A Quantitative Genetic Analysis of Caregiver-reported and Observed Fear, Anger, and Sadness in Middle Childhood

by

Sierra Clifford

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Approved May 2013 by the
Graduate Supervisory Committee:

Kathryn Lemery, Chair
Michelle Shiota
Nancy Eisenberg

ARIZONA STATE UNIVERSITY

December 2013

ABSTRACT

The purpose of the current study was to use structural equation modeling-based quantitative genetic models to characterize latent genetic and environmental influences on proneness to three discrete negative emotions in middle childhood, according to mother-report, father-report and in-home observation. One primary aim was to test the extent to which covariance among the three emotions could be accounted for by a single, common genetically- and environmentally-influenced negative emotionality factor. A second aim was to examine the extent to which different reporters appeared to be tapping into the same genetically- and environmentally-influenced aspects of each emotion. According to mother- and father-report, moderate to high genetic influences were evident for all emotions, with mother- and father-report of fear and father-report of anger showing the highest heritability. Significant common environmental influences were also found for mother-report of anger and sadness in both univariate and multivariate models. For observed emotion, anger was moderately heritable with no evidence for common environmental variance, but sadness, object fear and social fear all showed modest to moderate common environmental influences and no significant genetic variance. In addition, cholesky decompositions examining genetic and environmental influences across reporter suggested that despite considerable overlap between mother-report and father-report, there was also reporter-specific variance on anger, sadness, and fear. Specifically, there were significant common environmental influences on mother-report of anger- and sadness that were not shared with father-report, and genetic influences on father-report of sadness and fear that were not shared with mother-report. In-home observations were not highly correlated enough with parent-report to support multivariate

analysis for any emotion. Finally, according to both mother- and father-report, a single set of genetic and environmental influences was sufficient to account for covariance among all three negative emotions. However, fear was primarily explained by genetic influences not shared with other emotions, and anger also showed considerable emotion-specific genetic variance. In both cases, findings support the value of a more emotion-specific approach to temperament, and highlight the need to consider distinctions as well as commonalities across emotions, reporters and situations.

TABLE OF CONTENTS

Page

Appendix

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

Temperament, a multifaceted construct defined in terms of individual differences in reactivity and regulation across affective and behavioral domains (Rothbart & Bates, 2006), has been increasingly recognized as one important contributor to personality and adjustment in childhood and across the lifespan (Shiner, Buss, McClowry, Putnam, Saudino, & Ventner, 2012). There are moderate but consistently reported relations between temperament, socio-emotional competence, and resilience to stress in childhood and adolescence (Calkins, Blandon, Williford, & Keane, 2007; Eisenberg et al., 2004; Sanson, Hemphill, & Smart, 2004). Furthermore, although temperamental traits such as negative emotional reactivity, impulsivity and social inhibition are within the range of normal human variation and do not necessarily indicate maladaption or dysregulation (Rothbart & Bates; Derryberry, Reed, & Pilkenton, 2003), high levels of these traits in conjunction with lower self-regulation in childhood and adolescence do longitudinally predict internalizing and externalizing behavior problems (Eisenberg et al., 2009; Eisenberg, Spinrad, & Eggum, 2010; Oldehinkel, Hartman, Ferdinand, Verhulst, & Ormel, 2007). Research has also highlighted the potential role of childhood temperament in underlying vulnerability or resilience to psychopathology (Muris & Ollendick, 2005; Nigg, 2006; Perez-Edgar & Fox, 2005; Watson, Gamez, & Simms, 2005), perhaps especially under conditions of environmental stress (Compas, Connor-Smith & Jaser, 2004; Nigg, 2006). Even in normative environmental contexts, there is evidence that differences in emotional reactivity, sensitivity to reward or punishment, and self-regulation may moderate children's responses to parenting style and disciplinary practices

(Kochanska, 1991; Kochanska, 1995; Kochanska, Aksan, & Joy, 2007; Tschan, Kaiser, Chesney, Alkon, & Boyce, 1996), supporting the Goodness of Fit theory principal that children develop most successfully when parents adapt their own parenting style and family environment to complement their children's temperament (Thomas & Chess, 1977).

In light of the range of processes and outcomes that may be influenced by temperament, the ability to accurately assess temperament is important for both basic and applied developmental and clinical research. However, there has been disagreement regarding the most appropriate way to measure temperament, and in particular the widespread use of caregiver-report questionnaires (Kagan, 1992; Rothbart & Bates, 2006). Multimethod approaches have long been recommended (Kagan, 2007; Rothbart & Bates, 2006), and the use of observational measures in particular is often emphasized as a way to circumvent some of the limitations of caregiver-report questionnaires, including reporter biases such as mood- or personality-influenced reporting (e.g. Durbin & Wilson, 2011). However, less is known about the properties of many observational measures of temperament in comparison to established caregiver-report measures, and it is not currently clear to what extent observational measures assess situationally stable, dispositional traits, as opposed to context-specific reactions (Rothbart & Bates, 2006).

A second issue is the level of specificity at which temperament is analyzed. Much of the current research considers temperament only at the level of broad positive and negative reactivity, as opposed to directly examining children's tendencies toward discrete positive or negative emotions such as contentment or anger. However, those studies that do consider specific emotions often find differential prediction of later

temperament (Rothbart, Evans & Ahadi, 2000) and related outcomes such as externalizing and internalizing behavior problems (Calkins et al., 2007; Eisenberg et al., 2009; Lemery, Essex, & Smider, 2002; Oldehinkel et al., 2007). Depending on how broad components of temperament are defined and constructed, their physiological and genetic underpinnings and implications for child development may be very different.

Behavior genetic analyses provide one way to help clarify some of these questions. By considering correlations between individuals with different degrees of genetic relatedness, quantitative genetic studies allow an estimation of the variance within a given population which can be attributed to genetic or environmental influences, including the extent to which two traits are correlated for genetic or environmental reasons (Plomin, DeFries, McClearn, & McGuffin, 2008). If standardized observational assessments of temperament do index stable and constitutionally based individual differences, as opposed to day to day fluctuations in reactivity and regulation, these measures should show evidence of genetic or common environmental variance, and the presence of shared genetic or environmental influences between caregiver-report and observational measures of temperament suggests that both measures are to some degree indexing the same stable, constitutional trait, with discrepancies perhaps best attributed to immediate context or measurement error. In contrast, the presence of significant genetic or common environmental factors unique to observational measures indicates that the two types of measurement tap systematic and equally valid but different aspects of temperament, highlighting the importance of multi-method research.

In addition, quantitative genetic analyses can provide information about whether temperament is best assessed at the broad factor level or at the level of more specific

traits. To the extent that anger-, sadness- and fear-proneness represent different phenotypic manifestations of an underlying constitutional proclivity toward negative emotion, these lower-order traits should show considerable genetic overlap; in contrast, the presence of substantial genetic influences unique to each measure suggests systematic constitutional differences which may be lost when different aspects of negative emotionality are aggregated, indicating that a more specific level of analysis may be necessary. Although negative emotionality is one of the most commonly researched aspects of temperament on a phenotypic level, there have been no examinations of the genetic and environmental covariance among the traits of anger-, sadness- and fear-proneness in middle childhood, and studies examining the heritability of temperament using observed measures are uncommon outside of infancy and early childhood. Thus, the aim of the current study is to use quantitative genetic modeling to examine the genetic and environmental influences on the covariance between caregiver-report and in-home observations of negative emotionality in middle childhood, both at the broad composite level and at the level of discrete fear, anger and sadness, as well as the etiological overlap between fear, anger and sadness within type of measurement.

Temperament: Definition and Theoretical Basis

Temperament has been characterized in many different ways over time, from a relatively narrow focus on emotional reactivity, intensity and prevailing affective tendencies (Allport, 1961) to broader definitions encompassing not only emotionality but activity level, approach and withdrawal tendencies, and self-regulation (Kagan & Fox, 2006; Rothbart & Bates, 2006; Shiner et al., 2012; Thomas & Chess, 1977). Some theorists such as Buss and Plomin (1975) and Thomas and Chess (1977) have

emphasized early emergence, stability and hereditary or biological bases of temperament, to the extent that Buss and Plomin (1975) originally considered heritability and rank-order stability from infancy to adulthood to be necessary components of any temperamental trait, although current perspectives have moved away from such absolute requirements (Shiner et al., 2012). Kagan and Fox (2006) describe temperament as an emergent property of multiple interacting systems of neurochemistry and neural circuitry, with a prominent place given to systems related to inhibition and approach. Rothbart and Bates (2006), on the other hand, emphasize the importance of regulatory components of temperament and the interplay between biological and environmental factors, and between different aspects of temperament, in shaping the full profile of individual differences over time. Goldsmith and Campos (1986) take a restricted perspective in considering differences in the experience and regulation of primary emotions, including activity level as a potential indicator of general emotional arousal, but not behavioral tendencies such as impulsivity. Unlike many theorists (e.g. Buss & Plomin, 1975), Goldsmith and Campos also do not consider heritability a requirement of temperament. Such different emphases have contributed to different underlying assumptions and approaches to the study of temperament. For instance, Kagan (1992; 2007) emphasizes the value and objectivity of physiological measures of temperament, and argues that individual patterns of neurotransmission and neural structure may be associated with a multitude of distinct, categorical temperaments (Kagan & Fox, 2006). In contrast, Rothbart and Bates (2006) and Buss and Plomin (1975) focus on a few continuous, higher-order components of temperament, and defend the validity and utility of caregiver-report (Rothbart & Bates, 2006).

At the same time, considerable consensus has emerged on certain key elements. Across nearly all conceptualizations, temperament is thought to emerge early in development, to show relative stability across times and situations, and to be based at least in part on biological or constitutional differences between individuals (Rothbart & Bates, 2006; Shiner et al., 2012). These areas of consensus are largely supported by current research, although in many ways the development and nature of temperament is more complex than originally realized (Shiner et al., 2012). Many dimensions of temperament are highly heritable, especially but not only when assessed using caregiver-report instruments (Goldsmith, Buss & Lemery, 1997; Mullineaux, Deater-Deckard, Petrill, Thompson & DeThorne, 2009; Saudino, 2005), and there is evidence that neurological and hormonal factors can contribute to consistent differences in reactivity and regulation (DeYoung & Gray, 2009). However, there is also evidence that social and family factors including socioeconomic status and parenting play a role in the development of traits including fear, irritability and self-regulation in childhood and adolescence (Davidov & Grusec, 2006; Lengua & Kovacs, 2005; Moilanen, Shaw, Dishion, Gardner, & Wilson, 2010). Although most aspects of temperament do show low stability during infancy and early childhood (Lemery, Goldsmith, Klinnert, & Mrazek, 1999), they reach moderate stability by middle childhood, with continuity over time increasing throughout adolescence and adulthood (Neppl et al., 2010; Roberts & Del Veccio, 2000). However, moderate and gradual change also continues throughout the lifespan (Roberts & Del Veccio, 2000), and normative maturational changes such as increasing capacity for self-regulation throughout childhood, heightened emotional intensity and variability in adolescence, and hormonal and neurochemical changes

triggered by puberty can all influence both the manifestation and the underlying structure of temperament (Dahl, 2009; Gunnar, Wewerka, Frenn, Long, & Griggs, 2009; Posner & Rothbart, 2009; Putnam, Rothbart, & Gartstein, 2008).

Indeed, both biological and environmental influences may contribute to both stability and change in temperament over the lifespan (Fish, Stifter & Belsky, 1991; Hagekull & Bohlin, 1998; Posner & Rothbart, 2009; Posner, Rothbart, Sheese & Voelker, 2011). For instance, the early instability of many behavioral and emotional aspects of temperament coincides with and may be due to the rapid maturation of biological systems involved in attention and regulation during the first few years of life, including the executive attention network (Posner & Rothbart, 2009; Posner et al., 2011) and the vagus nerve (Porges, 2007), and there is some evidence that continuity in environmental factors including parenting and marital quality may promote stability in temperament (Fish et al., 1991; Hagekull & Bohlin, 1998; Rothbart & Bates, 2006). Quantitative genetic studies also suggest that the rate and extent of normative maturational changes in temperament may themselves be genetically influenced (Saudino, 2005). Matheny (1989) reports not only higher concurrent similarity in behavioral inhibition between monozygotic (MZ) twins relative to dizygotic (DZ) twins but more similarity in direction and magnitude of change across 12 to 30 months. Other quantitative genetic studies suggest new genetic influences on observed extraversion, persistence, behavioral inhibition and activity level in early childhood that come online with age, although such differences may also be due to changes in measurement with age (Saudino, 2005).

Organization of Temperament in Middle Childhood

Some consensus has also emerged regarding certain major components of temperament, at least on a broad conceptual level. Factor analytic studies across a number of caregiver-report measures of temperament in middle childhood, including the Child Temperament Questionnaire (Sanson, Smart, Prior, Oberklaid & Pedlow, 1994; Thomas & Chess, 1977), the Middle Childhood Temperament Questionnaire (Hegvik, McDevitt, & Carey, 1982; McClowry, Hegvik, & Teglasi, 1993), and the Children's Behavior Questionnaire (Rothbart, Ahadi, Hershey, & Fisher, 2001), have consistently revealed components related to negative emotionality, to self-regulation, and to extraversion, high-activation positive emotion or high approach tendencies (Rothbart & Bates, 2006). Many temperament questionnaires in middle childhood also yield distinct factors related to activity level, agreeableness, and behavioral inhibition or shyness (McClowry et al., 1993; Presley & Martin, 1994; Rothbart & Bates, 2006). Although the factor structure of temperament depends to some degree on the reporter and the nature and range of items included in each questionnaire, the consistency of components related to negative emotionality and self-regulation across studies is notable (Rothbart & Bates, 2006).

Rothbart and colleagues' (2001) Children's Behavior Questionnaire (CBQ) assesses three broad factors, denoted as negative emotionality, effortful control and extraversion/surgency. Negative emotionality as assessed by the CBQ in middle childhood is composed of anger, fear, sadness, low soothability, and discomfort. Effortful control includes aspects of appropriate attentional allocation, effortful behavioral inhibition and low activation positive emotion (e.g. contentment). Extraversion/surgency

includes aspects of high intensity positive emotion, approach, and impulsivity, as well as low shyness (Rothbart et al., 2001). These three components are distinct in the sense that an individual's temperament on any one component may or may not coincide with their levels of the other two (Rothbart et al., 2001), but conceptually and empirically related in the sense that each component may influence the expression of and outcomes related to the others, both concurrently and over time (Rothbart & Bates, 2006). For instance, high levels of effortful control may enable the regulation of biologically-based tendencies toward emotional reactivity, approach or withdrawal that might otherwise manifest in uncontrolled or problem behaviors, and low effortful control may be less problematic in the absence of strong reactive or impulsive tendencies that demand regulation (Eisenberg et al., 2009; Frick & Morris, 2004). As such, it is likely the full profile of an individual's temperament, as much as the strength of any given trait or component, that matters most for child adjustment (Caspi & Shiner, 2006; Crocket, Moilanen, Raffaelli & Randall, 2006; Rettew, Althoff, Dumenci, Ayer & Hudziak, 2008; Van den Acker, Decovic, Prinzie & Asscher, 2010).

In addition, temperament and personality can be thought of as hierarchically structured (Markon, Krueger, & Watson, 2005; Tackett, Krueger, Iacono, & McGue, 2008), with broader, higher-order components such as negative emotionality, surgency or effortful control divided into increasingly specific, discrete subordinate dimensions (Rothbart & Bates, 2006). For example, the dimensions of negative emotionality in middle childhood (sadness, anger, fear, shyness, discomfort, and low soothability) may be broken down into more basic facets (e.g. social versus object fear). As such, the three-, four- and five-factor solutions found by different studies of temperament are not mutually

exclusive, but represent different levels of analysis (Markon et al., 2005; Tackett et al., 2008), and there may be unique variance and complex interrelations between traits considered at each level (Markon et al., 2005). A great deal of valuable temperament research has taken place at the level of higher order components, but it is also becoming recognized that a fine-grained approach to temperament can be more appropriate for some research questions, and may help to clarify inconsistencies in findings from studies at higher levels of abstraction (Rothbart & Bates, 2006).

Specificity in the Study of Temperament

Although broad components similar to negative emotionality, effortful control and extraversion/surgency are consistently reported, they are not precisely equivalent across studies. For instance, some researchers use a global measure of negative emotionality encompassing aspects of fear, anger and sadness (e.g. Lengua, West, & Sandler, 1998; Sallquist et al., 2009; Valiente et al., 2003), some choose to analyze a single aspect of negative emotionality (often anger/frustration; e.g. Degnan, Calkins, Keane & Hill-Soderlund, 2008; Diener & Kim, 2004; Morris, Silk, Steinberg, Sessa, Avenevolli & Essex, 2002), and others consider proneness to different negative emotions within the same study (e.g. Lemery et al., 2002; Oldehinkel et al., 2007). Whether traits related to fear, inhibition or shyness are assessed and analyzed separately or considered as part of a more general negative emotionality or surgency component may influence findings relating negative emotionality to childhood social and emotional adjustment and mood and behavioral problems in important ways.

Negative emotionality assessed broadly has been found to predict both internalizing and externalizing behavior problems (Caspi, Henry, McGee, Moffitt, &

Silva, 1995; Frick & Morris, 2004; Gilliom & Shaw, 2004; Gjone & Stevenson, 1997; Nigg, 2006; Sanson, Oberklaid, Pedlow & Prior, 1991), particularly in conjunction with low effortful control (Caspi, 2000; Eisenberg et al., 2009; 2010; Nigg, 2006), although relations with externalizing problems tend to be stronger and more consistent (Rothbart & Bates, 2006; Sanson et al., 2004). However, there is some inconsistency regarding the specificity of associations between negative emotionality, internalizing and externalizing, as well as some non-replications, and sample ages and measurement styles vary between studies (Rothbart & Bates, 2006; Sanson et al., 2004). For instance, in one early study, Lengua and colleagues (1998) report that both mother and child report of broad negative emotionality in middle childhood are related to conduct problems and depression. In contrast, teacher-report of negative emotionality from ages three to six years does not predict children's aggression (Russell, Hart, Robinson, & Olsen, 2003), although Valiente and colleagues (2003) report that combined teacher- and parent-report negative emotionality may moderate the relation between effortful control and externalizing problems at 11 years. Parent- and teacher-report difficult temperament, defined as a combination of high negative emotionality and demandingness, has also been found to longitudinally predict the covariance between internalizing and externalizing problems from middle childhood to early adolescence even after accounting for family and social risk factors, although difficult temperament is related to neither pure internalizing nor pure externalizing (Keiley, Lofthouse, Bates, Dodge, & Pettit, 2003).

