction in infants and children* (pp. 53-73). San-Francisco: Jossey-Bass.

Crockenberg, S., & Acredolo, C. (1983). Infant temperament ratings: A function of infants of mothers, or both? *Infant Behavior and Development, 6,* 61-72.

Crockenberg, S., & Smith, P. (2002). Antecedents of mother-infant interaction and infant irritability in the first 3 months of life. *Infant Behavior and Development, 25,* 2-15.

Cummings, E.M., Davies, P.T., & Campbell, S.B. (2000). *Developmental psychopathology and family process: Theory, research and clinical implications.* Guilford : New York.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243,* 1668-1674.

Demetriou, H., & Hay, D. (2004). Toddlers' reactions to the distress of familiar peers: the importance of context. *Infancy*, 6, 299-318.

Department of Health, (DoH, 2000). *Assessing children in need and their families: Practice Guidance*. London: HMSO.

Department for Work and Pensions, (DWP, 2005). Retrieved September 27, 2007 from *http://www.dwp.gov.uk/mediacentre/pressreleases/2005/mar/nov04_c%2Bf.pdf*

Dunn, J., Slomkowski, C., Beardsall, L., & Rende, R. (1994). Adjustment in middle childhood and early adolescence: Links with earlier and contemporary sibling relationships. *Journal of Child Psychology and Psychiatry, 35,* 491-504.

Dunn, J., Deater-Deckard, K., Pickering, K., O'Connor, T., Golding, J, & the ALSPAC study team. (1998). Children's adjustment and prosocial behaviour in step-, single-parent, and non-stepfamily settings: findings from a community study. *Journal of Child Psychology and Psychiatry, 39,* 1083-1095.

Durbin, C.E., Hayden, E.P., Klein, D.N., & Olino, T.M. (2007). Stability of laboratory-assessed temperamental emotionality traits from ages 3 to 7. *Emotion, 2,* 388-399.

Earls, F., & Jung, K.G. (1987). Temperament and Home Environment Characteristics as Causal Factors in the Early Development of Childhood Psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 491-498.

Edhborg, M., Seimyr,L., Lundh, W., & Widstrom, A.M. (2000). Fussy child-difficult parenthood? Comparisons between families with a 'depressed' mother and non-depressed mother 2 months postpartum. *Journal of Reproductive and Infant Psychology, 18,* 225-238.

Egger, H.L., Ascher, B.H., & Angold, A. (2002). *The preschool age psychiatric assessment (PAPA) parent interview –version 1.3.* Centre for developmental epidemiology, department of psychiatry and behavioural sciences, Duke University Medical Centre.

Egger, H.L., & Angold, A. (2006). Common emotional and behavioural disorders in preschool children: presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry, 47,* 313-337.

Egger, H.L., Erkanli, A., Keeler, G., Potts, E., Walter, B.K., & Angold, A. (2006). Test-retest reliability of the preschool age psychiatric assessment (PAPA). *Journal of the American Academy of Child and Adolescent Psychiatry, 45,* 538-549.

Elias, P., McKnight, A., & Kinshott, G. (1999). SOC 2000: Redefining skill revision of the standard occupational classification. *Skill Task Force Research Paper 19.*

Else-Quest, N.M., Shibley Hyde, J., Goldsmith, H.H., Van Hulle, C.A. (2006). Gender differences in temperament:A meta-analysis. *Psychological Bulletin, 132,* 33-72.

Escalona, S.K. (1968). *The Roots of Individuality.* UK: Tavistock Publications

Fergus, E.L., Miller, R.B., Luckenbaugh, D.A., Leverich, G.S., Findling, R.L., Speer, A.M., & Post, R.M. (2003). Is there progression from irritability/dyscontrol to major depressive and manic symptoms? A retrospective community survey of parents of bipolar children. *Journal of Affective Disorders, 77,* 71-78.

Field, A. (2005). *Discovering statistics using SPSS.* (2nd ed.) London: Sage Publications.

Fish, M., & Crockenberg, S. (1981). Correlates and antecedents of nine-month infant behaviour and mother-infant interaction. *Infant Behaviour and Development, 4,* 69-81.

Flouri, E. (2008). Temperament influences on parenting and child psychopathology: Socio-economic disadvantage as moderator. *Child Psychiatry Human Development*

Forbes, E.E., Cohn, J.F., Allen, N.B., Lewinsohn, P.M. (2004). Infant affect during parent-infant interaction at 3 and 6 months: Differences between mothers and fathers and influence of parent history of depression. *Infancy, 5,* 61-84.