In contrast to the associations between broad negative emotionality and externalizing problems, alone and in conjunction with internalizing (Frick & Morris, 2004; Mun, Fitzgerald, von Eye, Puttler, & Zucker, 2001; Rothbart & Bates, 2006;

Sanson et al., 2004; Sanson & Prior, 1999), fear, shyness and withdrawal appear to be more strongly related to pure internalizing problems (Kagan & Snidman, 1999; Mun et al., 2001; Nigg, 2006; Oldehinkel, Hartman, deWinter, Veenstra, & Ormel, 2004; Schwartz, Snidman & Kagan, 1999), and may be protective against externalizing (Keiley et al., 2003; Russell et al., 2003; Sanson & Prior, 1999; Schwartz, Snidman & Kagan, 1996). For instance, in one study of preschoolers, Diener and Kim (2004) find that mother report of anger and frustration was positively related to externalizing behavior and negatively related to prosocial behavior, independently and in interaction with self-regulation, whereas shyness was unrelated to prosocial behavior and protective against externalizing. It is also possible that negative emotional reactivity and fearfulness interact to predict children's risk for problem behaviors; in one longitudinal study of boys from low income families, a composite of experimenter-coded and mother-report negative emotionality predicted higher initial levels of both internalizing and externalizing problems, as well as steadily high trajectories of externalizing symptomatology from 2 to 6 years in conjunction with negative maternal control and low levels of fear, and steadily high trajectories of internalizing symptomatology in conjunction with negative maternal control and high levels of fear (Gilliom & Shaw, 2004).

However, as noted earlier, some conceptions of negative emotionality have included aspects of fear or sadness (e.g. Lengua et al., 1998; Russell et al., 2003), whereas others have focused more narrowly on anger and frustration (e.g. Degnan et al., 2008; Morris et al., 2002), and studies which distinguish between anger, fear and sadness often report differential associations with concurrent and later behavior problems (Eisenberg et al., 2009; Lemery et al., 2002; Oldehinkel et al., 2007; Rothbart & Bates,

2006). Eisenberg and colleagues (2004) report that teacher-report anger, but not sadness or parent-report anger, moderates the relations between effortful control or impulsivity with externalizing in early to middle childhood, whereas sadness was related to internalizing problems through its association with lower behavioral adaptability. Lemery and colleagues (2002) also find evidence for differential prediction, even after overlapping items between temperament and adjustment scales are removed, such that relations between parent-report of preschool-age children's temperament and hostile-aggressive behavior are strongest for anger and, inversely, for effortful control. In contrast, sadness and, to a lesser degree, fear and anger predict children's anxious behavior. Similar results are reported by Oldehinkel and colleagues (2007) in a sample followed from preadolescence to early adolescence, with fear associated with internalizing problems and frustration associated with both internalizing and externalizing, although relations between frustration and externalizing were again found to be stronger. Some evidence also exists for associations between sadness and externalizing behaviors in childhood and adolescence (Eisenberg et al., 2005), and between anger and internalizing (Lemery et al., 2002; Morris et al., 2002; Oldehinkel et al., 2007), particularly over time (Eisenberg et al., 2009), but these associations are weaker and may be mediated through social processes such as peer rejection (Eisenberg et al., 2009). In general, predispositions to sadness and fear or anger are either found to be directly linked more strongly to internalizing or externalizing problems, respectively (Muris & Ollendick, 2005), or to act as moderators of other risk factors such as lower effortful control (e.g. Calkins et al., 2007; Muris & Ollendick, 2005; Oldehinkel et al.,

2007) or environmental stress (e.g. Bates, Pettit. Dodge, & Ridge, 1998; Leve, Kim & Pears, 2005).

Furthermore, specific negative and positive emotional tendencies in infancy and early childhood differentially predict temperament in middle childhood (Rothbart et al., 2000). Anger/frustration during laboratory tasks at 13 months predicts higher anger/frustration and lower soothability at 7 years, but is unrelated to fear or sadness, and early approach tendencies and activity level are associated with both anger/frustration and aspects of extraversion/surgency, including high-intensity pleasure and impulsivity (Rothbart et al., 2000). In contrast, fearfulness at 13 months predicts later fear, shyness and sadness, as well as low-intensity pleasure (Rothbart et al., 2000), which is considered an aspect of effortful control (Rothbart et al., 2001). It is not clear whether these associations are due to age-related differences in manifestation of the same underlying temperament or the result of elicitation of and interaction with environmental factors such as parenting, but continuity between specific negative emotional tendencies and differences in the prediction of self-regulation, impulsivity and high- versus low-intensity pleasure do suggest that differences are meaningful and stable enough to justify analyzing individual differences in discrete negative emotions separately.

There are also theoretical reasons to consider temperament on a more specific level. The question of whether emotions represent discrete patterns of physiological and cognitive activation (e.g. Ekman, 1992) or arise from a more general combination of affect, arousal and attribution (e.g. Russell, 1980) has yet to be fully resolved, but on a behavioral and cognitive level, discrete emotions do have distinct antecedents, functions and consequences for behavior and adjustment (Izard, 2007; Scarantino & Griffiths,

2011). Sadness, anger and fear, all negative emotions, are considered in functionalist and evolutionary perspectives to serve different social and adaptive functions (Nesse & Ellsworth, 2009; Tooby & Cosmides, 2008). Generally, sadness may promote withdrawal and conservation of resources after loss, fear is a response to perceived threat or danger which may lead to avoidance or aggression if avoidance is not possible, and anger facilitates aggression, promotes the acquisition or seeking of resources and in social species such as humans may promote social recalibration following a perceived wrong (Nesse & Ellsworth, 2009; Tooby & Cosmides, 2008). The different motivational aspects suggest that there is merit in differentiating between emotions in the prediction of particular outcomes (e.g. internalizing versus externalizing problems or social aggression versus withdrawal).

For example, anger has been conceptualized as an approach emotion in the sense that it is provoked by blocked goals and elicits approach-related behaviors (Berkowitz, 1989; E. Harmon-Jones & Allen, 1998; Tooby & Cosmides, 2008), which may explain the moderate positive associations between anger and positive emotionality or extraversion reported in some studies of temperament (Donzella et al., 2000; E. Harmon-Jones, C. Harmon-Jones, Abramson, & Peterson, 2009; C. Harmon-Jones, Schmeichel, Mennitt, & E. Harmon-Jones, 2011; Nigg, 2006; Rydell, Berlin & Bohlin, 2003). Factors tapping positive emotionality or extraversion sometimes include items related to approach or impulsivity (e.g. extraversion/surgency or sociability; Gartstein & Rothbart, 2003; Rothbart et al., 2001; Sanson, Smart, Prior, Oberklaid & Pedlow, 1994; Shiner, 1998), which have been implicated in both intense positive emotion or reward sensitivity

and anger (Carver & E. Harmon-Jones, 2009; Lara & Askikal, 2006; Rothbart, Derryberry & Posner, 1994).

There is also evidence from human and animal studies for distinct biological systems underlying different positive and negative emotions, considering peripheral physiological response patterns, neural structure and activation, and chemical neurotransmission (Berridge, 2003; Larsen, Berntson, Poehlmann, Ito, & Cacioppo, 2008; Panksepp & Watt, 2011; Vytal & Hamann, 2010), although findings in some cases are mixed (Murphy, Nimmo-Smith, & Lawrence, 2003; Phan, Wager, Taylor, & Liberzon, 2002). The brain is highly interconnected and dynamic, with extensive connections between regions implicated in sensation and perception, emotion, memory, attention and regulation, which develop and may become more integrated or more differentiated with age (Fair et al., 2009; Gao et al., 2009; Rothbart & Bates, 2006), and no single neural circuit or neurotransmitter can be said to underlie any dimension of temperament or affect (Kagan & Fox, 2006; Paris, 2005; Zuckerman, 1995). However, a growing body of research offers evidence of multiple neural and peripheral systems involved in both broad and narrow aspects of temperament and personality (DeYoung & Gray, 2009; Fortunato, Dribin, Granger, & Buss, 2008; Rothbart & Bates, 2006; Zuckerman, 1995).

On a general level, frontal EEG asymmetry has been linked to differences in positive and negative affect in humans, such that individuals who show right frontal asymmetry have been found to display higher withdrawal tendencies and withdrawal-associated negative emotions, whereas left frontal asymmetry is associated with higher approach, positive emotionality and anger (Davidson, 2002; E. Harmon-Jones, Vaughn-

Scott, Mohr, Sigelman, & C. Harmon-Jones, 2004; Kagan & Fox, 2006). In addition, some researchers such as Kagan and Fox (2006) have suggested that serotonin and dopamine are the primary neurotransmitters associated with broad differences in approach and withdrawal tendencies, and DeYoung and Gray (2009) have suggested that adult personality may be broadly organized into two higher order factors, stability and adaptability, which may be influenced by serotonin and dopamine, respectively. Reward sensitivity and approach tendencies do appear to be influenced in part by a dopaminergically-modulated circuit that includes the ventral tegmental area, basolateral amygdala and nucleus accumbens (Depue & Collins, 1999). Dopamine is also involved in reward-based learning, exploratory behavior and anticipatory behavior (Alcaro, Huber, & Panksepp, 2007; Ashby, Isen & Turken, 1999; Depue & Collins, 1999), and dopaminergic activity in the prefrontal cortex and anterior cingulate cortex modulates the effortful regulation of attention and cognitive flexibility (Braver & Barch, 2002; da Silva Alves et al., 2011; Posner et al., 2011), which may provide the foundations of self-regulation in childhood (Posner et al., 2011). In contrast, serotonin is involved with a range of inhibitory processes and widely related to withdrawal, sensitivity to punishment and aversive learning (Heisler et al., 2007; Lucki, 1998; Spoont, 1992); low serotonergic activity has also been implicated in vulnerability to depression and anxiety (Holmes, 2008; Krishnan & Nestler, 2008; Lucki, 1998), general tendencies towards negative affect (Lesch, 2003; Spoont, 1992), and impulsive aggression (Manuck, Flory, McCaffrey, Matthews, Mann, & Muldoon, 1998; Lesch & Merschdorf, 2000; Lucki, 1998).

The hypothalamic-pituitary-adrenal (HPA) axis is also implicated in both fear and anxiety and threat-based aggression through its role in sensitivity and reactivity to threat and stress (Carter, 1986; McEwen, 1998), and higher reactive cortisol in laboratory testing and emotion-elicitation tasks, as well as during exposure to stressful environmental conditions such as adverse childcare, has been linked to higher tendencies toward negative affect and withdrawal (Blaire, Peters, & Granger, 2004; Buss, Schumacher, Dolski, Kalin, Goldsmith, & Davidson, 2003; Dettling, Parker, Lane, Sebanc, & Gunnar, 2000; Fortunato et al., 2008). Some evidence also exists that salivary alpha amylase is associated specifically with approach and positive emotion in toddlers during laboratory assessment, although replication is needed (Fortunato et al., 2008). However, higher baseline cortisol has been linked with inhibited temperament in some (Buss et al., 2003; Kagan, Reznick, & Snidman, 1987; Schmidt et al., 1997), but not all (Schmidt, Fox, Schulkin, & Gold, 1999) studies, with inhibited temperament related to higher baseline cortisol in the home and angry or aggressive tendencies related to elevated cortisol in response to starting school (de Haan, Gunnar, Tout, Hart, & Stansbury, 1997). There is also evidence that infants high in distress to limitations, rather than distress to novelty, show the highest increases in cortisol in response to maternal separation (Gunnar, Larson, Hertsgaard, Harris, & Broderson, 1992), and preschoolers high in surgency and low in effortful control have been found to show elevated cortisol reactivity and anger in response to losing a competitive game (Donzella et al., 2002), suggesting that relations between temperament and cortisol reactivity to stress can be ambiguous and dependent on the social context and the nature of the stressor (Kagan & Fox, 2006; Rothbart & Bates, 2006).

Although there is evidence for broad differences in reward and punishment sensitivity, approach and withdrawal that may contribute to common influences on negative and positive emotional predispositions (DeYoung & Gray, 2009; Fox et al., 2005; Rothbart, Derryberry & Posner, 1994), there is also evidence that distinct patterns of neural activation and neurotransmission may be involved in more specific emotional tendencies (Vytal & Hamann, 2010; Damasio, Grabowski, Bechara, Damasio, Ponto, Parvizi, & Hichwa, 2000; Panksepp, 1993; 2010). Anger has been found to involve the dopaminergic approach circuits (Lara & Askikal, 2006; Rothbart et al., 1994), as well as ventromedial hypothalamus and dorsal periaqueductal gray (Panksepp, 1982; 2010), whereas fear and anxiety are associated with a network which appears to be largely modulated by serotonin and norepinephrine (DeYoung & Gray, 2009; Lara & Askikal, 2006) and includes the lateral and central amygdala and hippocampus, anterior cingulate cortex, medial septal area and orbitofrontal cortex (Canli, Sivers, Whitfield, Gotlib, & Gabrieli, 2002; Davidson et al., 2002; Johnson et al., 1999; Sugiura et al., 2000 Tauscher, Bagby, Javanmard, Christensen, Kasper, & Kapur, 2001). But one meta-analysis of neuroimaging studies in humans also points to specific subcortical regions which are consistently activated during the elicitation of discrete emotions, including the left inferior frontal gyrus (IFG) for anger, the left medial frontal gyrus (MFG) for sadness and the left amygdala for fear (Vytal & Hamann, 2010). Differential activation in these and other regions also successfully differentiated between discrete emotions; relatively greater activation in the left MFG distinguished sadness from fear and anger, activation in the right parahippocampal gyrus and left IFG was associated with higher anger than sadness and fear, respectively, and greater activation in the left amygdala and putamen

was associated with higher fear than sadness and anger, respectively (Vytal & Hamann, 2010). Genetically or environmentally influenced differences in any of these regions may be associated with differences in the frequency or intensity of individual differences in discrete emotions, and temperament.

Measurement of Temperament

Both caregiver-report and observational measures are used in the study of temperament from infancy and early childhood, although studies of adolescents and adults more commonly rely on self-report (Rothbart & Bates, 2006). Observational measures of temperament might take the form of structured observational assessment of children during tasks or situations designed to elicit specific responses, or observer-report ratings of children's temperament summarized at the end of an assessment period. Both caregiver-report and laboratory observation of temperament are effective predictors of child adjustment, socioemotional development, and mood and behavioral problems (e.g. Caspi & Silva, 1995; Eisenberg et al., 2001; Hayden, Klein, & Durbin, 2005; Murphy, Shepard, Eisenberg & Fabes, 2004; Rothbart & Bates, 2006). However, convergence between the two types of measurement tends to be small, with correlations between caregiver-report and laboratory observations often in the range of .00 to .40 for both positive and negative emotion (Durbin & Wilson, 2011; Forman, O'Hara, Larsen, Coy, Gorman, & Stuart, 2003; Goldsmith & Campos, 1990; Kochanska, Coy, Tjebkes, & Husarek, 1998; Seifer, Sameroff, Barrett, & Krafchuck, 1994; Stifter et al., 2008). Further, one study in which parents were brought into the laboratory, trained and instructed to observationally rate their children's emotions found that convergence with experimenter-report remained low, and a second found significant associations between

parent predictions of child distress and child behavior in laboratory episodes related to fear and social inhibition but not frustration or disappointment (Kiel & Buss, 2006), suggesting that discrepancies between caregiver- and observer-report cannot be ascribed to context effects alone (Durbin & Wilson, 2011).

Observational measures, both within and outside a standardized laboratory setting, are often considered the gold standard of measurement in the field, whereas caregiver-report questionnaires have been subject to criticism on the grounds that these measures introduce multiple potential sources of systematic bias (Kagan & Fox, 2006; Rothbart & Bates, 2006; Seifer, Sameroff. Dickstein, Schiller, & Hayden, 2004). Kagan (1992; 2007) argues that observational and physiological measures of temperament are imperfect but relatively objective, whereas caregiver-report and self-report measures of personality and temperament are vulnerable to reporters' idiosyncratic interpretations of question wording, social acceptability bias, and lack of an objective basis for comparison, although others have noted that parents may perform better than observers on assessing their children's emotional variability (Durbin & Wilson, 2011).

Caregiver-report measures do have several limitations. One common problem in twin and adoption studies has been parents' tendency to contrast one twin or sibling against the other, leading to biased heritability estimates (Saudino, 2005; 2003). Caregiver reports can also be influenced by caregivers' expectations of children's temperament and behavior (Seifer et al., 2004) and caregivers' own mood and dispositional qualities (Durbin & Wilson, 2011; Forman et al., 2003), although some researchers have argued that the assumption that the evidence for depressed mothers having distorted perceptions of their children is unsupported (Richters, 1992). Durbin and

Wilson (2011) found that parents' ratings of children's negative emotions in the laboratory were more highly correlated with parental depression and parent personality characteristics than with observer ratings of children's negative and positive emotion, suggesting that mood-based positive and negative biases are not limited to retrospective recall. However, it is important to note that parents may have been taking past behavior into account even within a laboratory setting, leading to assessments of child behavior that are more trait-like than the ratings of laboratory observers unfamiliar with the children. In addition, although many studies have found associations between maternal depression (Leerkes & Crockenberg, 2003), negative affect (Gartstein & Marmion, 2008) and anxiety (Bates & Bayles, 1984) and concordance between parent- and observer-ratings of temperament, there is also evidence that parent-report and laboratory observations of temperament are equally good longitudinal predictors of teacher-rated disruptive school behavior (Hayden et al., 2005), suggesting that even given the existence of rater bias, parent-report of child temperament remains a valid method of assessment.

Finally, it is worth noting that parents and observers both may be more adept at noticing externally directed or high intensity emotions such as anger or excitement, or emotions such as fear with more obvious behavioral manifestations, relative to internally directed or low intensity emotions such as sadness and contentment. Correlations between mother- and father- report of five year olds' temperament using the CBQ have been found to be highest for shyness, impulsivity and inhibitory control ($r > .70$), and lowest for discomfort ($r = .28$) and smiling and laughter ($r = .34$), with moderate agreement for most other scales (Rothbart et al., 2001). Such results suggest that parents are more consistent in detecting aspects of childhood temperament with more obvious

behavioral manifestations, although discrepancies may also be due to differences in parental elicitation of children's emotions (Rothbart et al., 2001). There is also evidence of higher convergence between caregiver-report and laboratory observation (Hayden et al., 2005), and between mother- and father-report (Christensen, Margolin, & Sullaway, 1992; Duhig, Renk, Epstein, & Phares, 2000) on children's externalizing behavior than internalizing behavior, which lends credence to the possibility that parents are more accurate in detecting externally-directed emotion and behavior. If so, to the extent that more externally-directed emotions are overrepresented in broad components of negative and positive emotionality, findings related to these broad factors may not reflect the full range of possible relations between temperament and child development. Thus, although predictions based on broadly-assessed negative and positive emotionality are often informative, further specificity is both possible and valuable, even within the context of stable and consistently derived higher-order factors.

Many of these limitations can be reduced greatly by using measures which assess the frequency of specific behaviors within a limited frame of time, as opposed to global judgments of temperament or emotion (Hwang & Rothbart, 2003; Rothbart & Goldsmith, 1985). Questionnaires assessing specific behaviors are less vulnerable to contrast effects (Hwang & Rothbart, 2003), and some studies using the Infant Behavior Questionnaire (IBQ; Rothbart, 1981) and and Toddler Behavior Assessment Questionnaire (TBAQ; Goldsmith, 1996) find no evidence of contrast effects (Goldsmith, Buss, & Lemery, 1997; Goldsmith, Lemery, Buss, & Campos, 1999), although the CBQ does show evidence of contrast effects in some (Goldsmith et al., 1997; Majdandzic, van den Boom, & Heesbeen, 2008) but not all (Mullineaux et al., 2009) studies. Such measures may also

address Kagan's (1992; 2007) concerns regarding idiosyncratic interpretation and lack of objective standards for children's behavior, as items focus on specific behaviors (e.g. response to being strapped in a carseat), and do not require raters to compare their children's behaviors to other children's behaviors.

Furthermore, as Rothbart and Bates (2006) emphasize, there are also considerable advantages to the use of caregiver-report measures. Caregivers are familiar with children's behavior and emotional reactivity across a wider range of times and situations, theoretically rendering caregiver-report measures less influenced by context effects and day-to-day variation (Majdandzic et al., 2008; Mangelsdorf, Schoppe, & Burr, 2000; Rothbart & Goldsmith, 1985), and thus provide relevant information about a construct such as temperament which is characterized by consistency in behavioral, emotional and regulatory tendencies (Shiner et al., 2012). Caregivers may also be more likely to observe lower-frequency child behaviors across a variety of naturalistic, ecologically valid settings, potentially offering a more complete picture of children's reactivity and regulation than that which can be observed in more restricted laboratory settings (Rothbart & Bates, 1998; 2006; Rothbart & Goldsmith, 1985).

Finally, although lack of convergence between caregiver-report and observational measures has been largely interpreted as reflecting error or bias inherent in caregiver-report (e.g. Kagan, 1992; 2007), it is uncertain to what extent such discrepancies are due to unreliability in caregiver-report as opposed to observation (Rothbart & Bates, 2006). Even in a standardized laboratory setting, observation carries its own sources of unreliability, including observer dispositional characteristics and interactions with the child, ambiguity in coding low intensity behaviors or reactions, and context or carryover

from previous events or episodes (Rothbart & Bates, 2006; Rothbart & Goldsmith, 1985). Standardized laboratory assessments examine children's behavior in artificial and restricted situations which may not reflect the full range of behavior and emotional response (Rothbart & Bates, 2006; Rothbart & Goldsmith, 1985; Stifter et al., 2008). Although many laboratory assessments, such as the Laboratory Temperament Assessment Battery (Lab-TAB; Goldsmith & Rothbart, 1999), strive to replicate typical situations facing children in day-to-day life, it is unclear to what extent children's behavior is influenced by the novelty and artificiality of the situation (Rothbart & Bates, 2006). In addition, laboratory assessments are often able to consider only a limited window of time, which may also restrict the range and intensity of behavior and emotion displayed (Rothbart & Goldsmith, 1985; Rothbart & Bates, 1998). Because of the sensitivity of observational measures to environmental context and daily fluctuations in mood and behavior, the use of multiple episodes can be necessary to achieve adequate reliability; in one study of preschoolers' activity level, measured using actometers, the reliability of single episodes was low but increased dramatically with aggregation, rising from .13 to .75 (Eaton, 1994). Aggregating data across multiple episodes has also been found to increase convergence with caregiver-report of activity level, positive and negative emotionality (Eaton, 1994; Forman et al., 2003), although it is worth noting that correlations typically remain modest.

It may well be that caregiver-report and observational measures capture distinct, but equally valid, aspects of temperament. Laboratory observations offer a highly objective, controlled measure of moment-to-moment differences in children's reactivity and behavior in response to novel tasks or situations and interactions with strangers,

whereas caregiver-report measures may provide more insight into children's temperament in the context of daily environments and interactions with caregivers (Rothbart & Bates, 2006). All methods of measurement have strengths and weaknesses, and multiple methods complement each other and allow for a more complete picture of childhood temperament. At the same time, understanding the characteristics of temperament as assessed by different measures, including the presence of distinct and shared genetic and environmental influences on caregiver-report and laboratory observation of temperament, can prove useful for interpreting both relations and discrepancies between them.

Genetics of Temperament

There is strong support for the idea that temperament is both mediated by physiological differences between individuals and at least partly hereditary (Goldsmith et al., 1997; Goldsmith et al., 1999; Rothbart et al., 2011), with both quantitative and molecular genetic research pointing to the possibility of both direct genetic effects (e.g. Laucht, Becker, Blomeyer & Schmidt, 2007; Goldsmith et al., 1997) and gene-environment interplay (e.g. Hankin et al., 2011; Jang, Dick, Wolf, Livesly, & Paris, 2005; Krueger, South, Johnson, & Iacono, 2008). Gene-environment interplay encompasses both gene-environment interaction, in which the phenotypic expression of an individual's genes depends on the environments to which that individual is exposed (Rutter, Moffitt, & Caspi, 2006), and gene-environment correlation, in which an individual's genes and the environments to which they are exposed or select into are not independent (Scarr & McCartney, 1983). In the case of temperament, there is evidence that both processes may

operate (Rothbart & Bates, 2006), although gene-environment correlation has not been widely studied.

Quantitative genetic research considers the importance of overall genetic and environmental influences at a population level, without attempting to isolate specific genes or environmental factors contributing to phenotypic variation in a trait (Lemery-Chalfant, 2010; Neale & Cardon, 1992). Estimates of the heritability (proportion of phenotypic variance explained by genetic factors) of both caregiver-report and laboratory observation of temperament are often substantial across childhood (Goldsmith et al., 1997; Mullineaux et al., 2009; Saudino, 2005), generally ranging from 20% to 60% across multiple dimensions of temperament (Saudino, 2005), and longitudinal stability in dimensions including behavioral inhibition, activity level, negative emotionality, anger and inhibitory control is largely accounted for by genetic factors (Gagne, Vendlinski, & Goldsmith, 2009; Gagne & Goldsmith, 2011; Ganiban, Saudino, Ulbricht, Neiderhiser, & Reiss, 2008; Hoekstra, Bartels, Hudziak, Van Beijsterveldt, & Boomsma, 2008; Saudino, Plomin, & DeFries, 1996). However, some caregiver-report measures of temperament, such as the Emotionality, Activity, Sociability/ Emotionality, Activity, Sociability, Impulsivity scale (EAS/EASI; Buss & Plomin, 1975; 1984) and the Colorado Childhood Temperament Inventory (CCTI; Rowe & Plomin, 1977), are prone to overestimation of heritability due to twin contrast effects (Goldsmith et al., 1997; Hwang & Rothbart, 2003), and only studies which directly model contrast effects or use measures demonstrated to be relatively less vulnerable to such biases (e.g. the IBQ and TBAQ; Hwang & Rothbart, 2003) will be considered here.

In addition, it is important to note that heritability is a general statistic describing the total proportion of phenotypic variance in a trait that can be explained by genetic differences between individuals, and as such is dependent on both genetic and environmental heterogeneity within a population (Neale & Cardon, 1992; Lemery-Chalfant, 2010). Thus, heritability is specific to the sample in question at a given point in time, and there is evidence that the heritability of a range of traits including externalizing behavior problems, antisocial behavior, internalizing problems, physical health and IQ does differ on the basis of environmental factors including punitive discipline, parenting warmth and negativity, marital quality and socioeconomic status (Button, Maughan, & Eley, 2008; Feinberg, Button, Neiderhiser, Reiss, & Hetherington, 2007; Johnson & Krueger, 2005; South & Krueger, 2008; Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003), although not all studies have shown evidence of moderated heritability (Kendler, Aggen, Jacobsen, & Neale, 2003; Rutter et al., 2006).

Only limited research to date has tested such measured environment by heritability interactions in relation to temperament or personality, but Boomsma, van Baal, and Koopmans (1999) report that the heritability of disinhibited personality is contingent on religious upbringing, such that genetic influences on disinhibition were higher for twins who reported a non-religious upbringing, (although the difference was only significant for males) suggesting that cultural factors can constrain the expression of genetic tendencies toward disinhibited behavior.

In addition, two quantitative genetic studies of personality in adolescents and adults have found that genetic contributions to emotionality tend to be stronger in more positive environments, even after accounting for gene-environment correlation (Jang et

al., 2005; Krueger et al., 2008). Krueger and colleagues (2008) report that estimates of genetic variance for both negative and positive emotionality are highest for adolescents who report high mutual positive regard with parents, and that youth report of conflict with parents similarly moderated genetic contributions to negative, but not positive, emotionality. Jang and colleagues (2005) found a similar pattern between emotional instability and perceived family conflict in adults, with lower heritability at higher levels of family conflict. Unfortunately, the results of these studies are difficult to interpret because although the personality traits studied are similar to negative and positive emotionality in childhood, the use of concurrent (Krueger et al., 2008) and, more problematically, retrospective (Jang et al., 2005) self-report introduces the possibility that participants' reports of family environment were affected by differences in emotionality. Further research in childhood using more objective measurement of the family environment is needed before strong conclusions can be drawn, but the findings from these adult personality studies do suggest that genetic predispositions toward higher positive and lower negative emotionality may be constrained in more stressful environments.

In contrast to the broad population-based approach of quantitative genetic research, molecular genetic studies use individual neurochemical differences as a theoretical basis for seeking associations between traits such as behavioral inhibition or impulsivity and functional polymorphisms in genes related to neurotransmitter synthesis, reception and reuptake. Widely studied genes include the serotonin transporter gene (SLC6A4; Heils et al., 1996; Lesch et al., 1994; Lesch et al., 1994), the gene coding for the dopamine D4 receptor (DRD4; Ding et al., 2002; Van Tol et al., 1992), and the gene

coding for catechol-*O*-methyl transferase (COMT; Lachman, Papolos, Saito, Yu, Szumlanski & Weinshilboum, 1996), an enzyme responsible for the breakdown of catecholamines including dopamine and serotonin (Chen et al., 2004; Tunbridge, Harrison & Weinberger, 2006; Yavich, Forsberg, Karayiorgou, Gogos & Mannisto, 2007). Polymorphisms in these genes have been associated with physiological differences that may be related to temperamental predispositions (Forbes, Brown, Kimak, Ferrell, Manuck, & Hariri, 2009; Munafo, Brown & Hariri, 2008; Mier, Kirsch & Meyer-Lindenberg, 2009). For instance, the short allele of 5-HTTLPR, a widely researched polymorphism in the serotonin transporter gene which has been previously associated with risk for depression under conditions of high environmental stress (Karg, Burmeister, Shedden & Sen, 2011; Uher & McGuffin, 2008; 2010), is also associated with greater amygdala volume, reactivity and resting state activation (Munafo et al., 2008), and disrupted functional connectivity between the amygdala, anterior cingulate and medial prefrontal cortex (Lemogne, Gorwood, Boni, Pessiglioni, Lehericy & Fossati, 2011; Pacheco, Beevers, Benavides, McGeary, Stice & Schnyer, 2009). In addition, the longer, 7-repeat allele of a 48 base pair variable number tandem repeat (VNTR) in Exon III of DRD4 has been related to higher ventral striatum reactivity in response to reward (Forbes et al., 2009), as well as lower startle reactivity and modulation of the startle reflex by affectively valenced pictures in adult males (Roussos, Giakoumaki, & Bitsios, 2009).

Some initial research has also found these candidate genes to be related to differences in child temperament and personality traits, including higher behavioral inhibition in children possessing at least one copy of the short allele of 5-HTTLPR (Battaglia et al., 2005; Fox et al, 2005; Hayden et al., 2007) and higher novelty seeking in

adolescents with a copy of the 7-repeat allele of the DRD4 Exon III VNTR (Laucht et al., 2007). Other studies have focused on interaction between genetic variants and environmental risk in predicting aspects of temperament such as irritability or persistence under stressful conditions (e.g. Ivorra, Sanjuan, Jover, Carot, de Frutos, & Molto, 2010; Amstadter et al., 2012). In one of the few studies with the sample size needed to test molecular gene by environment interaction, Hankin and colleagues (2011) report that 5-HTTLPR interacts with parenting to predict positive emotionality, such that carrying two copies of the short allele of 5-HTTLPR is associated with the highest self-reported positive emotionality in girls who experience supportive parenting, and the lowest positive emotionality in girls who experience unsupportive parenting, suggesting that particular genetic variants can be associated with either risk or resilience depending on the environmental circumstances.

Due to limitations in sample size and the lack of direct replication in the majority of molecular genetic studies of temperament to date, results of these studies do need to be interpreted cautiously. However, more consistent associations between candidate genes and putative biological endophenotypes such as amygdala reactivity (Munafo et al., 2008) and activation in the prefrontal cortex (Mier et al., 2009) suggest that the possibility of associations between these genes and temperament should be considered critically, but not dismissed out of hand.

Molecular genetic studies of gene-environment interaction have highlighted the possibility that certain polymorphisms may be associated with differential risk or resilience to environmental stressors (e.g. Hankin et al., 2011), whereas quantitative genetic studies suggest that the relative importance of genetic influences on risk and

resilience can depend on environmental conditions (e.g. Krueger et al., 2008). These two positions are not contradictory. Molecular genetic studies consider associations at the individual level, whereas quantitative genetic research is concerned with total genetic and environmental variance in the population, and offers little information about the influence of specific genes or environmental processes influencing on individual outcomes (Rutter et al., 2006). The combined picture offered by quantitative and molecular research suggests both that genetic differences are important in shaping temperament and personality and that the environment plays a role in influencing how genetic differences manifest (Rutter et al., 2006; Shanahan & Hofer, 2005).

Heritability of Caregiver-Report Temperament

Although caregiver-report studies of temperament generally yield moderate genetic variance, and often show negligible shared environmental variance (Gagne et al., 2009; Saudino, 2005), some dimensions of temperament do appear to be more heritable than others, and differential patterns of genetic and environmental variance are occasionally found within the subordinate traits. For instance, effortful control and its component dimensions of inhibitory control, persistence and attentional focusing tend to be moderately to highly heritable in toddlerhood and middle childhood (Gagne & Goldsmith, 2011; Goldsmith et al., 1997; Lemery-Chalfant, Doelger, & Goldsmith, 2008; Mullineaux et al., 2009; Wang, Deater-Deckard, Cutting, Thompson, & Petrill, 2012), with negligible common environmental variance reported in some studies (Goldsmith et al., 1997; Lemery-Chalfant et al., 2008; Mullineaux et al., 2009), although others have reported modest to moderate common environmental influence on inhibitory control and attentional focusing (Gagne & Saudino, 2010; Wang et al., 2012). In contrast, low

intensity pleasure, considered a dimension of effortful control (Rothbart et al., 2001), has been found to be best explained by both moderate heritability and moderate common environmental variance in middle childhood (Mullineaux et al., 2009).

The presence of common environmental influences on low intensity pleasure is consistent with infant and toddler studies showing that aspects of positive emotionality including broad positive affect, soothability, smiling and laughter, pleasure, adaptability and sociability often have substantial common environmental variance (Goldsmith & Campos, 1986; Goldsmith et al., 1997; 1999; Silberg et al., 2005), with heritability ranging from moderate for smiling and laughter, adaptability and sociability in infancy (Goldsmith et al., 1999; Silberg et al., 2005) to negligible for soothability in infancy (Goldsmith & Campos, 1986; Goldsmith et al., 1999) and pleasure in toddlerhood (Goldsmith et al., 1997). Positive emotion in middle childhood is split between two broad components of temperament, such that contentment and low intensity pleasure are aspects of effortful control, whereas exuberance, high intensity pleasure and extraversion are aspects of surgency (Rothbart et al., 2001; Rothbart & Bates, 2006). Although neither effortful control nor surgency shows evidence of common environmental variance when considered at a superordinate level (Goldsmith et al., 1997; Lemery-Chalfant et al., 2008; Mullineaux et al., 2009), the presence of significant common environmental influences when aspects of positive emotionality are considered separately suggests that examining genetic and environmental influences on only higher order components of temperament may overlook patterns of differential heritability across dimensions.

The heritability of surgency is also moderate at the broad component level (Goldsmith et al., 1997), with dimensions of impulsivity, activity level and approach all

found to have substantial genetic components (Bezdjian, Baker, & Tuvblad, 2011; Deater-Deckard et al., 2010; Saudino, 2009), with the remainder of the variance accounted for by unique environmental influences. Activity level has been found to be one of the most highly heritable aspects of temperament when assessed via caregiver-report on the IBQ and TBAQ (Goldsmith et al., 1997; Saudino & Eaton, 1991; Saudino, 2009), although one study does yield a more moderate estimate of heritability using the IBQ (Goldsmith et al., 1999). In addition, although the IBQ and the TBAQ appear to be relatively less affected by rater bias (Hwang & Rothbart, 2003), activity level in general appears to be one of the dimensions of temperament most likely to show evidence of contrast effects (Saudino & Eaton, 1991), and one study using a combined twin and family design has found that shared environmental influences on EAS activity level become evident when contrast effects are accounted for (Saudino, McGuire, Reiss, Hetherington, & Plomin, 1995).

Finally, caregiver-report of broad negative emotionality is moderately to highly heritable in infancy, toddlerhood and middle childhood, with estimates of additive genetic variance typically ranging from 40% to 70% and the remaining variance typically accounted for by unique environmental influences (Goldsmith et al., 1997; Goldsmith et al., 1999; Mullineaux et al., 2009; Singh & Waldman, 2010; Tackett, Waldman, Van Hulle, & Lahey, 2011), although one study reports evidence for moderate nonadditive genetic variance even after modeling twin contrast effects (Singh & Waldman, 2010). Findings at the subordinate level are similar for mother-report of CBQ fear, anger and sadness, with heritability within the range of 56% to 83% for fear, 71% to 75% for sadness, and 66% for anger (Goldsmith & Lemery, 2000; Mullineaux et al., 2009), and

distress to limitations and distress to novelty in infancy have also been found to show substantial and only modestly overlapping genetic variance (Goldsmith et al., 1999). However, one study finds notable discrepancies between mother and father-report of both broad negative emotionality and sadness in middle childhood, with mother-report yielding high estimates of genetic variance and no significant common environmental influences, whereas father-report negative emotionality and sadness are fully explained by high common environmental variance and unique environmental influence (Mullineaux et al., 2009). Such discrepancies may be a result of rater bias, but may also be explained by the very different social contexts under which mothers and fathers interact with children.

In addition, some studies have found moderate common environmental influences on caregiver-report anger according to the TBAQ, the Differential Emotions Scale (DES; Izard, 1972), and caregiver-report of fighting and angry outbursts in early childhood (Emde, Robinson, Corley, Nikkari, & Zahn-Waxler, 2001; Goldsmith et al., 1997). Goldsmith and colleagues (1997) also report that caregiver-report social fear in toddlerhood using the TBAQ is best accounted for by both moderate genetic and common environmental influences. Although common environmental influences are not always reported for anger (Deater-Deckard et al., 2010; Mullineaux et al., 2009), these findings are intriguing and bear further investigation, especially as there is evidence for significant common environmental influences on mother- and father-report of EAS emotionality in adolescents when twin contrast effects are accounted for (Ganiban et al., 2008; Saudino et al., 1995).