Fox, N.A., Henderson, H.A., Rubin, K.H., Calkins, S.D., & Schmidt, L.A. (2001). Continuity and discontinuity of behavioural inhibition and exuberance: psychopathological and behavioural influences across the first four years of life. *Child Development, 72, (1),* 1-21.

Frick, P.J. (2004). Integrating research on temperament and childhood psychopathology:Its pitfalls and promise. *Journal of Clinical Child and Adolescent Psychology, 33,* 2-7.

Gartstein, M.A., & Rothbart, M.K. (2003). Studying infant temperament via the revised infant behaviour questionnaire. *Infant Behaviour and Development, 26,* 64-86.

Gilliam, W.S. (2005). Prekindergarteners left behind: expulsion rates in state prekindergarten systems. Retrieved September 27, 2007 from *http://www.hartfordinfo.org/issues/wsd/education/NationalPreKExpulsionPaper.pdf*

Gilliom, M., & Shaw, D.S. (2004). Codevelopment of externalising and internalising problems in early childhood. *Development and Psychopathology, 16,* 313-333.

Gjone, H., & Stevenson, J. (1997). A longitudinal twin study of temperament and behaviour problems: common genetic or environmental influence. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 1449-1456.

Goldsmith, H.H. (1996). Studying temperament via construction of the Toddler Behaviour Assessment Questionnaire. *Child Development, 67,* 218-235.

Goldsmith, H.H., & Rieser-Danner, L. (1990). Assessing early temperament. In C. Reynolds & R. Camphaus (Eds.), *Handbook of psychological and educational assessment of children:Vol.2, Personality behaviour and context* (pp. 345-378).

Goldsmith, H.H., & Rothbart, M.K. (1994). *The Laboratory Assessment Battery, Prelocomotor 2.03.* Department of Psychology, University of Wisconsin.

Goldsmith, H.H., Lemery, K.S., & Essex, M.J. (2004). Roles for temperament in the liability to psychopathology in childhood. In L.F.DiLalla (Ed.). *Behaviour genetics principles: Perspectives in development, personality, and psychopathology.* (pp. 19-39). Washington, DC: American Psychological Association.

Goldsmith, H.H., Buss, A.H., Plomin, R., Rothbart, M.K., Thomas, A., Chess, S., Hinde, R.A., & McCall, R.B. (1987). Roundtable: what is temperament? four approaches. *Child Development, 58,* 505-529.

Goldsmith, H. H., Lemery, K. S., & Essex, M. J. (2004). Temperament as a liability factor for childhood behavioral disorders: The concept of liability. In L. DiLalla (Ed.). *Behavior genetics principles: Perspectives in development, personality, and psychopathology.* (pp. 19-39). Washington, DC: American Psychological Association

Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38,* 581-586.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 1337-1345.

Goodman, R., & Scott, S. (1999). Comparing the strengths and difficulties questionnaire and the child behaviour checklist: is small beautiful? *Journal of Abnormal Psychology, 27,* 17-24.

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the strengths and difficulties questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *British Journal of Psychiatry, 177,* 534-539.

Green, H., McGinnity, A., Ford, T., & Goodman, R. (2005). *Mental Health of Children and Young People in Great Britain, 2004.* London:HMSO

Guardian, 17th March 2008. Violence on the increase in schools teachers warn. Matthew Taylor. http://www.guardian.co.uk/education/2008/mar/17/schools.uk

Guardian, 22nd Dec 2008. Violence levels in schools very worrying: Tories warn. Anthea Lipsett. http://www.guardian.co.uk/education/2008/dec/22/schools-violence-tories

Hane, A.A., Fox, N.A., Polak-Toste, C., Ghera, M.M., & Guner, B.M. (2006). Contextual basis of maternal perceptions of infant temperament. *Developmental Psychology, 42,* 1077-1088.

Harold, G., Aitken, J., & Shelton, K. H. (2007). Interparental conflict and children's academic attainment: A longitudinal analysis. Journal of Child Psychology and Psychiatry, *48,* 1223-1232.

Harrington, R.C., Fudge, H., Rutter, M., Pickles, A., & Hill, J. (1990). Adult outcomes of childhood and adolescent depression: psychiatric status. *Archives of General Psychiatry, 47,* 465-473.