The presence of heritable influences on caregiver-report negative emotionality in infancy and childhood is well established, but research examining more specific dimensions is still needed, particularly in large samples using measures which are less vulnerable to twin contrast effects. In addition, there is still a need to extend quantitative genetic research of specific dimensions of negative emotionality beyond the univariate framework, as an understanding of genetic or environmental overlap between proneness to experience discrete negative emotions has the potential to inform research into the etiology of negative emotionality and related behavior problems, as well as the most appropriate level of analysis in future genetically informed research.

Heritability of Observed Temperament

Although quantitative genetic research of temperament in infancy and childhood has primarily relied on caregiver- and teacher-report measures, a number of studies have also used experimenter-ratings (e.g. the Infant Behavior Record scales; Bayley, 1969) and laboratory observations of temperament (e.g. the Lab-TAB). Twin studies have found evidence of genetic influence on multiple dimensions of temperament in infancy and early childhood using observational measures (Goldsmith et al., 1997), although not all findings are consistent across studies, and the small sample sizes and diverse measure of temperament used in many studies can make direct comparison of results difficult. Even with these limitations, however, the objectivity of observational measures and their potential to shed light on different facets of child temperament than those seen in the context of the parent-child relationship make these studies valuable both as a comparison to caregiver-report research and on their own merits.

Dimensions of temperament related to self-regulation, including task orientation, attention, and inhibitory control, have been examined across infancy and middle childhood (Deater-Deckard, Petrill, & Thompson, 2007; Gagne & Goldsmith, 2011; Gagne & Saudino, 2010; Goldsmith et al., 1997; Lemery-Chalfant et al., 2008), with the majority of prior research concentrating on experimenter ratings of task orientation or task persistence, a component of temperament encompassing persistence and sustained attention during testing situations (Goldsmith et al., 1997; Matheny, 1980). Task orientation shows evidence of low to moderate (16%-47%) but significant genetic variance across 14, 20 and 24 months in both twin and adoption samples (Braumgart, Plomin, DeFries, & Fulker, 1992; Saudino, Plomin & DeFries, 1996), and moderate to high heritability between 4 and 7 years (28%-83%; Deater-Deckard et al., 2007; Goldsmith & Gottesman, 1981; Lemery-Chalfant et al., 2008), with no significant common environmental influences at any age. However, with the exception of Lemery-Chalfant and colleagues' (2008) findings of high broad-sense heritability (83%) and modest unique environmental influences (17%) in 8 year old twins, the majority of variance in task persistence across both infancy and middle childhood is typically explained by unique environmental factors (50%-81%; Braumgart et al., 1992; Deater-Deckard et al., 2007; Goldsmith & Gottesman, 1981; Saudino, Plomin & DeFries, 1996).

In addition, inhibitory control has also been studied in toddlerhood and early childhood using structured laboratory observations of response inhibition, with conflicting results regarding the genetic and environmental underpinnings of observed inhibitory control (Gagne & Saudino, 2010; Gagne & Goldsmith, 2011). In one study of two year old twins, Gagne and Saudino (2010) report that observed inhibitory control is

best explained by unique environmental (62%) and additive genetic (38%) components. In contrast, a separate, larger study of three year old twins using two of the same three Lab-TAB episodes shows no evidence of significant genetic influences (Gagne & Goldsmith, 2011), with inhibitory control accounted for by both common (37%) and unique (63%) environmental variance. The reasons for this discrepancy are unclear, especially as inhibitory control is not yet widely studied, and early childhood is a time of considerable instability in temperament and development in neurobiological systems underlying executive attention and effortful control (Posner & Rothbart, 2009; Roberts & DelVeccio, 2000). However, despite the lack of genetic variance in inhibitory control reported by Gagne and Goldsmith (2011), the majority of observational studies to date do suggest the presence of heritable influences on at least some aspects of self-regulation (Deater-Deckard et al., 2007; Gagne & Saudino, 2010; Goldsmith et al., 1997; Lemery-Chalfant et al., 2008).

Observational studies have also found evidence for genetic influences on activity level, extraversion and behavioral approach, most commonly in infancy and early childhood using the IBR ratings (Goldsmith et al., 1997). There is not always a strict differentiation between aspects of temperament related to sociability and positive emotionality, and some observational measures such as IBR affect/extraversion encompass aspects of both these traits, although infant activity level is a separate component according to the IBR (Matheny, 1980). Experimenter ratings of affect/extraversion and person interest in infancy and active adjustment (a composite of activity level, sociability and adaption to novelty) at age 7 show moderate heritability (31%-66%; Braumgart et al., 1992; Goldsmith & Gottesman, 1981; Saudino et al., 1996),

with the remainder of the variance accounted for by unique environmental factors. In addition, although some early studies do not find significant genetic influences on IBR activity level in infancy (Matheny, 1980) or at age four (Goldsmith & Gottesman, 1981), later research suggests that activity level is modestly to moderately heritable in infancy according to IBR ratings (20%-57%; Braumgart et al., 1992; Goldsmith & Gottesman, 1981; Saudino et al., 1996) and moderately heritable in both home (32%) and laboratory (59%) contexts according to objective mechanical measures, with moderate common environmental influences (54%) also evident for mechanically measured activity in the home (Saudino, 2009).

In contrast, positive emotion, when assessed separately from extraversion, has been found to yield considerable estimates of common and unique environmental influences, with moderate (Roisman & Fraley, 2006) or negligible (Volbrecht, Lemery-Chalfant, Aksan, Zahn-Waxler, & Goldsmith, 2007) genetic variance. In a study using a composite of infants' positive and negative emotion assessed according to the IBR, Roisman and Fraley (2006) report significant common (23%) and unique (38%) environmental, as well as additive genetic (39%) factors, whereas caregiver-report of negative emotionality in the same study showed moderate heritability (46%) but no indication of common environmental variance. A second study assessing positive emotionality at 12 and 25 months according to an aggregate measure of two episodes of the Lab-TAB and caregiver-report on the IBQ and TBAQ finds that positive emotion across both ages is fully accounted for by common (56%) and unique (44%) environmental influences (Volbrecht et al., 2007). Although quantitative genetic research examining positive emotion is still limited, these findings are consistent with caregiver-

report research finding substantial common environmental and limited genetic variance in positive emotionality.

Unlike other aspects of temperament, behavioral inhibition, stranger fear and shyness have often been assessed not only with experimenter ratings but with laboratory observations of children's interactions with unfamiliar adults or peers (Goldsmith et al., 1997; McGuire, Clifford, Fink, Basho. & McDonnell, 2003). These studies have been relatively consistent in showing evidence of high to moderate heritability and no significant role for the shared environment (DiLalla, Kagan, & Reznick, 1994; Emde et al., 1992; McGuire et al., 2003; Robinson, Kagan, Reznick, & Corley, 1992), as well as some evidence for genetic influences on longitudinal stability and change (Matheny, 1989; Robinson et al., 1992). For instance, stranger fear at nine months shows high heritability according to both behavioral observations during a stranger approach episode (68%) and caregiver-report on the IBQ Distress to Novelty scale (58%), although the small sample size of the behavioral observation subsample suggests results should be interpreted cautiously (Goldsmith et al., 1999). In addition, behavioral inhibition has also been found to show high (70%) heritability at 24 months in an unfamiliar peer interaction (DiLalla et al., 1994), and moderate heritability across the ages of 14 to 24 months (42%-56%) in interactions with adult strangers (Emde et al., 1992; Robinson et al., 1992), with the heritability of individual facets of inhibition ranging from 30% for peer avoidance to 70% for latency to touch toys (DiLalla et al., 1994). Finally, in middle childhood, a combined twin and sibling study of observed behavioral inhibition in nonsocial, adult and peer contexts reports a broad-sense heritability of 59% for exploratory behavior across all contexts, with context-specific heritability estimates ranging from 51% for adult

interactions to 71% for peer interactions (McGuire et al., 2003). However, despite relatively high stability from early infancy to childhood (Rothbart & Bates, 2006), social inhibition in infancy and early childhood appears to be related to fearfulness or reactive aversion to novelty, whereas in middle childhood and adolescence behavioral inhibition may be more related to low extraversion or approach tendencies (Putnam et al., 2008). In the absence of longitudinal research from infancy to childhood, it is uncertain to what extent the same genetic factors underlie inhibition at each age.

Few studies to date have considered negative emotion using observational measures, and early research in infancy shows little evidence of significant genetic or shared environmental influences on negative or positive emotional scales during child testing (Emde et al., 1992), although Goldsmith and Gottesman (1981) do report evidence of genetic influences on fearfulness at age seven. In addition, reactivity to restraint at 24 months (Emde et al., 2001), experimenter ratings of irritability at age four and anger between the ages of four and seven (Deater-Deckard et al., 2007; Goldsmith & Gottesman, 1981) and laboratory observations of anger at 12 and 26 months (Gagne & Goldsmith, 2011) have yielded estimates of moderate heritability (25%-38%), with some studies also suggesting moderate common environmental influences (34-45%) at 20 and 24 months (Emde et al., 2001) and 36 months (23%; Gagne & Goldsmith, 2011).

Finally, one study finds that a latent factor of negative emotionality, composed of laboratory observations of frustration and experimenter ratings of negative hedonic tone, shows significant genetic influences at 14 (65%) and 20 months (40%), and significant common environmental influences at 20 (38%) and 24 months (51%), as well as nonshared environmental influences unique to each measurement occasion (Rhee et al.,

2012). Furthermore, shared genetic influences account for the moderate stability between 14 and 20 months, and shared common environmental influences account for the stability between 20 and 24 months, with a significant decrease in heritability from 14 to 24 months (Rhee et al., 2012). This pattern is unusual for quantitative genetic research, which has previously found the heritability of multiple dimensions of temperament to remain stable or increase with age (Saudino, 2005). However, the majority of prior work has been done with caregiver-report measures, which may show higher heritability due to both rater bias and differences in context such as parents' ability to take into account a wider range of children's behaviors over a longer period of time. In addition, the laboratory episodes used by Rhee and colleagues (2012) are designed to primarily assess aspects of anger and frustration, which has shown evidence of common environmental influences in both caregiver (Emde et al., 2001; Goldsmith et al., 1997) and observational (Emde et al., 2001; Gagne & Goldsmith, 2011) studies, and it may be that a negative emotionality factor that more directly taps sadness or fear would show a different pattern of genetic and environmental variance across the first two years.

Genetic and Environmental Covariance Between Observed Dimensions of Temperament

Some recent research has also used multivariate quantitative genetic models to examine the extent to which phenotypic covariation between related dimensions of temperament can be attributed to the same genetic or environmental influences. Although these bivariate and multivariate genetic designs are still relatively uncommon in the study of temperament, they can provide important information concerning the hierarchical

structure of temperament and personality (see Krueger et al., 2002, for an example with the externalizing spectrum).

In one study using a trivariate Cholesky decomposition to examine genetic and environmental influences shared between laboratory observations of anger and inhibitory control, Gagne and Goldsmith (2011) find that genetic factors play a role in anger/frustration at both 12 (38%) and 36 months (32%), although the longitudinal stability of laboratory-assessed anger is minimal; substantial unique environmental variance is also evident in anger at 12 (62%) and 36 (45%) months, and common environmental variance at 36 months (23%). In addition, the correlation between anger/frustration and inhibitory control at 36 months was explained by overlapping common and, to a lesser extent, unique environmental factors, with as much as 73% of the common environmental influences on observed anger at 36 months shared with observed inhibitory control (Gagne & Goldsmith, 2011). In contrast, parent-report anger at both time points and inhibitory control at 36 months all showed significant additive genetic variance, with 12 month anger and 36 month inhibitory control found to be particularly highly heritable; both additive genetic (56%) and unique environmental (34%) variance accounted for the overlap between anger and inhibitory control at 36 months, but no common environmental influence on either dimension of temperament was found (Gagne & Goldsmith, 2011).

Another twin study of attentional control and anger/frustration in a sample of four to eight year old twins also suggests that the correlation between laboratory-assessed anger and at least some facets of self-regulation may be largely environmental (Deater-Deckard et al., 2007). Unlike inhibitory control in early childhood (Gagne & Goldsmith,

2011), sustained attention during testing tasks in early to middle childhood was found to show significant genetic variance (28%; Deater-Deckard et al., 2007), consistent with research finding high heritability of observed attention in middle childhood (Lemery-Chalfant et al., 2008). However, Deater-Deckard and colleagues (2007) also found no genetic overlap between attention and anger/frustration. Significant genetic variance was found for anger (25%), but the correlation between anger and attentional focusing in this study was fully accounted for by nonshared environmental influences, although the authors note that the low sample size (N = 259 twin pairs, 105 MZ) may have limited their power to detect common environmental influences (Deater-Deckard et al., 2007). Despite conceptual and genetic differences between inhibitory control and attentional control, both studies suggest that although anger/frustration is phenotypically correlated with self-regulation, the two aspects of temperament are at least to some degree genetically distinct; proneness to anger in early childhood cannot be entirely attributed to genetically-influenced deficits in self-regulation, and many of the same family- and individual-level environmental influences are likely to underlie both heightened anger and lower inhibitory control (Deater-Deckard et al., 2007; Gagne & Goldsmith, 2011).

There is also some evidence from a study of 584 twins, assessed longitudinally from 12 to 25 months, that positive emotion and empathic helping are correlated for environmental reasons, whereas positive emotion and empathy-related hypothesis testing share minimal but significant genetic variance (Volbrecht et al., 2007). Like positive affect, Volbrecht and colleagues (2007) find that empathic helping is almost entirely accounted for by common (57%) and unique (43%) environmental influences. However, although the phenotypic correlation between empathic helping and positive affect was

significant, it was modest (r = .16), and only 3% of common and 2% of unique environmental variance in helping was shared with positive affect. In contrast, empathy-related hypothesis testing was best explained by additive genetic (62%) and unique environmental factors (38%), and the similarly modest (r = .12) correlation between positive emotion and hypothesis testing is accounted for by genetic influences (Volbrecht et al., 2007). The finding that positive emotion and empathy are largely genetically and environmentally distinct in this sample is interesting, as it emphasizes the fact that even in cases when two aspects of temperament are largely accounted for by common environmental influences at a particular age, such influences are not necessarily the same. In addition, evidence that highly related aspects of the same component of temperament, such as empathy-related helping and hypothesis testing, can show very different patterns of genetic and environmental variation emphasizes the utility of a fine-grained approach to temperament in behavior genetic research.

Genetic and Environmental Covariance Between Parent-Report and Observed Temperament

Multivariate quantitative genetic models have also been applied to decomposing the covariance between different measures of the same dimensions of temperament, allowing an estimation of the extent to which these measures tap the same genetic or environmental variance. Like genetically informed studies of the overlap between temperamental dimensions, these studies are rare, but have the potential to be informative about the role of method and context in the measurement of temperament.

In a study of inhibitory control in two year old twins, Gagne and Saudino (2010) examined the correlation between caregiver-report of TBAQ inhibitory control and

laboratory observation of simple response inhibition across three episodes of the lab-TAB, and found evidence for both genetic covariance between measures and genetic and environmental variance unique to each. Specifically, caregiver-report inhibitory control was found to have additive genetic (58%), common environmental (26%) and nonshared environmental (16%) variance, whereas observed inhibitory control was best explained by nonshared environmental (62%) and additive genetic (38%) components (Gagne & Saudino, 2010). Forty-seven percent of the genetic influences on inhibitory control were shared with observed inhibitory control, and this shared genetic factor entirely accounted for the correlation between the two measures (Gagne & Saudino, 2010). Such findings provide support for inhibitory control as an aspect of temperament that is not only heritable but consistent across situations and methods of assessment, and suggest that both caregiver-report and laboratory observation are capable of tapping the same underlying individual differences, although the notable genetic variance unique to each measure also points to the importance of differences in method and context.

In addition, in a study examining activity level in two-year-old twins, as measured by caregiver-report on the TBAQ, observer-report using the IBR, and actigraphs at home and in the laboratory, Saudino (2009) reports both significant genetic covariance between all measures and measure-specific genetic variance that cannot be accounted for by differences in eliciting context. Although genetic covariance between caregiver-report and actigraph-measured activity level in the home was substantial (38%), and fully accounted for the modest phenotypic correlation between the two methods *(r* = .25), genetic variance in caregiver-reports and actigraphs was still primarily measure-specific. In contrast, genetic influences on observer ratings were almost entirely shared with

actigraph-measured activity level in the laboratory (.95), and this shared genetic variance accounted for 72% of the phenotypic correlation (r = .67), with the remaining 28% attributable to unique environmental factors (Saudino, 2009). In a separate sample of 463 twin pairs (150 MZ) between the ages of seven and nine, Wood, Rijsdijk, Saudino, Asherson, and Kuntsi (2008) examined activity level according to actigraph measurement and parent and teacher ratings of hyperactivity-impulsivity. A common latent factor explained a small but significant portion of the variance in each measure, and was almost entirely accounted for by genetic factors (92%), suggesting that genetic influences are particularly important for consistency in activity level across measures (Wood et al., 2008). However, each measure was also explained by unique genetic (18%-38%) and, in the case of actigraph measurements, common environmental (37%) factors, again underlining the ability of multimethod research to tap into distinct, systematic aspects of child behavior not picked up by all superordinate components.

Conclusion

Although many quantitative genetic studies of observed temperament do report substantial additive genetic variance, heritability estimates tend to be lower than those with caregiver-report data, and in many cases the majority of variance in temperament is explained by unique environmental factors (e.g. Braumgart et al., 1992; Deater-Deckard et al., 2007; Gagne & Goldsmith, 2011; Gagne & Saudino, 2010; Saudino, Plomin & DeFries, 1996). These differences likely reflect a number of factors. Research using experimenter-ratings and standardized laboratory observation is not vulnerable to the inflation of heritability estimates caused by artificially low or negative DZ twin correlations (Saudino, 2005), but caregiver-report measures have tended to yield higher

estimates of heritability even in studies showing no evidence of such contrast effects (e.g. Gagne & Goldsmith, 2011; Mullineaux et al., 2009), suggesting that the greater genetic variance often found by caregiver-report studies is not only an artifact of twin contrasts. In addition, measures of observed behavior are by their nature more susceptible to the influence of situation-specific factors such as tester differences and daily events, contributing to higher estimates of unique environmental variance. However, this does not mean the large estimates of unique environmental variance reported by many observational studies should be dismissed as error or inconsistency; although additive genetic factors largely account for consistency of temperament and behavior over time (e.g. Deater-Deckard et al., 2007), human behavior takes place in the context of the same day-to-day environmental events likely to underlie the greater twin differences in observational studies, and differences in reactivity to such events is itself an important component of many dimensions of temperament (Rothbart & Bates, 2006).

The differences in estimates of genetic and environmental variance underlying the same aspect of temperament (e.g. both attentional control and inhibitory control; Deater-Deckard et al., 2007; Gagne & Goldsmith, 2011; Gagne & Saudino, 2010; Lemery-Chalfant et al., 2008) from study to study also illustrate the principle that heritability, as an index of twin differences and similarities, is specific to the population in question and might well be affected by sample characteristics, the level of measurement used, the range of behaviors recorded, and the particular construct examined (Lemery-Chalfant, 2010). As such, patterns of genetic and environmental overlap between aspects of temperament are more informative than precise estimates of heritability. For instance, Saudino and colleagues' (2009) work with activity level suggests that in addition to

method effects, different measures of the same dimension of temperament may tap genuinely different, and genetically distinct, aspects of child behavior. However, quantitative genetic research examining covariance between caregiver-report and observational measures of other aspects of temperament is still limited, and activity level is not necessarily generalizable to traits in affective or regulatory domains.

Although negative emotionality has been among the most studied aspects of temperament in the phenotypic literature, no quantitative genetic study to date has examined laboratory observations of negative emotionality in middle childhood, or considered genetic and environmental influences on the covariance between the subordinate traits that make up negative emotionality. In addition, middle childhood has been understudied in temperament research, particularly genetically-informed temperament research, which has concentrated on infancy and early childhood, perhaps in part because of the emphasis on early-emerging hereditary traits. However, middle childhood is a time when temperament becomes increasingly stable (Roberts & DelVeccio, 2000), and children face an increasing range of environmental challenges outside the home (e.g. in academic settings and peer groups), which may both influence and be influenced by child temperament (Rothbart & Bates, 2006).

The Current Study

The aim of this study was to examine the genetic and environmental influences on three subordinate-level aspects of negative emotionality (anger, sadness and fear) in middle childhood, using both observational and caregiver-report measures. The first major goal was simply to use univariate quantitative genetic model fitting to describe the heritability of anger, sadness and fear in middle childhood, with univariate

models fit separately for mother-report, father-report, and in-home observations of each emotion. Secondly, I intended to examine the genetic and environmental influences on the covariance between mother-report, father-report, and in-home observation for each emotion, using multivariate Cholesky decompositions to parse the variance shared across reporters into genetic and environmental components. Finally, I intended to use multivariate Cholesky decompositions to describe the genetic and environmental covariance between anger, sadness and fear, and to assess the extent to which these emotions might be accounted for by a single common set of latent genetic and environmental influences, with separate models fit for mother-report, father-report and in-home observation.

Hypotheses

1) The first aim of the current study was to examine the phenotypic convergence between mother-report, father-report, and in-home observation of proneness to three discrete negative emotions (fear, anger and sadness) in middle childhood, and to examine the heritability of fear, anger and sadness assessed according to each type of measurement. For this aim, I hypothesized:
    1. Consistent with prior literature, mother- and father-report of each dimension of negative emotionality would show moderate convergence, but convergence between caregiver-report and observational measures would be modest, with the lowest convergence expected for sadness and the highest for fear across all measures.
    2. Mother-report of sadness, anger and fear, and father-report of anger and fear, would all be moderately to highly heritable, whereas in-home observations are

expected to be primarily explained by unique environmental influences, with modest to moderate genetic variance. Consistent with limited prior research (Mullineaux et al., 2009), father-report of sadness is expected to be primarily explained by common environmental influences, with only limited genetic variance.

3. Across all reporters, fear is expected to be the most highly heritable dimension of temperament, and anger is expected to show modest but significant common environmental variance.

2) The second aim of the current project was to examine the extent to which the covariance between mother-report, father-report and in-home observation of fear, anger and sadness can be explained by overlapping genetic or environmental influences. This aim is largely exploratory, with the goal of characterizing the genetic and environmental relations between measures without making strong predictions. However, for this aim, I hypothesized:

1. For all dimensions of negative emotionality, covariance between reporters would be largely accounted for by shared genetic variance, although anger is also expected to show significant common environmental covariance across reporters. Although the majority of variance in father-report of sadness is expected to be explained by common environmental factors, this variance is not expected to overlap with other measures.
2. Consistent with the idea that mother-report and father-report are expected to capture distinct aspects of child temperament, mother-report and father-report are expected to share genetic variance with in-home observations which is

unique to that parent, rather than shared across all reporters.

3) The third aim of the current project was to examine whether or not proneness to the discrete negative emotions of sadness, anger and fear in middle childhood can be accounted for by a single, common genetically- and environmentally-influenced negative emotionality factor, and if so, the extent to which each negative emotion is represented by this factor. For this aim, mother-report, father-report, and in-home observation were considered separately. I did expect that it would be possible to fit a common negative emotionality factor, which was expected to be highly heritable. However, I did not expect that this common factor would be sufficient to account for sadness, anger and fear for two reasons:
    1. Across all types of measurement (mother-report, father-report and in-home observation), sadness, anger and fear were expected to display substantial genetic and, in some cases, common environmental variance which could not be accounted for by a common factor.
    2. Genetic and environmental relations among variables were expected to be more complex than a single factor could represent (for example, sadness and fear might share genetic variance which is not shared with anger), and thus, a less restrictive model would be required to provide a good fit to the data.

Chapter 2

METHOD

Participants

Participants in this study were drawn from the Wisconsin Twin Project (WTP), an ongoing longitudinal study of temperament, emotion and psychopathology following twins from toddlerhood through adolescence (Lemery-Chalfant et al., 2006). Twins were recruited from hospital birth records in the state of Wisconsin between the years of 1989 and 2004. At Phase 3 of the WTP, when twins were approximately eight years old (*M* = 7.87 years, *SD* = .93), 1866 twins (32.2% MZ, 33.3% same sex DZ, 34.5% opposite sex DZ; 50% male) were selected for a follow up assessment. The sample at Phase 3 is primarily Caucasian (85.8%), with the remaining participants identifying as African American (6.2%), Native American (1.5%), Asian (.2%), biracial (1.8%) and other (3.5%), which is representative of the population of Wisconsin. The total family income ranges from below $10,000 per year to over $200,000 per year, with the mean family income between $50,001 and $60,000 per year. The level of education ranges from 6 years of formal education to a graduate degree for both mothers and fathers, with the majority of participants reporting either some college, trade or technical school education or a college degree (*M* = 14.88 years of formal education for mothers and 14.43 years of formal education for fathers).

Procedure

When the twins were approximately eight years of age, families took part in the Phase 3 follow-up assessment, which included two phone interviews, mailed questionnaire packets and a 4-hour home visit (Lemery-Chalfant et al., 2006). As part of

the first phone interview, both mothers and fathers reported on twins' temperament using the CBQ, along with demographic information and multiple measures of child behavior, symptoms of psychopathology, parenting and home environment.

As part of the home visit, twins took part in sixteen videotaped and coded episodes of the middle childhood version of the Laboratory Temperament Assessment Battery (Lab-TAB; Goldsmith, Reilly, Lemery, Longley & Prescott, 2001) designed to assess temperament dimensions including activity level, anger, sadness, contentment, exuberance, persistence, impulsivity, object fear, and social fear/inhibition. Twins were tested in the home, and twins were tested separately in different rooms of the house. Episodes were administered in a different order for each twin, but for each episode, the same child tester worked with both twins, and the same episode was administered in the same room of the house for each twin. Time was provided between episodes for twins to rest and return to a neutral baseline state. In addition, the sequence of episodes was designed to avoid running potentially stressful episodes consecutively, in order to minimize carryover effects and child fatigue or distress (Goldsmith et al., 2001).

Measures

Zygosity. Zygosity was assessed through mother-report on the Zygosity Questionnaire for Young Twins (Goldsmith, 1991) when the twins were 2 to 3 years old (Lemery-Chalfant et al., 2006). This questionnaire is a detailed 32-item caregiver-report instrument assessing similarities and differences between twins (e.g differences in hair color, texture and shade). The Zygosity Questionnaire for Young Twins demonstrates high (over 95%) agreement with zygosity assessed by genotyping (Forget-Dubois et al., 2003; Price, Freeman, Craig, Petrill, Ebersole, & Plomin, 2000).

Caregiver-Report Negative Emotion. Negative emotionality was assessed using mother-report and father-report on the Children's Behavior Questionnaire (CBQ; Rothbart et al., 2001), a 180-item caregiver-report questionnaire designed for children between the ages of four and seven. The CBQ contains scales assessing 15 lower-order dimensions of child temperament within the past six months on a 1-7 Likert scale, with 1 being "extremely untrue of your child" and 7 being "extremely true of your child" (Rothbart et al., 2001). In previous studies, the CBQ has been found to show good to adequate internal consistency (ranging from .67 to .94, *M* = .77), high longitudinal stability, and good construct and convergent validity (Ahadi, Rothbart, & Ye, 2003; Rothbart et al., 2001; Rothbart et al., 2000).

In the current study, 10-item scales assessing object fear, shyness, anger/frustration and sadness were used. Sample questions for these scales include "is afraid of loud noises," "sometimes seems nervous when talking to adults s/he has just met," "gets angry when told s/he has to go to bed," and "cries sadly when a favorite toy gets lost or broken", respectively. In addition, a higher-order mean composite of negative emotionality factor was created from the fear, anger and sadness scales. In the current study, there was good to adequate internal consistency at the subscale level for both mother-report of anger (alpha = .827), sadness (alpha = .693), fear (alpha = .746), and shyness (alpha = .893), and father-report of anger (alpha = .818), sadness (alpha = .674), fear (alpha = .753), and shyness (alpha = .884).

Observed Negative Emotion. Observed negative emotionality was assessed using a home-based version of the Lab-TAB designed for use with children between the ages of six and eight (Goldsmith et al., 2001). The Lab-TAB is a standardized assessment

of temperament consisting of sixteen episodes intended to tap discrete emotional reactions, as well as other aspects of temperament including impulsivity, regulation and activity level, under naturalistic conditions. Episodes typically range from three to ten minutes in length, and are divided into coding epochs which range from five to thirty seconds, depending on the episode. The intensity of children's facial, vocal, and postural displays of emotion is coded in each epoch, along with other relevant behaviors (e.g. intensity of approach and avoidance, duration of gaze aversion). A similar preschool version of the Lab-TAB has been validated in both laboratory and home settings (Gagne, Van Hulle, Aksan, Essex, & Goldsmith, 2011), and demonstrates moderate convergence between Lab-TAB composites and observer post-visit ratings (correlations ranging from .21 to .76; Gagne et al., 2011), although correlations with parent-report of child temperament on the CBQ are more modest.

Selected Episodes. In the current study, four Lab-TAB episodes which regularly evoke both sadness and anger responses from children, 'I'm Not Sharing,' 'Impossibly Perfect Stars,' 'Transparent Box,' and 'Wrong Gift', were used to assess both observed anger and sadness. Fear was assessed using two episodes, 'Scary Mask,' which targets object fear, and 'Storytelling,' which targets social fear. It should be noted that multiple variables (e.g., presence of overall negativity and positivity) are coded across all episodes, but only variables that specifically measure discrete anger, sadness or fear are considered in the current study.

Not Sharing. Not Sharing is an episode designed to evoke feelings of sadness and anger in response to unfair treatment, and consists of a child tester unequally sharing a bag of candy with the child, with progressively greater inequality as the episode

continues. This episode is of variable length, depending on the child and the experimenter, and ends when all the candy has been distributed. Coding in Not Sharing takes place in 10 second epochs, and coded variables specific to discrete sadness and anger include latency to first clear anger and first clear sadness in seconds, intensity of combined bodily and expressed anger and combined bodily and expressed sadness in each epoch on a scale of 0-3, and intensity of anger vocalizations and sadness vocalizations on a scale of 0-3.

***Impossibly Perfect Stars*.** Impossibly Perfect Stars is a 4 minute episode intended to evoke feelings of anger/frustration and sadness in response to criticism and being asked to complete a repetitive and impossible task. During this episode, the child is asked to repeatedly draw stars by a child tester, who responds by critiquing each star and asking the child to try again. Coding in Impossibly Perfect Stars takes place in 10 second epochs, and variables specific to discrete sadness and anger include latency to first clear anger and first clear sadness in seconds, intensity of bodily anger and bodily sadness in each epoch on a scale of 0-2, intensity of expressed anger and expressed sadness in each epoch on a scale of 0-2, and intensity of anger and sadness vocalizations on a scale of 0-2.

***Transparent Box*.** Transparent Box is a four-minute episode designed to evoke feelings of anger/frustration, in which the child is presented with several appealing toys in a locked transparent box, along with a set of keys which cannot open the box. Coding in Transparent Box takes place in 10 second epochs, and variables specific to discrete anger and sadness include latency to first clear anger and first clear sadness in seconds,

intensity of bodily anger and bodily sadness in each epoch on a scale of 0-2, and intensity of expressed anger and expressed sadness in each epoch on a scale of 0-2.

Wrong Gift. Wrong Gift is an episode targeted toward feelings of sadness, disappointment, frustration and ability to regulate negative affect. In this episode, instead of a preferred gift, the child is given a box containing an unappealing gift that the child has previously ranked as their least favorite. This episode is of variable length but no shorter than two and a half minutes, beginning when the child is told he or she will receive a prize and ending after the child is allowed to pick the correct gift. Coding in Wrong Gift takes place in 5 second epochs, and variables specific to discrete sadness and anger include latency to first clear anger and first clear sadness in seconds, intensity of bodily anger and bodily sadness in each epoch on a scale of 0-2, and intensity of expressed anger and expressed sadness in each epoch on a scale of 0-2.

Scary Mask. During the Scary Mask episode, the child interacts with a friendly stranger wearing a frightening mask. The episode is divided into six epochs of 15 seconds each, beginning when the child first notices the stranger's face and continuing as the stranger takes off the mask, begins a conversation with the child, and finally asks the child to touch and wear the mask. Variables coded in Scary Mask include latency to first fear response in seconds, intensity of expressed fear and bodily fear in each epoch on a scale of 0-3, intensity of approach and avoidance in each epoch on a scale of 0-3, intensity of vocal fear expression on a scale of 0-2, duration of gaze aversion, and presence of startle.

Storytelling. During the storytelling episode, the child is asked to give a speech about the previous day's events in front of an audience of multiple child testers. The

episode is divided into 10 second epochs, but may range from one to over twelve minutes, depending on the child, and always includes at least one prompt by the child tester (e.g. 'is there anything else you would like to tell me?'). Variables coded in Storytelling include latency to first fear response in seconds, latency to begin speaking, intensity of expressed fear and bodily fear in each epoch on a scale of 0-3, intensity of avoidance in each epoch on a scale of 0-3, percent of time speaking, presence of partially voiced or whispered speech, and number of disfluencies or hesitations.

For each episode, there were five or six trained coders. Tapes were coded by a single coder, with 10% of tapes double-coded by a master coder. Kappa values in the fearfulness episodes ranged from .79 to .89, and kappa values for sadness variables ranged from .70 to .93. However, anger variables tended to have lower inter-rater reliability, with kappa values ranging from .63 and .69 for expressed anger in the 'Not Sharing' and 'Impossibly Perfect Stars' episodes, respectively, to .84 for bodily anger in 'Transparent Box.'

Lab-TAB Composite Formation.

Lower-order Composites. The first step was the formation of lower-order composites from raw data in accordance with the Lab-TAB manual (Goldsmith et al., 2001) and guidelines suggested by Gagne and colleagues (2011). Coded variables used in the formation of lower-order composites included latency scores to first clear anger, sadness or fear response, and scores measuring the intensity of discrete anger, sadness or fear in each epoch (see Table 1 for a summary of variables used to form lower-order composites in each episode). Typically, intensity of expressed, bodily, and vocal emotion was coded separately, but in Not Sharing, intensity of expressed and bodily emotion were

combined into a single variable. Unless otherwise noted, the process of lower-order composite formation was the same for each episode, although the number of lower-order composites formed differed based on the variables coded in each episode.

Latency scores were windsorized to three standard deviations outside the mean, and then reverse-coded to transform them to speed scores. For two episodes, Transparent Box and Wrong Gift, no latency scores were originally coded. For Wrong Gift, latency to first response is less meaningful because there is a fixed point where a negative response is elicited (the child first sees the disappointing gift). However, for Transparent Box, rough latency scores for anger and sadness were computed as the time in seconds until halfway through the earliest epoch in which a child showed a facial or bodily emotional response.

After latency scores were transformed to speed scores, the next step was to form episode-level mean intensity composites from children's anger, sadness, or fear responses coded in each epoch. For each emotion, separate mean intensity composites were created from expressed, bodily and vocal parameters whenever these parameters were available. Peak intensity scores were also considered for expressed, bodily and vocal anger, sadness and fear, but these scores were not used in later composite formation due to lower variability and lower intercorrelations among theoretically-related variables (e.g., among variables tapping speed and intensity of anger) both within and across-episode**.** Sum scores were not considered due to variable episode length across participants for Not Sharing and Storytelling, and because it was important not to conflate epochs in which scores could not be coded on a parameter with epochs in which no response was shown.

After speed and mean intensity scores were computed for all lower-order parameters in an episode, these speed and mean intensity composites were checked for

low variability and deviations from normality using descriptive statistics. Three parameters were identified which showed insufficient variability due to very low frequency of responses above zero. In Storytelling, 1,009 participants out of a total of 1032 showed a mean expressed fear of 0 across all epochs. In Transparent Box, 521 participants out of 743 showed a mean expressed sadness of 0 across all epochs, and 551 showed a mean bodily sadness of 0, with a further 80 participants showing a mean expressed sadness of .04 (on a 0-3 coding scale) and 90 showing a mean bodily sadness of .04. These parameters were not considered further in analyses. Several mean intensity scores exceeded the recommended cutoff of 2.00 for skewness or 7.00 for kurtosis (Muthén & Kaplan, 1985; parameters that show deviations from normality are listed in Table 1). However, skewed or kurtotic parameters were not transformed until after coder differences were tested and mean intensity scores were adjusted for coder differences.

Rescaling and Transforming Parameters. After computing mean intensity scores for expressed, bodily and vocal anger, sadness and fear in each episode, these parameters were tested for mean differences between coders, in order to determine whether or not any coder was consistently rating children's intensity of anger, sadness or fear as significantly higher or lower than other coders. Between-coder differences in mean intensity scores were examined using one-factor ANOVAs with coder number as the predictor variable and mean intensity of each response as the dependent variable. For example, if any coder was systematically over-rating the intensity of children's expressed anger relative to other coders, then the mean of children's expressed anger across tapes coded by that coder would be expected to be significantly higher than the mean expressed anger across tapes coded by other coders.

When mean differences between coders were found, the scores of coders who showed significant mean differences were rescaled to the metric of a referent coder using a series of steps. First, the coder who coded the most tapes was selected as the referent coder, unless the coder who coded the most tapes also showed limited variance (standard deviation of less than .10). If the coder with the highest number of tapes showed low variance, then the coder with the second highest number of tapes was selected. If there were two coders with the same number of tapes, the one with the highest variance was chosen. After the referent coder was selected, the scores of the other coders were standardized using a z-score transformation. These standardized scores were then rescaled back onto the raw scores of the referent coder. In order to rescale standardized scores back onto the metric of the referent coder, standardized scores were first multiplied by the raw standard deviation of the referent coder, then summed with the raw mean of the referent coder (Lemery et al., 1999). Similar to standardizing by using z-scores, this method of rescaling preserves the properties of the data while making scores more comparable across coders.