Hay, D.F., Hudson, K., & Liang, W. (in press). Links between Preschool Children's Prosocial Skills and Aggressive Conduct Problems: The Contribution of ADHD Symptoms.

Hay, D.F., Pawlby, S., Waters, C.S., & Sharp, D. (2008). Antepartum and postpartum exposure to maternal depression: different effects on different adolescent outcomes. *Journal of Child Psychology and Psychiatry,49,* 1079-1088.

Hay, D.F., Pawlby, S., Waters, C.S., Perra, O., & Sharp, D. (2010). Mothers' antenatal depression and their children's antisocial outcomes. *Child Development, 81,* 149-165.

Henderson, N.D. (1982). Human behaviour genetics. *Annual Review of Psychology, 33,* 403-440.

Hirshfield-Becker, D.R., Biederman, J., Calltharp, S., Rosenbaum, E.D., Faraone, S.V., & Rosenbaum, J.F. (2003). Behavioural inhibition and disinhibition as hypothesised precursors to psychopathology: implications for paediatric bipolar disorder. *Biological Psychiatry, 53,* 985-999.

Home Office (2008). Young people and crime: Findings from the 2006 Offending, Crime and Justice Survey. http://www.homeoffice.gov.uk/rds/pdfs08/hosb0908.pdf

Hudziak, J.J. et al., (2004). Genetic and environmental contributions to the child behaviour checklist: obsessive-compulsive scale: A cross-cultural twin study. *Archives of General Psychiatry, 61,* 608-616.

Hubert, N.C., Wachs, T.D., Peters-Martin, P., & Gandour, M.J. (1982). The study of early temperament: Measurement and conceptual issues. *Child Development, 53,* 571-600.

Isabella, R.A., Ward, M.J., & Belsky, J. (1985). Convergence of multiple sources of information on infant individuality: Neonatal behavior, infant behavior, and temperament reports. *Infant behavior and Development, 8,* 283-291.

Jackson, D.N. (1967) *Personality research form manual.* New York: Research Psychologist Press.

Janson, H. & Mathiesen, K.S. (2008). Temperament profiles from infancy to middle childhood: Development and associations with behavior problems. *Developmental Psychology, 44,* 1314-1328.

Jenkins, J.M, Rashbash, J., & O'Connor, T.G. (2003). The role of shared family context in differential parenting. *Developmental Psychology, 39,* 99-113.

Kagan, J. (1994). *Galen's prophecy: Temperament in human nature.* New York: West View Press.

Kagan, J. (1998). Biology and the child. In W.S.E. Damon and N.V.E. Eisenberg (Eds.), *Handbook of child psychology: social, emotional and personality development* (5th edn.), vol. 3. New York: Wiley, 177-235.

Kagan, J., & Snidman, N. (1991). Infant predictors of inhibited and uninhibited profiles. *Psychological Science, 2,* 40-44.

Kagan, J., Snidman, N., Zentner, M., & Peterson, E. (1999). Infant temperament and anxious symptoms in school age children. *Development and Psychopathology, 11,* 209-224.

Kashani, J.H., Holcomb, W.R., & Orvaschel, H. (1986). Depression and depressive symptoms in preschool children from the general population. *American Journal of Psychiatry, 143,* 1138-1143.

Kashani, J.H., Allan, W.D., Beck, Jr., N.C., Bledsoe, Y., & Reid, J.C. (1997). Dysthymic disorder in clinically referred preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 1426 - 1423.

Kiernan, K.E. (1997). Becoming a young parent: A longitudinal study of associated factors. *British Journal of Sociology, 48,* 406-428.

Kline, P. (1999). *The handbook of psychological testing* (2nd edition). London:Routledge.

Kochanska, G. (1997). Multiple pathways to conscience for children with different temperaments: From toddlerhood to age 5. *Developmental Psychology, 33,* 135-146.

Kochanska, G. (1998). Mother-child relationship, child fearfulness, and emerging attachment: A short-term longitudinal study. *Developmental Psychology, 34,* 480-490.

Kolko, D.J., & Kazdin, A.E. (1993). Emotional/behavioural problems in clinic and non-clinic children: correspondence among child, parent and teachers reports. *Journal of Child Psychology and Psychiatry, 34,* 991-1006.

Komsi, N., Raikkonen, K., Pesonen, A.K., Heinonen, K., Keskivaara, A.L., & Strandberg, T.E. (2006). Continuity of temperament from infancy to middle childhood. *Infant Behavior & Development, 29,* 494-508.

Kovacs, M. (1981)). Rating scales to assess depression in school aged children . *Acta Paedopsychiatry, 46,* 305-315.

Lahey, B. (2004). Commentary: Role of temperament in developmental models of psychopathology. *Journal of Clinical Child and Adolescent Psychology, 33,* 88-93.

Leibenluft, E., James, R., Blair, R., Charney, D.S., & Pine, D.S. (2003). Irritability in pediatric mania and other childhood psychopathology. *Annals of the New York Academy of Sciences, 1008,* 201-218.

Lemery, K.S., Essex, M.J., & Smider, N.A. (2002). Revealing the relation between temperament and behaviour problem symptoms by eliminating measurement confounding expert ratings and factor analyses. *Child Development, 73,* 867-882.

Lemery, K.S., Goldsmith, H.H., Klinnert, M.D., & Mrazek, D.A. (1999). Developmental models of infant and childhood temperament. *Developmental Psychology, 35,* 189-204.

Lemola, S., Stadmayr, W., & Grob, A. (2009). Infant irritability: The impact of fetal alcohol, maternal depressive symptoms, and low emotional support from the husband. *Infant Mental Health Journal, 30,* 57-81.

Lengua, L.J. (2003). Associations among emotionality, self-regulation, adjustment problems, and positive adjustment in middle childhood. *Applied Developmental Psychology, 24,* 595-618.

Lengua, L.J., West, S.G., & Sandler, I.N. (1998). Temperament as a predictor of symptomatology in children: Addressing contamination of measures. *Child Development, 69,* 164-181.

Levy, F., Hay, D., McStephen, M., Wood, C., & Waldman, I. (1997). Attention deficit hyperactivity disorder: A category or a continuum? Genetic analysis of a large scale twin study. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 734-744.

Locke HJ; Wallace KM. Bonjean CM, Hill RJ, & McLemore SD (1967). *Sociological measurement: An inventory of scales and indices.* San Francisco: Chandler Publ. 211-214.

Loeber, R., & Hay, D.F. (1997). Key issues in the development of aggression and violence from childhood to early adulthood. *Annual Review of Psychology, 48,* 371-410.

Loeber, R., Burke, J., & Pardini, D.A. (2009). Perspectives on oppositional defiant disorder, conduct disorder, and psychopathic features. *Journal of Child Psychology and Psychiatry, 50,* 133-142.

Loeber, R., Wung, P., Keenan, K., Giroux, B., Stouthamer-Loeber, M., Van Kammen, W.B., & Maughan, B. (1993). Developmental pathways in disruptive child behaviour. *Development and Psychopathology, 5,* 103-133.

Loranger, A. W., Susman, V. L., Oldham, J. M., et al (1987) The Personality Disorder Examination: a preliminary report. *Journal of Personality Disorders, 1,* 1-13.

Luby, J.L., Heffelfinger, A.K., Mrakotsky, C., Hessler, M.J., Brown, K.M., & Hildebrand, T. (2002). Preschool major depressive disorder (MDD): preliminary validation for developmentally modified DSM-IV criteria . *Journal of the American Academy of Child and Adolescent Psychiatry, 41,* 928-937.

Luby, J.L., Heffelfinger, A.K., Mrakotsky, C., Brown, K.M., Hessler, M.J., Wallis, J.M., & Spitznagel, E.L. (2003a). The clinical picture of depression in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 42,* 340-348.

Luby, J.L., Heffelfinger, A.K., Mrakotsky, C., Brown, K.M., Hessler, M.J., & Spitznagel, E.L. (2003b). Alterations in stress cortisol reactivity in depressed preschoolers relative to

psychiatric and no-disorder comparison groups. *Archives of General Psychiatry, 60,* 1248-1255.

Mangelsdorf, S., Gunnar, M., Kestenbaum, R., Lang, S., & Andreas, D. (1990). Infant proneness-to-distress temperament, maternal personality, and mother-infant attachment: Associations and goodness of fit. *Child Development, 61,* 820-831.

Mangold (2007). *Interact Software Programme.* Germany: Mangold International.

Maughan, B., Collishaw, S., Meltzer, H., Goodman R. (2008) Recent trends in UK child and adolescent mental health (2008) *Social Psychiatry and Psychiatric Epidemiology, 43,* 305-310.