Higher-order Within-episode Composites. After parameters were rescaled to adjust for coder differences, the next step was to form episode-level mean composites for each emotion from lower order speed and expressed, bodily and vocal parameters. First, descriptive statistics were used to examine deviations from normality in the rescaled parameters, and square-root transformations were used for parameters that exceeded recommended cutoffs for skewness and kurtosis (2.00 and 7.00, respectively; Muthén & Kaplan, 1985).

After applying square-root transformations, zero-order correlations among lower-order speed and intensity parameters were used to examine whether variables were highly correlated enough to allow the formation of mean composites. Final mean composites for anger, sadness and fear within each episode are summarized in Table 1. All lower-order speed and intensity parameters were standardized using z-score transformations prior to higher-order composite formation. In general, zero-order correlations revealed that within each episode, lower-order parameters assessing a single emotion (e.g., speed to first anger, mean intensity of bodily anger, and mean intensity of expressed anger) were moderately to highly positively correlated with each other (*r* ranging from .288 to .698), with the exception of speed and intensity of vocal anger and sadness in Stars and Not Sharing, which were relatively less related to other parameters (*r* ranging from .154 to .486).

For Stars, correlations between bodily, expressed and speed parameters ranged from .288 to .551 for Stars anger. Vocal anger was relatively less correlated with expressed (.162) and speed (.217), but moderately correlated with bodily anger (.486), and so a final mean composite was formed from all four parameters (speed, intensity of expressed anger, intensity of bodily anger, and intensity of vocal anger). Correlations among sadness parameters were similar in Stars, with correlations ranging from .203 to .698 among speed, expressed, bodily and vocal sadness, and the lowest correlations again evident for vocal sadness. Thus, the final mean sadness composite Stars included speed, anger, bodily and vocal parameters. For Not Sharing, anger parameters were also moderately to highly intercorrelated (.246-.563), with the exception of a high correlation between speed to vocal anger and intensity of vocal anger (.887), and a final mean

composite was formed from speed to first expressed/bodily anger, speed to first vocal anger, mean intensity of combined expressed/bodily anger, and mean intensity of vocal anger. However, for sadness in Not Sharing, vocal parameters were not highly correlated with speed or mean intensity of combined expressed/bodily sadness (.134-.196), and so the final sadness composite in Not Sharing was formed from and mean intensity of combined expressed/bodily sadness, which were moderately correlated with each other (.492). In addition, for Transparent Box, all sadness parameters had too little variability to be considered in further analyses, leaving a total of three parameters assessing anger: mean intensity of bodily anger, mean intensity of expressed anger, and speed. These parameters were moderately to highly correlated (.434-.677), supporting the formation of a mean anger composite. Finally, in Wrong Gift, only two parameters (bodily intensity and expressed intensity) were coded for anger and sadness. Bodily and expressed intensity were highly correlated within-emotion (r = .604 for anger and .614 for sadness), and were used to form mean anger and sadness composites. For Scary mask, a mean fear composite was formed from speed, expressed and bodily fear, all of which were moderately correlated (.340-.366). However, for Storytelling, only two indicators of fear had sufficient variability (speed and bodily intensity); these parameters had a moderate correlation (.420), and were combined into a mean composite of social fear.

Higher-order Cross-episode Composites. After creating higher-order within-episode composites for anger, sadness and fear, zero-order correlations between final within-episode emotion composites were run to examine whether episode-level anger, sadness and fear composites could be aggregated into cross-episode anger, sadness and fear composites. All higher-order across-episode composites are summarized in Table 2,

and correlations between the final within-episode composites are reported in Table 3. In general, observed emotion composites were not strongly related at the cross-episode level. For anger, three episodes (Impossibly Perfect Stars, Not Sharing, and Transparent Box) showed moderate cross-episode correlations (.241-.311). For sadness, cross-episode correlations were lower, with Not Sharing and Impossibly Perfect Stars correlated by .174, and correlations between Wrong Gift sadness and the other sadness composites were only modest (.067-.07). Because capturing variance in sadness shared across episodes was important, it was decided to aggregate scores on Not Sharing and Impossibly Perfect Stars even though the correlation was relatively low, but Wrong Gift was not strongly related enough to the other variables to aggregate across episode.

Thus, episode-level anger composites for Impossibly Perfect Stars, Not Sharing, and Transparent Box, and sadness composites for Impossibly Perfect Stars and Not Sharing, were standardized using z-score transformations, and used to compute cross-episode mean composites of anger and sadness respectively. Social fear assessed in Storytelling and object fear assessed in Scary Mask were moderately correlated, (r = .262), but these two episodes were kept separate, both on the basis of conceptual differences between social and object fear and because these two composites showed a different pattern of zero-order correlations with other observed composites and parent-report data (see Tables 6 and 7).

CBQ Negative Emotionality Composites. For parent-report data, anger, sadness and fear were all moderately-to-highly positively correlated within reporter (.205 – .580), and shyness was moderately correlated with fear (.226 and .240 for mother- and father-report) but only modestly correlated with anger (r_M = .098; r_F = .062) and sadness (r_M =

.149; r_F = .140). Mean composites for mother-report and father-report negative emotionality were formed from anger, sadness and fear, but shyness was not included, both because of relatively lower correlations and because prior research suggests that shyness is more a facet of surgency/extraversion than negative emotionality in middle childhood (Rothbart et al., 2001).

Data Analysis Plan

Overview. Quantitative genetic research uses structural equation modeling-based path analysis and model-fitting techniques to decompose the total phenotypic variance in a trait into latent additive genetic (A), non-additive genetic (epistasis and dominance; D), common environmental (C) and unique environmental (E) components. Conceptually, the A component reflects the average effect of all individual genes across the genotype, whereas the D component accounts for interaction between alleles at the same (dominance) or different (epistasis) loci. Together, A and D encompass all influences that lead to greater phenotypic resemblance between individuals who are more closely genetically related. For instance, because monozygotic (MZ) twins are genetically identical, whereas dizygotic (DZ) twins share 50% of their segregating genes on average, if 100% of the resemblance in a trait was accounted for by additive genetic influences then phenotypic correlations between MZ twins would be expected to be approximately twice as high as phenotypic correlations between DZ twins. On the other hand, MZ twin correlations higher than twice DZ correlations suggest the action of non-additive genetic influences (Neale & Cardon, 1992), or, in some cases, rater bias (Saudino, 2003).

In contrast, the C component encompasses all influences which increase the phenotypic similarity of individuals reared in the same household, independent of their

degree of genetic relatedness. DZ twin correlations higher than half of MZ twin correlations suggest the action of common environmental factors, and if 100% of the phenotypic resemblance in a trait were due to the influence of the common environment, MZ and DZ twin correlations would be expected to be approximately equal. Finally, all non-genetic factors that serve to make individuals different from one another, including context effects and measurement error, are attributed to the unique or nonshared environmental component. For instance, all differences between MZ twins reared in the same household can be attributed to the unique environment.

Analyses for Aim 1: Univariate ACE Models. The classic univariate biometrical model, the ACE model, is constructed on the basis of the above logic. Figure 1 represents the ACE model for two twins reared in the same household. The latent A, C and E components in this model are estimated from the observed variances and covariances of MZ and DZ twins' scores on a single phenotype of interest, in conjunction with the genetic and environmental correlations between twins. The ACE model is a multigroup structural equation model, with covariances modeled differently for MZ and DZ twins on the basis of different degrees of genetic relatedness. Specifically, because MZ twins share 100% of their genes, whereas DZ twins share 50% of their additive genetic influences, the correlation between latent additive genetic components is fixed to 1.0 for MZ twins, and .5 for DZ twins. The common environment is fully shared between cotwins, and therefore, the correlation between latent C components is fixed to 1.0 for both MZ and DZ twins. Because of model identification limitations with only twin data, it is not possible to estimate both C and D components within the same model. However, after the ACE model has been fit, it is also possible to fit an alternate ADE model, which accounts

for nonadditive genetic effects. In the ADE model, the additive genetic correlation is fixed to 1.0 and .5 for MZ and DZ twins, respectively, and the dominant genetic correlation is fixed to 1.0 and .25 for MZ and DZ twins, respectively, because DZ twins inherit the same alleles at a locus 25% of the time (Neale & Cardon, 1992). The E component, which by definition encompasses influences which are unique to each twin, is uncorrelated for both MZ and DZ twins. Finally, because the magnitude of genetic and environmental effects on a trait is not expected to differ between cotwins assigned at random, the a, c and e pathways are all fixed to be equal across cotwins.

In the current study, the first step was to fit univariate ACE or ADE models to estimate the genetic and environmental influences on sadness, anger, and fear. This was done separately for mother-report, father-report, and observed sadness, anger, fear. All models were fit using the statistical program OpenMX (Boker et al., 2011), an R-based program which uses Maximum Likelihood estimation procedures. After the full ACE or ADE models were fit, nested models were tested by systematically dropping parameters, and the fit of the reduced models was compared to that of the full model in order to find the most parsimonious solution. Model fit was assessed according to chi-square difference test and Akaike's Information Criterion (AIC; Akaike, 1987). The AIC is a fit index which is suitable for the comparison of non-nested models (Williams & Holahan, 1994). Smaller AIC values indicate a model with more support, with proposed guidelines for interpretation suggesting that a change in AIC of less than 2.00 from the best-fitting model provides considerable support for the more restricted model, a change of between 4.00 and 7.00 provides lower support, and a change in AIC of more than 10.00 does not support the reduced model (Burnham & Anderson, 2004). A nonsignificant difference in

fit implies that the reduced model accounts for the observed data as well as the full model, whereas a significant decrement in fit indicates that the dropped parameter was necessary to adequately reproduce the observed data. However, because the E component contains measurement error, it is always retained in the model, and D is never estimated without A because it is unlikely that all genetic influences are interactive, with no additive effects.

Analyses for Aim 2: Multivariate ACE Models. In the same way that univariate models decompose the variance in a single trait into genetic and environmental components, bivariate Cholesky decompositions can be used to decompose the covariance between two traits into latent A, C (or D), and E components. This is done by considering the extent to which one twin's score on the first trait covaries with the other twin's score on a second trait as a function of genetic relatedness. Conceptually, if proneness to sadness and proneness to anger are influenced by the same set of genes, one twin's anger score should be related to the other twin's score on sadness, and the strength of this relation should be higher for MZ than DZ twin pairs. Figure 2 shows the multivariate Cholesky decomposition, which allows an estimation of the extent to which latent genetic and environmental factors influencing one phenotype are shared with one or more other phenotypes, regardless of the heritability of each phenotype. In this model, the first set of latent factors (A1, C1 and E1) encompasses genetic, common and unique environmental influences on the first phenotype, which may also be shared with the second and third phenotypes. The second set of latent factors (A2, C2, and E2) encompasses genetic and environmental influences on the second phenotype which are independent of the first, but may be shared with the third. The third set of latent factors

represents those genetic and environmental influences that are unique to the third phenotype.

In the current study, Cholesky decompositions were fit examining the genetic and environmental covariance across reporters. This was done separately for anger, sadness, fear and shyness. In each case, mother- and father-report of temperament were expected to share the greatest amount of phenotypic overlap, whereas in-home observation was expected to be more independent. As with the univariate model, reduced multivariate models were tested by systematically dropping parameters, and selecting the most parsimonious model with no significant decrement in fit.

Analyses for Aim 3: Independent Pathway and Common Pathway Models. The third aim of the current study was to examine the extent to which anger, sadness, and fear could be accounted for by a common, latent negative emotionality factor, explained by a single set of shared genes and environmental influences. Although the Cholesky decomposition allows an estimation of the genetic and environmental influences on the covariance between multiple phenotypes, it is descriptive in the sense that it places no restrictions and tests no predictions regarding the ways in which a set of phenotypes may be related. The Cholesky decomposition not only allows the number of independent A, C and E influences on a set of traits to be equal to the number of traits, but allows multiple genetic and environmental factors to account for the covariance between traits. If the Cholesky decomposition is the best fitting model, it implies not only that the traits in question are relatively distinct, but that covariance between them is complex and best explained at the subordinate rather than superordinate level. For example, if there were two distinct additive genetic factors influencing sadness, one of which was shared with

anger and the other of which was shared with fear, the Cholesky model would be expected to provide the best fit to the data. In contrast, the independent pathway model (Figure 3) and the common pathway model (Figure 4) both test the hypothesis that the covariance between a set of phenotypes can be fully accounted for by a single set of shared genetic and environmental influences.

Of these two models, the independent pathway model is the less restrictive, as it posits that the shared genetic and environmental influences are directly related to each distinct phenotype, rather than acting through a single latent factor. Thus, this model assumes that the same genetic and environmental factors account for the covariance among anger, sadness and fear, but allows the magnitude of the pathways from these shared genetic and environmental factors to each emotion to differ, such that, for example, covariance between anger and sadness might be explained primarily by common environmental influences, and covariance between sadness and fear by genetic influences. In contrast, the common pathway model, which is nested within the independent pathway model, assumes that there is a common negative emotionality factor which itself has a single estimate of heritability, common environmental variance, and nonshared environmental variance, with each specific emotion loading on this factor to a greater or lesser extent. As in phenotypic factor analysis, factor loadings are estimated for each phenotype, and the variance in this common factor is decomposed into A, C and E (or A, D and E) components. In addition, the residual variance is also decomposed into A, C and E components, allowing an estimation of the genetic and environmental variance in each phenotype which is independent of the common factor. As a consequence, the common pathway model requires that any genetic or environmental

factors allowed to explain covariance between two emotions must also be shared with the third. To the extent that covariance between anger, sadness and fear is explained by a single common set of genetic and environmental influences, and the structure of these genetic and environmental influences is similar for each emotion, the common pathway model will provide a good fit to the data.

The common pathway model is the model that most closely represents the conceptualization of negative emotionality as a single, broad, genetically and environmentally influenced dimension of temperament. If the majority of constitutional, stable variance in fear, anger and sadness is accounted for by the same broad factor of temperament, then the common factor would be expected to provide a good fit for the data and account for substantial proportions of genetic and common environmental variance in each negative emotion, with residual variance primarily attributable to E (unique environmental influences, context effects and measurement error). In contrast, a poorly fitting common factor model, or the presence of substantial genetic or common environmental components unique to fear, anger or sadness, both suggest that there is meaningful variance in proneness to discrete negative emotions which is not well-accounted for by a broad negative emotionality factor.

Thus, the third aim of the current project was to test the fit of the common factor model for fear, sadness and anger, relative to the Cholesky and independent pathway models. This was done separately for anger, sadness and fear according to each reporter. The fit of a set of models was tested, beginning with the least restrictive trivariate Cholesky decomposition, progressing to the independent pathway model, and finally the most highly restrictive common pathway model. As with prior analyses, model fit was

assessed according to the chi-square difference test and the AIC. After the best fitting full model has been selected, the significance of individual parameters was tested and the most parsimonious model with no significant decrement in fit was selected.

Chapter 3

RESULTS

Preliminary Analyses

Means, standard deviations, and skewness and kurtosis for caregiver-report and observed temperament are presented in Table 4. No variables exceeded recommended cutoffs for skewness and kurtosis (2.00 and 7.00, Muthén & Kaplan, 1985), and all variables with scores exceeding three standard deviations from the mean were windsorized prior to analyses. Phenotypic correlations between mother- and father-report of anger, sadness, fear, shyness and broad negative emotionality are presented in Table 5. Because twins are clustered within families, all correlations were run in MPlus using the type=complex and cluster options (Muthén & Muthén, 1998-2011) in order to account for twin dependence while examining the full sample. As expected, phenotypic correlations between anger, sadness, and fear are high-to-moderate for both mothers and fathers, although the somewhat lower correlations between anger and fear (r_M = .256; r_F = *.205*), relative to anger with sadness (r_M = .560; r_F = .580) and sadness with fear (r_M = .420; r_F = .354), are noteworthy. In addition, phenotypic correlations show relatively high agreement between parents for all negative emotions (.475 – .622) and for shyness (.634). Correlations between observed anger, sadness, object and social fear are presented in Table 6. Correlations between observed emotions were often modest, with the highest correlation found for social and object fear (r = .262). Observed anger and sadness were correlated by .206, but it is important to note that these composites were formed from anger and sadness within the same two episodes, and part of this correlation is likely due to shared episode context. Finally, correlations between caregiver-report and in-home

observations of temperament are presented in Table 7. In contrast to the moderate-to-high convergence between mother- and father-report, convergence between caregiver-report and observed aspects of negative emotionality is uniformly low, with no correlation between parent-report and observed temperament higher than .100, aside from the inverse relation between parent-report of shyness and observed anger (-.227 and -.213 for mothers and fathers respectively).

Twin Intraclass Correlations

Twin intraclass correlations (ICCs) are presented in Table 8 for MZ and DZ twins, both aggregated across gender and considered separately for male, female and opposite sex twin pairs. The relative strength of MZ compared to DZ twin ICCs can be used to provide a rough index of heritability, and can suggest the potential presence of rater contrast effects, non-additive genetic variance, and sex differences in heritability. In this case, mother-report of anger and sadness and father-report of sadness show DZ correlations higher than half the MZ twin correlations, suggesting both additive genetic and common environmental components. In contrast, MZ twin correlations approximately twice as high as DZ twin correlations for mother-report of fear and father-report of anger and fear suggested primarily additive genetic and nonshared environmental variance, whereas the low DZ correlations for mother- and father-report of shyness suggested that it would be necessary to fit an ADE model and test for rater contrasts. In contrast to parent-report, twin ICCs for observed temperament show little evidence of heritability for any emotion except anger, with common environmental components likely to be important for observed sadness, object fear and social fear, suggesting that an ACE model would be most appropriate.

Quantitative Genetic Analyses

Saturated Models. Before fitting univariate ACE or ADE models, saturated models were fit in order to test for sex differences and rater contrast effects. Fully saturated multigroup models allowing means, variances and covariances to be freely estimated for male MZ, female MZ, male DZ, female DZ, and opposite sec DZ twins were fit, and tested against a series of restricted models constraining means and variances to be equal across twin pairs and zygosity groups, and constraining means, variances and covariances to be equal across sex. Covariance between twins is expected to differ for MZ and DZ groups, but unequal variances across zygosity groups indicate that parents may be contrasting the DZ twins against each other (Saudino et al., 1995). In the current study, separate saturated models were fit for mother-report, father-report and in-home observation of each emotion.