Maughan, B., Rowe, R., Messer, J., Goodman, R., Meltzer, H. (2004). Conduct Disorder and Oppositional Defiant Disorder in a national sample: developmental epidemiology. *Journal of Child Psychology and Psychiatry, 45,* 609-621.

Mayer, J.D. (1993). A system-topics framework for the study of personality. *Imagination, Cognition and Personality, 13,* 99-123.

McCall, R. B. (1986). Issues of stability and continuity in temperament research. In R. Plomin & J. Dunn (Eds.). *The study of temeperament: Changes, continuities and challenges.* Hillsdale, NJ: Erlbaum.

McDevitt, S.C., & Carey, W.B. (1978). Measurement of temperament in 3- to 7-year-old children. *Journal of Child Psychology and Psychiatry, 19,* 245-253.

Mellor, D. (2004). Furthering the use of the strengths and difficulties questionnaire: reliability with younger child respondents. *Psychological Assessment, 16,* 396-401.

Meltzer, H., Gatward, R., Goodman, R., & Ford, T. (2000). *Mental Health of Children and Adolescents in Great Britain.* London: The Stationery Office.

Mitchell, J., McCauley, E., Burke, P.M., & Moss, S.J. (1988). Phenomenology of depression in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 12-20.

Moffitt, T.E., Caspi, A., Harrington, H., Milne, B.J., Melchior, M., Goldberg, D., et al. (2007). General anxiety disorder and depression:Childhood risk factors in a birth cohort followed to age 32. *Psychological Medicine, 37,* 441-452.

Nigg, J.T. (2006). Temperament and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 47,* 395-422.

Nock, M.K., Kazdin, A.E., Hiripi, E., & Kessler, R.C. (2007). Lifetime prevalence, correlates, and persistence of oppositional defiant disorder: results from the national

comorbidity survey replication. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 48,* 703-713.

Oberklaid, F., Prior, M., Sanson, A., Sewell, J., & Kyrios, M. (1990). Assessment of temperament in the toddler age group. *Pediatrics, 85,* 559-566.

Office of National Statistics (ONS; 2007). Retrieved September 27, 2007 from *http://www.statistics.gov.uk/CCI/nugget.asp?ID=1748&Pos=2&ColRank=1&Rank=176*

Oldehinkel, A.J., Hartman, C.A., De Winter, A.F., Veenstra, R., & Ormel, J. (2004). Temperament profiles associated with internalising and externalising problems in preadolescence. *Development and Psychopathology, 16,* 421-440.

Olson, S.L., Bates, J.E., & Bayles, K. (1989). Predicting long term developmental outcomes from maternal perceptions of infant and toddler behaviour. *Infant Behaviour and Development, 12,* 77-92.

Orvaschel, H., & Puig-Antich, J. (1987). *Schedule for Affective Disorders and Schizophrenia for School-Age Children: Epidemiologic 4th Version.* FT Lauderdale FL: Nova University, Center for Psychological Study.

Patterson, G.R. (1982). *Coercive family process.* Eugene, OR: Castalia.

Pauli-Pott, U., Mertesacker, B., Bade, U., Bauer, C., & Beckmann, D. (2000). Context of relations of infant negative emotionality to caregiver's reactivity/sensitivity. *Infant Behavior and Development, 23,* 23-29.

Pauli-Pott, U., Mertesacker, B., Bade, U., Haverlock, A., & Beckermann, D. (2003) Parental perceptions and infant temperament development. *Infant Behavior & Development, 26,* 27-48.

Pauli-Pott, U., Mertesacker, B., & Beckmann, D. (2004). Predicting the development of infant emotionality from maternal characteristics. *Development and Psychopathology, 16,* 19-42.

Pawlby, S., Hay, D.F., O'Keane, V.,Waters, C.S., & Sharp, D. (2009). Antenatal depression predicts depression in adolescent offspring: prospective longitudinal community-based study. *Journal of Affective Disorders*, 107, S54-S55.

Pedlow, R., Sanson, A., Prior, M., & Oberklaid, F. (1993). Stability of maternally reported temperament from infancy to 8 years. *Developmental Psychology, 29,* 998-1007.

Plomin, R., DeFries, J.C., & Fulker, D.W.(1988). *Nature and nurture during infancy and early childhood.* Cambridge: Cambridge University Press.

Plomin, R. (2006). Genetics and children's experiences in the family. *Journal of Child Psychology and Psychiatry, 36,* 33-68.

Prior, M., Sanson, A., & Oberklaid, F. (1989). The Australian temperament project. In G. Kohnstamm, J. Bates, & M. Rothbart (Eds.), *Temperament in childhood* (pp. 537-556). New York: Wiley.

Putnam, S.P., & Stifter, C.A. (2008). Reactivity and regulation: The impact of Mary Rothbart on the study of temperament. *Infant Child and Development, 17,* 311-320.

Putnam, S.P., Rothbart, M.K., & Gartstein, M.A. (2008). Homotypic and heterotypic fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant and Child Development, 17,* 387-405.

Richman, N., & Graham, P.J. (1971). A behavioural screening questionnaire for use with 3-year-old children: preliminary findings. *Journal of Child Psychology and Psychiatry, 12,* 5-33.

Richters, J. & Pellegrini, D. (1989). Depressed mothers' judgements about their children: An examination of the depression-distortion hypothesis. *Child Development, 60,* 1068-1075.

Riese, M.L. (1987). Longitudinal assessment of temperament from birth to 2 years: A comparison of full-term and preterm infants. *Infant Behavior and Development, 10,* 347-363.

Robinson, J., Kagan, J., Resnick, J., & Corley, R. (1992). The heritability of inhibited and uninhibited behaviour. *Developmental Psychology, 28,* 1030-1037.

Romano, E., Zoccolillo, M., & Paquette, D. (2006). Histories of child maltreatment and psychiatric disorder in pregnant adolescents. *Journal of the American Academy of Chidl and Adolescent Psychiatry, 45,* 329-336.

Rothbart, M.K. (1981). Measurement of temperament in infancy. *Child Development, 52,* 569-578.

Rothbart, M.K., & Bates, J.E. (1998). Temperament. In W.Damon., & N. Eisenberg. (Eds.), *Handbook of child psychology: volume 3. Social, emotional and, personality development.* New York: Wiley.

Rothbart, M.K., & Derryberry, D. (1981). Development of individual differences in temperament. In M.E. Lamb & A.L. Brown (Eds.), *Advances in developmental psychology* (Vol. 1, pp.37-86). Hillsdale, NJ:Erlbaum.

Rothbart, M.K., & Hwang, J. (2002). Measuring infant temperament. *Infant Behavior & Development, 25, pp 113-116.*

Rothbart, M.K., & Goldsmith, H.H. (1985). Three approaches to the study of infant temperament. *Developmental Review, 5,* 237-260.

Rothbart, M.K., & Putnam, S.P. (2002). Temperament and socialisation. In L.Pulkinnen & A.Caspi (Eds.), *Paths to successful development: personality in the life course.* (pp.19-45). Cambridge UK: Cambridge University Press.

Rothbart, M.K., Ahadi, S.A., & Evans, D.E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology, 78,* 122-135.

Rothbart, M.K., Ahadi, S.A., & Hershey, K.L. (1994). Temperament and social behaviour in childhood. *Merrill-Palmer Quarterly, 40,* 21-39.

Rothbart, M.K., Chew, K.H., & Gartstein, M.A. (2001). Assessment of temperament in early development. In L.T.Singer & P.S. Zeskind (Eds). *Biobehavioural assessment of the infant.* (pp.190-208). New York: Guilford Press.

Rothbart, M.K., Ahadi, S.A., Hershey, K.L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development, 72,* 1394-1408.

Rushton, J.P., Chrisjohn, R.D, & Fekken,G.C. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences, 2,* 293-302.

Rutter, M. (2000). Psychosocial influences: critiques, findings and research needs. *Development and Psychopathology, 12,* 375-405.

Rutter, M. (2009). Understanding and testing risk mechanisms for mental disorders. *Journal of Child Psychology and Psychiatry, 50,* 44-52.

Rutter, M., & Sroufe, L.A. (2000). Developmental psychopathology: Concepts and challenges. *Development and Psychopathology, 12,* 265-296.

Rutter, M., Moffitt, T.E., & Caspi, A. (2006a). Gene-environment interplay and psychopathology: multiple varieties but real effects. *Journal of Child Psychology and Psychiatry, 47,* 226-261.

Rutter, M., Kim-Cohen, J, & Maughan, B, (2006b). Continuities and discontinuities in psychopathology between childhood and adult life. *Journal of Child Psychology and Psychiatry, 47,* 276-295.