In each case except for anger and observed object fear, the -2 log likelihood chi-square test of fit indicated that it was possible to equate means, variances and covariances across sex, and means and variances across zygosity. For mother-report, father-report and observed anger, means could not be equated between male and female twin pairs, but there was no evidence of sex differences in variance or covariance, indicating that A, C and E components of variance did not need to be modeled separately for males and females. In the case of observed object fear, the means and variance of same sex male and female twins could be equated within twin pair, but equating the variance of opposite sex male and female twins led to a significant reduction in model fit according to the chi-square test ($\Delta \chi^2(10) = 19.12$, $p = .04$, $\Delta AIC = -.88$), and means, variances and covariances of twin pairs could not be equated across sex ($\Delta \chi^2(12) = 25.69$, $p = .01$,

ΔAIC = 1.69), indicating that it would be necessary to test a model that allowed the heritability of object fear to differ by sex. In all models tested, variance could be constrained to be equal across MZ and DZ twins, indicating that although some variables show DZ twin correlations lower than half the magnitude of MZ correlations, there is no evidence that these low DZ correlations are due to rater contrast effects.

Univariate ACE and ADE Models. Standardized estimates of A, C (or D), and E components for parent-report of each negative emotion, as well as fit statistics for the full and best fitting reduced models, are presented in Table 9. ADE models were tested for all variables in which twin interclass correlations suggested that non-additive effects are plausible (e.g. DZ correlations lower than half MZ correlations), including mother-report and father-report of fear and shyness. According to the best-fitting models for both mother- and father-report, anger, sadness and fear were all moderately-to-highly heritable, with additive genetic influences lowest for mother-report of anger (.45) and sadness (.45) and highest for father-report of fear (.80). Significant common environmental variance was also evident for mother-report of both anger (.26) and sadness (.27), but not for fear, which was best explained by an AE model with high additive genetic variance (.74). For father-report, both anger (.74) and fear (.80) were highly heritable, with no evidence of common environmental variance. However, although father-report of sadness also showed considerable additive genetic variance (.53), the common environmental component (.19) could not be dropped from the model without a significant loss of fit. Finally, as expected from low DZ twin correlations, ADE models were required for both mother-report and father-report of shyness. Non-additive genetic variance (D) was estimated at .68 for mother-report and .71 for father-report, and

additive genetic influences on shyness were estimated at .00 for both mother- and father-report, with the remainder of variance accounted for by nonshared environmental influences and measurement error. Because it is conceptually unlikely for all genetic influence on any phenotype to be non-additive, and because the twin design has limited ability to distinguish additive from non-additive genetic variance on the basis of phenotypic resemblance between twins, the full ADE model was retained, with the D component best interpreted as a broad heritability estimate encompassing both additive and non-additive genetic effects.

Table 10 summarizes A, C and E components and fit statistics for the full and best fitting models for in-home observations of anger, sadness, object fear, and social fear. Because the saturated model for observed object fear indicated that variance could not be equated across opposite sex twins, the first model tested allowed for A, C and E components to differ between male and female twins. However, this model did not fit significantly worse than a model in which A, C and E components were constrained to be equal across sex ($\Delta \chi^2(4) = 2.46$, $p = .65$, $\Delta AIC = -5.54$); because modeling sex differences in the multivariate Cholesky framework affects the interpretability of results (Neale, Røysamb, & Jacobson, 2006), results for the univariate ACE model with variances constrained across sex are reported.

In contrast to findings for parent-report, anger was the only observed emotion to show any significant genetic variance (.47), with the remainder of variance in observed anger explained by nonshared environmental influences and measurement error (.53). For observed sadness, object fear, and social fear, the CE model presented the best fit to the data. Common environmental variance was modest but significant for sadness (.22) and

object fear (.17), and moderate for social fear (.40), but as expected for observed data, the majority of variance was accounted for by nonshared environmental influences and measurement error (.60-.83).

Multivariate ACE and ADE Models. After all univariate models were fit, a series of multivariate Cholesky decompositions were fit to examine the genetic and environmental covariance shared between reporters and between emotions. Because the aim of the multivariate Cholesky decomposition is to parse the variance shared between two or more phenotypes, a phenotype was only included in the analysis if it had a strong enough correlation with at least one other variable ($r > .150$) for shared variance to be meaningful. Thus, four bivariate models were fit examining covariance between mother- and father-report for anger, sadness, fear and shyness, but no attempts were made to examine genetic and environmental covariance between parent-report and observed data. In addition, two trivariate Cholesky decompsitions were fit examining genetic and environmental covariance among anger, sadness and fear, one for mother-report of temperament and one for father-report, and two four-variable Cholesky decompositions were fit for mother-report and father-report of anger, sadness, fear and shyness, respectively. Two bivariate Cholesky decompositions were also fit for observed data, one examining the covariance between anger and sadness and one examining the covariance between object and social fear. However, observed anger and sadness were not highly correlated enough with either object or social fear to consider them in a trivariate model.

Genetic and Environmental Covariance Across Reporter. First, four bivariate Cholesky decompositions were fit to examine the variance shared between mother-report and father-report for each aspect of temperament (anger, sadness, fear, and shyness). As

with the univariate models, after the full model was fit, paths were systematically dropped until the best-fitting, most parsimonious model was determined. Table 11 describes the results for the full and best fitting reduced models comparing A, C (or D) and E covariance across reporters, and the full and best fitting reduced models are also depicted in Figure 5. For anger, all additive genetic influences were fully shared between mother- and father-report, although this set of additive genetic influences explained a greater proportion of the variance in father-report (.74) than mother-report (.36). Mother-report of anger also had a significant independent common environmental component (.33), whereas father-report was fully explained by genetic and nonshared environmental factors. For sadness, there was evidence for both shared and reporter-specific genetic variance, with 32% of the variance in father-report sadness explained by additive genetic influences shared with mother-report, 41% explained by independent additive genetic influences, and 24% explained by nonshared environmental influences and measurement error. As with anger, mother-report of sadness was moderately heritable (.47) but also explained by common environmental influences not shared with father-report (.24). In contrast, the best-fitting model for mother- and father-report of fear was an AE-AE model with high heritability for mother-report (.75) and father-report (.80), and the genetic influence on father-report of fear was both shared (.33) and reporter-specific (.47). Finally, mother-report of shyness was explained by both additive genetic (.02) and non-additive genetic (.67) components, and father-report was explained by both non-additive genetic influences (.39) fully shared with mother-report and by a substantial independent additive genetic component (.31). For all emotions and for shyness, total nonshared environmental variance ranged from 25%-31% for mother-report and 20%-30% for

father-report. These nonshared environmental influences were primarily unique to each reporter, but in each case there was a very small (.02-.05) proportion of variance shared across reporters which could not be dropped without a significant loss of fit to the model.

Genetic and Environmental Covariance Across Emotion. After examining the genetic and environmental components of the covariance between mother- and father-report, a series of Cholesky decompositions were fit in order to describe the genetic and environmental variance shared between different aspects of negative emotion. First, two trivariate Cholesky decompositions were fit for anger, sadness, and fear, one for mother-report and one for father-report. Unfortunately, it was not possible to fit a multivariate Cholesky decomposition including observed anger, sadness and object or social fear, because object and social fear were not sufficiently highly correlated with observed anger or sadness ($r < .100$). However, two bivariate Cholesky decompositions were fit examining the covariance between anger and sadness, and social fear and object fear, respectively. Fit statistics for the full and best fitting reduced models examining variance shared across emotion are described in Table 12, and standardized A, C, and E components are reported in Table 13. Figures 6 and 7 depict the standardized and unstandardized A, C and E components for mother-report and father-report, respectively, and Figure 8 depicts the standardized and unstandardized A, C and E factors for in-home observations.

For mother-report anger, sadness and fear, the best fitting model was an ACE-ACE-AE model, with evidence for both shared and emotion-specific genetic variance. One set of additive genetic influences explained 47% of the variance in anger, 17% of the variance in sadness, and 9% of the variance in fear, and a second set of genetic influences

independent of anger was also found to explain 36% of the variance in sadness and 10% of the variance in fear, although the majority of genetic influences on fear (.55) were independent of both anger and sadness. In addition, common environmental influences were significant for both anger (.24) and sadness (.18), but not fear, and all common environmental influences could be explained by a single component fully shared across emotion. A small proportion of nonshared environmental variance was found to be shared between anger and sadness (.01) and between sadness and fear (.02), although the majority of nonshared environmental influence was specific to each emotion.

For father-report, the best-fitting model was an AE-AE-AE model, with both shared and emotion-specific additive genetic variance. Father-report of anger was highly heritable (.72), and additive genetic influences on anger also explained a substantial proportion of the variance in father-report of sadness (.35) and a small but significant proportion of the variance in father-report of fear (.05). A second set of additive genetic influences also explained 38% of the variance in sadness and 12% of the variance in fear, although again the majority of genetic influence on fear (.63) was not shared with other emotions. As with mother-report, the majority of nonshared environmental influence was specific to each emotion, although a small but significant proportion of nonshared environmental variance (.02) was shared between anger and sadness.

Note that four-variable Cholesky decompositions examining the impact of including shyness were also considered for parent-report (parameter estimates reported in Table 14; see Table 12 for model fit statistics)**,** with results for the variance shared among anger, sadness and fear remaining highly stable, but in both cases the genetic and environmental overlap between shyness and the other phenotypes was small. For

instance, although the same set of dominant genetic factors contributed to father-report of anger and shyness, this D factor contributed to 63% of the total variance in shyness in the final model but only 1% of the total variance in anger.

For in-home observations of anger and sadness, the best-fitting model was an AE-ACE model, with the phenotypic correlation between anger and sadness fully explained by additive genetic factors, although this shared genetic component explained 47% of the variance in anger and only 9% of the variance in sadness. Sadness was also found to have a significant common environmental component (.16), although both anger and sadness were primarily explained by nonshared environmental variance separate to each emotion. In contrast, the best-fitting model for observed object fear and social fear was a CE-CE model, with the same set of common environmental influences accounting for 18% of the variance in object fear and 41% of the variance in social fear. This shared C component fully accounted for the phenotypic correlation between object and social fear, although once again the majority of variance was accounted for by emotion-specific nonshared environmental components.

Independent Pathway and Common Pathway Models. After arriving at the final, most reduced Cholesky decomposition for mother-report and father-report of anger, sadness, and fear, the next step was to test whether two increasingly restrictive multivariate models, the independent pathway model and the common pathway model, could account for the genetic and environmental covariance between these three negative emotions. First, two independent pathway models were fit for mother- and father-report of anger, fear, and sadness, and these models were compared to the full and final, most reduced Cholesky decompositions using the AIC (Akaike, 1987). Table 15 summarizes

the fit of the full and most reduced independent pathway models for mother- and father-report of anger, sadness and fear in comparison to the corresponding full and most reduced Cholesky decompositions.

The common pathway model is nested within the independent pathway model, allowing comparisons to be made on the basis of the chi-square difference test as well as the AIC. Fit statistics and parameter estimates of the full and reduced independent pathway models and common pathway models are reported in Table 16, and the full and final independent pathway models for mother-report and father-report are depicted in Figures 9 and 10, respectively. For mother-report of anger, sadness, and fear, the independent pathway model did not fit significantly worse than the Cholesky decomposition (Δ AIC = .14, Δ df = 0 with the same number of parameters estimated), suggesting that the genetic and environmental covariance between these negative emotions can be represented by a single set of shared genetic, common environmental and nonshared environmental factors. However, the common pathway model did lead to a significant decrement in fit relative to the independent pathway model ($\Delta \chi^2(4) = 21.92$, $p < .001$, ΔAIC = 13.92). In the final independent pathway model for mother-report anger, sadness and fear (depicted in Figure 9), it was necessary to retain shared A and E components for all emotions, and a shared C component for anger and sadness, although the path from the shared C component to fear could be dropped without a significant decrement in fit. Interestingly, for sadness, all additive genetic (.53), common environmental (.17) and nonshared environmental (.29) variance was fully shared with other emotions, and for anger, all common environmental variance (.21) was accounted for by the shared C component. However, there was evidence of both shared and

emotion-specific additive genetic influences on both anger and fear, with 17% of the variance in anger and 18% of the variance in fear explained by the shared A factor, and 34% of the variance in anger and 56% of the variance in fear attributable to additive genetic influences not shared with other emotions. Although it was necessary to retain paths from the common nonshared environmental component to both anger and fear, the amount of variance explained by this factor was small (.01 and .03 for anger and fear, respectively), with 27% of the variance in anger and 24% of the variance in fear accounted for by nonshared environmental factors specific to that emotion.

For father-report of anger, sadness, and fear, the independent pathway model did not fit significantly worse than the Cholesky decomposition (Δ AIC = 2.76, Δ df = 0). Moreover, the highly restrictive common pathway model did not lead to a significant loss of fit relative to either the full independent pathway model ($\Delta \chi^2(4) = 5.67$, $p = .23$, ΔAIC = -2.34) or the full Cholesky decomposition (Δ AIC = .42, Δ df = 4), suggesting that for father-report, a single negative emotionality factor provides an adequate representation of the data (see Table 16). This factor, depicted in Figure 11, was found to be highly heritable (.88), with the remainder of variance accounted for by nonshared environmental influences. However, substantial emotion-specific additive genetic influences were also necessary to explain both anger (.38) and fear (.68), and all emotions had significant emotion-specific nonshared environmental components ranging from 16% for sadness to 25% for anger. Factor loadings on the common factor were higher for sadness (.6397) and anger (.5307) than fear (.332), which seemed to be relatively more independent of the other emotions.

Chapter 4

DISCUSSION

The goal of the current study was to use quantitative genetic modeling to characterize the genetic and environmental underpinnings of anger, sadness and fear in middle childhood, with the aim of addressing two key issues. First, the study was intended to examine the extent to which proneness to anger, sadness and fear can be said to represent the same broad, genetically- and environmentally-influenced predisposition toward negative emotionality, as opposed to one or all of these emotions being more genetically and environmentally distinct. Secondly, the study was intended to examine the extent to which different methods of measuring each of these emotions (mother-report, father-report, and in-home observation) appear to be tapping into the same basic underlying traits, as indexed by shared genetic and environmental variance across reporter. In both cases, findings highlight the need to consider distinctions as well as commonalities across emotions, reporters and situations. Regarding the first question, results support the validity of negative emotionality as a coherent, genetically-influenced trait, but also provide clear evidence for genetic influences on both anger and fear that are not shared with other emotions. The presence of such independent genetic variance on fear and anger indicates that a more fine-grained, emotion-specific approach to temperament is likely to be valuable, and provides support for the view of discrete emotions as at least partially biologically distinct. Regarding the issue of measurement, findings are more complicated, particularly for observed emotion, and highlight the need for greater consideration of the context in which an emotion is elicited and expressed,

especially when the aim is to assess temperament as a constitutional, biologically-based and situationally stable construct.

Phenotypic Convergence and Heritability of Negative Emotion in Middle Childhood

Because earlier genetically-informed studies of temperament have tended to focus on infancy and early childhood, one aim of the current study was simply to describe the univariate heritability and the phenotypic convergence across reporters for mother-report, father-report, and in-home observations of anger, sadness and fear in middle childhood. Consistent with prior research (Rothbart et al., 2001), there was moderate agreement between parents for all negative emotions and for shyness. As expected, the lowest convergence was found for sadness, although it was anger rather than fear that showed the highest agreement across reporters. However, convergence between parent-report and in-home observation was modest to negligible for both parents, although surprisingly, father-report was more strongly related to observed emotion than mother-report in all cases. One unanticipated exception to the pattern of very low convergence was the comparatively strong, although still low-moderate, inverse correlations between parent-report of shyness and observed anger.

Consistent with previous reports of moderate heritability of temperament in early and middle childhood (Gagne et al., 2009; Goldsmith et al., 1997; Saudino, 2005), both mother- and father-report of anger, sadness and fear showed moderate to high heritability at the univariate level, with the highest heritability evident for fear according to both mother and father-report, and the lowest for mother-report of anger and sadness and father-report of sadness. However, there were also some unanticipated findings, the most notable of which may be the presence of significant common environmental influences

on several dimensions of reported and observed temperament which have been primarily explained by additive genetic factors in prior studies (e.g., DiLalla et al., 1994; Goldsmith et al., 1997; McGuire et al., 2003; Mullineaux et al., 2009). Specifically, the current study found evidence for common environmental variance in sadness at the univariate level according to mother-report, father-report, and in-home observation, although it should be noted that the common environmental influence on father-report of sadness were not significant when considered in multivariate models. In addition, negligible heritability and significant common environmental variance were found for both in-home observations of both object fear and social fear, though not parent-reported shyness. However, contrary to expectations, common environmental variance in anger according to the current study was only evident for mother-report, with father-report anger explained primarily by additive genetic factors and observed anger found to be moderately heritable with no evidence of common environmental influences.

These results are contrary to my hypotheses, as prior research suggests the presence of modest but significant common environmental variance for both observed and mother-report of anger (Emde et al., 2001; Gagne & Goldsmith, 2001; Goldsmith et al., 1997) but not mother-report of sadness (Mullineaux et al., 2009), and anger has also been theoretically and empirically linked to other approach-related dimensions of temperament such as positive emotionality (e.g., Carver & E. Harmon-Jones, 2009; Deater-Deckard et al., 2010; Lara & Akiskal, 2006; Rothbart et al., 2000; Rydell et al., 2003) which have been found to be less heritable (Goldmith et al., 1997; 1999; Silberg et al., 2005). However, these findings are consistent with the literature, as there has been only limited research considering sadness at the subordinate level, and not all studies of

anger find common environmental influences (Deater-Deckard et al., 2010; Mullineaux et al., 2009). Moreover, according to previous research, aspects of temperament related to fear, shyness or behavioral inhibition are often among the most heritable when assessed in laboratory interactions with peers and adults (e.g., DiLalla et al., 1994; Goldsmith et al., 1999; McGuire et al., 2003), possibly because a child's predisposition towards shyness or withdrawal may be facilitated rather than suppressed by unfamiliar laboratory situations and interactions with strangers. The use of in-home rather than laboratory observations in the current study may have contributed to differences between this and previous studies regarding observed social and object fear, as the home may be a more comfortable context for some shy or fearful children. In addition, some differences between earlier and current findings may be explained by differences in sample characteristics, as the majority of earlier research has taken place in younger samples (e.g., Goldsmith et al., 1997; Gagne et al., 2009; Saudino, 2005), with few quantitative genetic studies examining lower-order dimensions of negative emotionality in middle childhood. Previous behavior genetic research is consistent in finding that heritability tends to increase rather than decrease with age (Saudino, 2005), suggesting that the older sample is unlikely to be the only explanation for the higher common environmental variance found in the current sample. However, previous research in this age range has either examined lower-order dimensions of temperament in a small sample which may have limited power to detect common environmental influences (e.g., Mullineaux et al., 2009) or examined anger, sadness and fear at the level of broad negative emotionality across wide age ranges from childhood to adolescence (Singh & Waldman, 2010; Tackett

et al., 2011), which may underestimate emotion-specific common environmental influences such as those found for sadness in this study.