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour.* London: Longmans.

Rutter, M., Tizard, J., Yule, W., Graham, P., & Whitmore, K. (1976). Research report:Isle of Wight studies, 1964-1974. *Psychological Medicine, 6,* 313-332.

Ryan, R.D., Puig-Antich, J., Ambrosini, P., et al. (1987). The clinical picture of major depression in children and adolescents. *Archives of General Psychiatry, 44,* 854-861.

Sameroff, A. J. (1995). General systems theories and developmental psychopathology. In D. Cicchetti & D. Cohen (Eds.), *Developmental Psychopathology,* Vol. 1 (pp. 659–695). New York: Wiley.

Sanson, A., & Rothbart, M.K. (1995). Child temperament and parenting. In Marc. H. Bornstein (Ed) *Handbook of parenting volume 4: Applied and practical parenting.* (pp 299-321). Lawrence Erlbaum Associates: Hove, UK.

Sanson, A., Prior, M., & Kyrios, M. (1990). Contamination of measures in temperament research. *Merrill-Palmer Quarterly, 36,* 179-192.

Sanson, A., Prior, M., & Oberklaid, F. (1984). Temperament in Australian toddlers. In: *Paper presented at the Third National Child Development Conference.*

Sanson, A., Prior, M., & Oberklaid, F. (1985). Normative data on temperament in Australian infants. *Australian Journal of Psychology, 37,* 185-195.

Sanson, A., Oberklaid, F., Pedlow, R., & Prior, M. (1991). Risk indicators: Assessment of infancy predictors of pre-school behavioural maladjustment. *Journal of Child Psychology and Psychiatry, 32,* 609-626.

Sanson, A., Prior, M., Oberklaid, F., Garino, E. & Sewell, J. (1987). The structure of infant temperament: factor analysis of the revised infant temperament questionnaire. *Infant Behaviour and Development, 10,* 97-104.

Saudino, K., Cherny, S., & Plomin, R.(2000). Parent ratings of temperament in twins: Explaining the too low dz correlations. *Twin Research, 3,* 224-233.

Shaffer, D., Fisher, P., Dulcan, M.K., & Davies, M. (1996). The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): Description, acceptability, prevalence rates, and performance in the MECA study. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 865 -877.

Shaffer, D., Lucas, C.P., & Richters, J.E. (1999). *Diagnostic Assessment in Child and Adolescent Psychopathology.* New York: The Guildford Press.

Shaw, D.S., Owens, E.B., Giovanelli, J., & Winslow, E.B. (2001). Infant and toddler pathways leading to externalising disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 36-43.

Shiner, R., & Caspi, A. (2003). Personality difference in childhood and adolescence: Measurement development and consequences. *Journal of Child Psychology and Psychiatry, 44,* 2-32.

Shirley, M.M. (1933). *The First Two Years: A Study of Twenty-Five Babies.* Minneapolis: University of Minnesota Press.

Smart, D., & Sanson, A. (2005). A comparison of children's temperament and adjustment across 20 years. *Family Matters, 72,* 50-57.

Snaith, R.P, & Taylor, C.M. (1985). Irritability Definition, assessment and associated factors. *British Journal of Psychiatry, 147,* 127-136.

Snaith, R.P., Constantopoulos, A.A., Jardine, M.Y., & McGuffin, P. (1978). A clinical scale for the self-assessment of irritability. *British Journal of Psychiatry, 132,* 164-171.

Speltz, M.L., McClellan, J., DeKlyen, M., & Jones, K. (1999). Preschool boys with oppositional defiant disorder: clinical presentation and diagnostic change. *Journal of the American Academy of Child and Adolescent Psychiatry,38,_* 838- 846.

Sroufe, L.A. (1997). Psychopathology as an outcome of development. *Development and Psychopathology, 9,* 251-268.

Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers, and children on internalizing and externalizing behaviour problems. *Journal of Clinical Child Psychology, 22,* 107-115.

Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245-251.

Stifter, C.A., Willoughby,M.T., & Towe-Goodman, N. et al., (2008). Agree or agree to disagree? Assessing the convergence between parents and observers on infant temperament. *Infant and Child Development, 17,* 407-426.

Stringaris, A. & Goodman, R. (2009). Three dimensions of oppositionality in youth. *Journal of Child Psychology and Psychiatry, 50,* 216-223.

Stringaris, A., Cohen, P., Pine, D.S., & Leibenluft, E. (2009). Adult outcomes of youth irritability: A 20 year prospective community-based study. *American Journal of Psychiatry, 166,* 1048-1054.

Tackett, J.L., & Krueger, R.F. (2005). Interpreting personality as a vulnerability for psychopathology: A developmental approach to the personality-psychopathology relationship. In B.L. Hankin & J.R.Z. Abela (Eds.), *Development of psychopathology: A vulnerability-stress perspective (pp. 199-214).* Thousand Oaks, CA: Sage.

Task Force on Research Diagnostic Criteria: Infancy and Preschool. (2003). Research diagnostic criteria for infants and preschool children: The process and empirical support. *Journal of American Academy of Child and Adolescent Psychiatry, 42,* 1504-1512.

Thomas, A. & Chess, S. (1977). *Temperament and Development.* New York: Brunner/Mazel.

Thomas, A. & Chess, S. (1984). Genesis and evolution of behavioral disorders: From infancy to early adult life. *The American Journal of Psychiatry, 141,* 1-9.

Thomas, A. & Chess, S. (1986). The New York Longitudinal Study: From infancy to early adult life. In R. Plomin & J. Dunn (Eds.), *The Study of Temperament: Changes, Continuities & Challenges* (pp. 39-52). London: Lawrence Erlbaum Associates.

Thomas, A., Chess, S., & Birch, H.G. (1968). *Temperament and behavior disorders in children.* London: University of London Press Ltd.

Thomas, A., Chess, S., Birch, H.G., Hertzig, M.E., & Korn, S. (1963). *Behavioural Individuality in Early Childhood.* New York: New York University Press.

Vaughan, B.E., Deinard, A., & Egeland, B. (1980). Measuring temperament in paediatric practice. *Journal of Pediatrics, 96,* 510-518.

Vaughan, B.E., Taraldson, B.J., Crichton, L., & Egeland, B. (1981). The assessment of infant temperament: A critique of the Carey infant temperament questionnaire. *Infant Behaviour and Development, 4,* 1-17.

Vaughn, B.E., Bradley, C.F., Joffe, L.S., Seifer, R., & Barglow, P. (1987). Maternal characteristics measured prenatally are predictive of ratings of temperamental 'difficulty' on the Carey infant temperament questionnaire. *Developmental Psychology, 1,* 152-161.

Vaughn, B.E., Taraldson, B.J., Cuchton, L., & Byron, Egeland. (2002). The assessment of infant temperament: A critique of the Carey infant temperament questionnaire. *Infant Behavior and Development, 25,* 98-112.

Wachs, T.D., & Bates, J.E. (2004). Temperament. In G. Bremner & A.Fogel (Eds.), *Blackwell Handbook of Infant Development* (pp. 465-499). Oxford: Blackwell Publishing Ltd.

Waters, C.S. (2008). *Young motherhood, maternal psychopathology and children's cognitive, behavioural and emotional development.* Unpublished doctoral dissertation. University of Cardiff, UK.

Watson, D., Gamez, W, Simms, L.J. (2006). Basic dimensions of temperament and their relations to anxiety and depression : A symptom based perspective. *Journal of Research in Personality, 39,* 46-66.

Williams, E.J. (1959). The comparison of regression variables. *Journal of the Royal Statistical Society (Series, B), 21,* 396-399.

Wing, J.K., Babor, T., Brugha, T., Burke, J., Cooper, J.E., Giel, R., Jablenski, A., Regier, D., & Sartorius, N. (1990). SCAN: Schedules for Clinical Assessment in Neuropsychiatry. *Archives of General Psychiatry, 47,* 589-593.

World Health Organisation. (1978). *Mental disorders: Glossary and guide to their classification in accordance with the ninth revision of the international classification of diseases.* World Health Organisation: Geneva.

World Health Organisation Division of Mental Health. (1992). *WHO SCAN Advisory Committee. Scan schedules for clinical assessment in neuropsychiatry, version 1.0.* World Health Organisation: Geneva.

Zero to Three. (1994). *Diagnostic Classification: 0-3: Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood.* Arlington, VA: National centre for clinical infant programs.

Zoccolillo, M., Pickles, A., Quinton, D., & Rutter, M. (1992). The outcome of childhood conduct disorder: Implications for defining adult personality disorder and conduct disorder. *Psychological Medicine, 22,* 971-986.