In addition, in the current study there were differences between mother-report and father-report in estimates of heritability and common environmental variance, which are somewhat different from findings reported in the limited amount of previous research (Mullineaux et al., 2009). Specifically, the single previous study to examine genetic influences on father-report of negative emotion in middle childhood found that mother-report of negative emotionality and its subordinate dimensions were all fully explained by additive genetic and nonshared environmental variance, whereas father-report of sadness was largely explained by common environmental influences with no evidence of heritability (Mullineaux et al., 2009). In contrast, the heritability of father-reported negative emotion in the current sample ranged from slightly to considerably higher than mother-report across all three emotions, with the difference most obvious at the univariate level for anger. Again, differences in sample characteristics are one possible explanation for differences between current and prior research, as Mullineaux and colleagues (2009) assessed a smaller sample with a higher percentage of female twins, and father-report in that study was only available for approximately half of the families. However, the higher heritability of father-report and lack of common environmental influence on anger in the current sample may also simply be an artifact of demographic or other systematic differences between families with and without father-report data. Although father-participation in the current study was high (approximately 80%), the proportion of participants without father-report data may still be high enough to affect estimates of mother-reported heritability. The heritability of other phenotypes such as

cognitive ability has been found to be influenced by sample characteristics that might be expected to differ between families with and without father-report data (e.g., SES; Turkheimer et al., 2003). However, although the literature considering moderated heritability is still too sparse to make strong directional predictions, especially where temperament is concerned, the heritability of negative emotionality has been found to be unaffected by chaos in the home and higher for children exposed to more crowded or unsafe conditions (Lemery-Chalfant, Kao, Swann, & Goldsmith, 2013). This suggests that if a lack of father-report data in the current study is indicative of a more chaotic or adverse environment, any difference would more likely be in the direction of higher rather than lower heritability. Follow up analyses are planned to test the possibility that differences in heritability between mother- and father-report are an artifact of demographic differences, both by examining differences in demographic factors (age, SES, percentage of female twins, racial and ethnic composition) and levels of each emotion depending on whether or not twins had father-report data, and by examining the univariate heritability of each emotion separately in families with and without father-report data. However, it may also be that fathers are picking up on an aspect of children's temperament that is genuinely different and more genetically influenced, although the meaning of this difference is not immediately clear.

Another unexpected finding at the univariate level was the extremely low DZ intraclass correlations for shyness, resulting in high estimates of non-additive genetic variance. This finding of high non-additive genetic variance in reported shyness is consistent with near-zero DZ correlations with fear, anger and shyness in previous studies using the CBQ (Goldsmith et al., 1997; 1999). In contrast to prior research using the EAS

(Hwang & Rothbart, 2003), the low DZ correlations for parent-report of shyness in the current study could not be explained by rater contrast effects, which would be indicated by unequal variance between MZ and DZ groups (Goldsmith et al., 1999; Saudino et al., 1995). It may be that these low DZ correlations are simply the result of strong non-additive genetic influences, but another possible explanation is that by middle childhood, shyness is subject to multiplier effects which act to magnify small differences over time through a series of cumulative interactions with the environment (e.g., Ceci et al., 2003), such that initially slightly more outgoing children are more likely to seek out or be exposed to social situations, leading to gains in proficiency and comfort with peers and decreases in behavioral manifestations of shyness, whereas initially slightly shyer children experience lower exposure to social situations and lower social success, which in turn leads to fewer opportunities to develop adaptive social skills and increasing levels of withdrawal over time. Such a multiplier effect would predict increasingly divergent levels of shyness over time in less genetically-related individuals, and this might be tested using a longitudinal examination of changes in shyness over time for MZ and DZ twins. However, in the absence of further investigation, caution in interpretation is warranted.

Finally, another issue that needs to be taken into consideration is the high nonshared environmental variance found for in-home observations of temperament, particularly in the case of sadness and object fear. High estimates of nonshared environmental variance are common in observational twin studies, with estimates typically ranging from 40%-80% depending on the dimension of temperament under consideration (Braumgart et al., 1992; Deater-Deckard et al., 2007; Gagne & Goldsmith, 2011; Gagne & Saudino, 2010; Goldsmith & Gottesman, 1981; Rhee et al., 2012;

Saudino, Plomin & DeFries, 1996), and findings from the current study are comparable, although as noted, heritability tends to be higher and nonshared environmental variance lower in studies of shyness or behavioral inhibition (e.g., DiLalla, Kagan, & Reznick, 1994; Emde et al., 1992; McGuire et al., 2003). However, the interpretation of these high estimates of nonshared environmental variance is difficult, as nonshared environmental variance may index meaningful contributions of the child's environment, whether due to immediate context effects or stable cross-situational influences, or it may index measurement error. Although the classic twin design does not allow a differentiation between meaningful contributions of the nonshared environment and measurement error, and the multiple factors that lead to differences between twins are difficult to measure, true nonshared environmental variance is nevertheless an important contributor to child development (Turkheimer, 2000). One way to examine whether high estimates of nonshared environmental variance in the current study indexes meaningful contextual influences rather than measurement error is to test whether nonshared environmental variance is reduced when only tapes coded by the most reliable coders (i.e., those closest to the master coder) are considered, as well as whether nonshared environmental variance increases when single episodes rather than cross-episode composites are considered. If nonshared environmental variance decreases when only episodes coded by the highest quality coder are considered, it suggests that high estimates of nonshared environmental variance are not inherent to the use of observational measures, but avoidable error caused by the inclusion of less reliable data. However, if similarly high estimates of nonshared environmental variance are found even when only the most reliable coders are considered, it indicates the presence of true differences in twins' emotional responses

which might be explained by the immediate episode context or by more stable individual differences. In addition, an increase in nonshared environmental variance when emotions are considered at the episode-specific rather than the cross-episode level suggests that aggregation is useful in arriving at a more consistent, stable measure of emotion even given distinct contexts and relatively low cross-episode correlations. Follow-up analyses are planned to test both of these questions in more detail.

Genetic and Environmental Influences on Covariance Across Reporters

After characterizing the univariate heritability of anger, sadness and fear, the second major goal of this study was to examine genetic and environmental influences on the covariance across reporters for each of these emotions. Bivariate Cholesky decompositions of the covariance between mother-report and father-report of temperament indicated that cross-reporter correlations for all emotions were almost entirely accounted for by additive genetic influences, with common environmental variance significant only for mother-report and non-shared environmental influences primarily-reporter specific. This is unsurprising, as common environmental influences were typically nonsignificant for father-report in the univariate models, but does suggest that aggregation across reporters is useful for arriving at a more consistent, heritable construct. However, to the extent that mothers are more likely to be primary-caregivers and spend more time with children outside of work (Pleck & Masciadrelli, 2004; Walker & McGraw, 2000), mothers may also be more apt to notice aspects of twins' behaviors directly influenced by the immediate, day-to-day environmental context, perhaps explaining higher findings of common environmental variance for mothers. One way to test this possibility might be to compare data from mothers who report being

homemakers to those who work full time outside the home, to see whether higher common environmental variance in temperament is found according to the reports of mothers who spend more time in the home.

Interestingly, although mother-report and father-report of anger appeared to have the most noticeable differences in heritability at the univariate level, and there were significant common environmental influences on mother-report of anger that were not shared with father-report, bivariate models suggested that fathers were reporting on much the same genetically-influenced aspects of children's anger as mothers. In contrast, despite relatively similar estimates of heritability and common environmental variance for mother-report and father-report of sadness and fear at the univariate level, father-report of sadness and fear both showed evidence of substantial unique genetic variance. If it had been found that the majority of independent variance in father-report of fear and sadness was accounted for by nonshared environmental influences, as was the case with anger, it might have suggested that discrepancies between reporters are primarily the result of measurement error, and perhaps that one parent or the other is simply a less nuanced or knowledgeable reporter of children's temperament. However, the presence of independent genetic variance in father-report of fear and sadness suggests that fathers are not merely better or worse at assessing the same underlying dimension of child temperament, but picking up on something genetically distinct and meaningful.

There are many potential explanations for why fathers might observe and report on different aspects of children's temperament from mothers. First, fathers may be more likely to interact with children in different situations than mothers, and consequently to observe or elicit different emotional responses. For example, despite recent increases

father participation in childrearing, fathers are still less likely to be primary caregivers (Pleck & Masciadrelli, 2004; Walker & McGraw, 2000), and more likely to engage in active play relative to other caregiving activities (Pleck & Masciadrelli, 2004; Yeung, Sandberg, Davis-Kean, & Hofferth, 2001; Schoppe-Sullivan, Kotila, Jia, Lang, & Bower, 2013), whereas mothers are more likely to be involved in supervision and physical care (Moon & Hoffman, 2008). As a consequence, fathers and mothers may be most familiar with different aspects of their children's emotional responses, which might be influenced by different genetic and environmental factors. For instance, in samples of children with disruptive behavior problems, mothers report higher parenting stress, more negative impact on their lives, and higher rates of problem behavior (Baker & Heller, 1996; Calzada, Eyberg, Rich, & Querido, 2004), suggesting that fathers may be less exposed to the stresses of parenting highly difficult children, although in the current study the focus was on normative temperament rather than externalizing problems, and mothers and fathers reported similar mean levels of children's anger and other negative emotions.

Secondly, fathers and mothers may interpret the same emotional responses differently, not only in terms of what emotion is being expressed but in terms of how socially appropriate or inappropriate it is for a seven year old child to cry easily or express anger. If reporters perceive their child's responses to be inappropriate or extreme, as opposed to age-typical, they might be more likely to view and rate their child as highly reactive (Kagan & Fox, 2006), and mothers' and fathers' perceptions of appropriate behavior may also influence parents' responses to the child. For instance, fathers have been found to be less tolerant than mothers of oppositional defiant behavior in girls, and less tolerant than mothers for children of both genders when only nonclinical levels of

disruptive behaviors are considered (Wright, Parent, Forehand, Edwards, Connors-Burrow, & Long, 2012). However, children may also be more likely to show and mask different emotions, or express their emotions differently, in the presence of mothers versus fathers, and when children do express negative emotions to fathers or to mothers, mothers and fathers might be more likely to respond or encourage their children to respond in different ways, such as offering comfort and encouraging expression versus encouraging instrumental coping strategies.

There is some evidence for differences between mothers and fathers in emotion socialization and support, with mothers more likely to engage in emotion-coaching behaviors and emotion-focused discussions with their children (Baker, Fenning, & Crnic, 2011; Cassano & Perry-Parrish, 2007; Fivush, Brotman, Buckner, & Goodman, 2000; Stocker, Richmond, Rhoades, & Kiang, 2007), and fathers may be more likely to meet children's negative emotions with punitive or minimizing responses (Cassano & Perry-Parrish, 2007; McElwain, Halberstadt, & Volling, 2007). This may be especially true when parents perceive their children's negative emotions to be less well-regulated, as Cassano and Zeman (2010) report that parents who were led to believe that their children's sadness regulation is below age-typical levels were more likely to use dismissing than supportive responses in a subsequent interaction, particularly in the case of fathers. In addition, there is some evidence that children are more likely to show compliance in interactions with fathers (Buhrmester, Camparo, Christensen, Shapiro Gonzalez, & Hinshaw; Calzada et al., 2004), suggesting that children themselves do respond differently in interactions with fathers and mothers. One possibility is that to the extent that fathers in the current sample are on average less emotionally expressive or less

likely to be involved in children's day to day lives, father-report of children's temperament may be affected by children's genetically-influenced shyness. This also represents one reason why father-report might be more similar to observational measures of emotion. It may not be that fathers are more accurate reporters of children's responses, but that the situations in which fathers observe children's responses are more similar to observed tasks, or that the way children respond to fathers is more similar to how they respond to experimenters.

However, despite this possibility, it should be noted that correlations between parent-report and observed measures were insufficient to support a multivariate genetic analysis for either mothers or fathers. Although not unexpected, this low phenotypic convergence between observed and reported emotion does have implications for the interpretation of results. Although different measures of temperament are often assumed to be measuring fundamentally the same construct, parents are likely to be reporting on aspects of children's temperament that are consistent and stable across situations (Rothbart & Bates, 2006), whereas with observed data, the context is an inextricable part of the observation. Aggregation across multiple observed contexts and episodes may be useful in arriving at a more consistent measure of children's temperament (e.g., Eaton, 1994), although in the current study the limited number of episodes and low cross-episode correlations prevented aggregation across more than two episodes for sadness and more than three for anger. However, even when aggregation is used, there may be aspects of children's in-the-moment emotional responses that are not well-captured by parent-report and vice versa, such as the influence of shyness on children's willingness to express anger to a friendly stranger.

This may be especially true for observed data, especially in light of the relatively low correlations seen in the current study not only between different observed emotions but between different episodes intended to assess the same emotion. As Rothbart and Bates (2006) note, highly negatively reactive children are not always negatively reactive; rather, these children show a predisposition toward responding with more intense negative emotion under relevant environmental conditions, and further attention to what those conditions might be has the potential to clarify the meaning of observed emotion and the best way to arrive at consistent observed measures of temperament. Some situations are likely better elicitors of emotion in general for the majority of individuals. For example, the criticism used in the episode Impossibly Perfect Stars in the current study appeared to be a relatively stronger, more salient elicitor of sadness than receiving a disappointing gift in this age group. However, there are also likely individual differences in the intensity and type of a child's emotional response to specific situations, at least partially based on personal experience (e.g., a child who was bitten by a dog retaining a fear of large animals), but perhaps also on more innate, biologically-based differences in temperament. As such, it may be important for observed temperament researchers to consider what aspects of a situation might be especially salient to children in a particular age range, and how children's reactions might be affected by aspects of reactivity and regulation that are not the primary emotion or dimension of temperament the situation is meant to assess.

The extent to which a task or situation contains social elements is likely a critical factor influencing children's emotional responses, as suggested by the finding that children rated by parents as high in shyness were significantly less likely to express anger

to an experimenter. However, the nature of the social situation itself is also likely to matter. For instance, reported shyness was related to children's behavior during anger episodes, but not during the episode designed to assess social fear. It may be that the structured nature of the storytelling task made it less threatening for some children who are shy in unstructured social situations, or that interactions with adults may be easier for some children who show high shyness with peers, whereas the potential performance-based or evaluative aspects of this task may have elicited shyness in some children who have little difficulty interacting with peers but are more prone to anxiety.

In addition, the inverse correlation between shyness and observed anger suggests another important consideration for observed studies of temperament and emotion: the degree to which a negative emotional response in a given situation may be normative or even adaptive. However, although the Laboratory Temperament Assessment Battery is designed to assess the full range of normative individual differences in standardized approximations of salient real-world contexts, and does not include episodes which have been found to primarily capture extreme or dysregulated emotion (Goldsmith et al., 2001), it is important to remember that children may show low levels of anger or sadness in such observed contexts for many reasons, and some of them may be adaptive (e.g., good regulation), but some may not (e.g., an inability or unwillingness to express emotion in social contexts due to high shyness). High levels of negative emotion are often considered to be maladaptive or risky (e.g., Nigg, 2006), but context-appropriate negative emotions do serve an adaptive function (Nesse & Ellsworth, 2009; Tooby & Cosmides, 2008), and some observed tasks may be more likely than others to tap into contextually and socially appropriate, adaptive negative emotion rather than negative

emotion as extreme or problematic. For instance, the ability to feel and express moderate anger in response to a situation that is genuinely unfair and indicative of socially inappropriate behavior on the part of another person, such as unequal sharing, may be indicative of protective traits such as self-esteem or willingness to stand up for oneself. In contrast, the most adaptive response to a task that is frustrating but not obviously unfair may be persistence without anger, and moderate levels of anger in such a context may be more likely to reflect socially inappropriate behavior on the child's part, indicative of high reactivity or lack of regulation. Indeed, Locke, Davidson, Kalin and Goldsmith (2009) report that context inappropriate anger within lab-TAB episodes intended to elicit positive emotion was only moderately related to context appropriate anger in Not Sharing, and associated with lower morning cortisol and flatter cortisol slopes, both of which are possible markers of HPA-axis dysregulation (Gunnar & Vasquez, 2001) which have been linked with higher risk for externalizing problems (Alink, van Ijzendoorn, Bakermans-Kranenburg, Mesman, Juffer, & Koot, 2008). In contrast, context appropriate anger in Not Sharing was associated with higher morning cortisol and steeper slopes (Locke et al., 2009), suggesting that children's anger responses in this episode are not necessarily associated with dysregulation and may even be adaptive.

Genetic and Environmental Influences on Covariance Across Emotions

The third aim of the current study was to test whether genetic and environmental influences shared across emotions could be accounted for by a single common factor, or whether more specificity would be required to explain relationships between aspects of negative emotionality at the subordinate level. Although there is reason to expect that distinct neural and biological systems may be involved in proneness to different emotions

such as anger or fear, many developmental theories and studies assume that genetic and environmental influences on anger, sadness and fear at the subordinate level act through a coherent, biologically-based higher order negative emotionality dimension (Rothbart & Bates, 2006). In order to examine the extent to which this assumption adequately represents the genetic and environmental underpinnings of proneness to discrete negative emotions and the relations between them, the current study tested two nested models, the independent pathway model and the common pathway model, both of which posit that a single set of shared genetic and environmental factors is sufficient to account for the covariance among all emotions. Findings were slightly different for mother-report and father-report, in the sense that covariance between father-report of anger, sadness and fear could be adequately represented by the highly restrictive common pathway model, whereas the less-restrictive independent pathway model was necessary to explain relations among these emotions according to mother-report. The finding that the common pathway model could not adequately represent the covariance between mother-report of anger, sadness and fear is likely due to the presence of a small but significant common environmental factor that was shared between mother-report of anger and sadness but not fear, as the independent pathway model allows the modeling of common environmental variance for some but not all emotions, whereas the common pathway model requires genetic and environmental influences on the common factor to be the same for all emotions. However, in both cases the implications are similar: there does appear to be a common set of factors that influences all three negative emotions to a greater or lesser degree, supporting the validity of a dispositional tendency toward broad negative emotional reactivity, which in the current sample is explained primarily by additive

genetic factors. However, it is important to note that not all emotions were equally well accounted for by this shared set of genetic influences. According to both mother-report and father-report, genetic influences on sadness were fully shared with both anger and fear, but a majority of the variance in fear, and a smaller but still substantial proportion of the variance in anger, was explained by independent, emotion-specific genetic influences.

The finding that fear was largely genetically distinct from anger and sadness regardless of reporter is not entirely surprising, in light of developmental literature suggesting that anger and fear do show some stability over time, but are not longitudinally related to each other from infancy to middle childhood (Rothbart et al., 2000), as well as differential prediction of outcomes such as internalizing and externalizing symptoms when anger and fear are considered separately (e.g., Lemery et al., 2002; Oldehinkel et al., 2007). In addition, this finding is consistent with a recent study examining the structure of observed temperament in preschoolers, in which there were separate factors for dysphoria (anger, sadness and hostility) and fear/behavioral inhibition (Dyson, Olino, Durbin, Goldsmith, & Klein, 2012), as well as with theoretical work drawing distinctions between fear as an anticipatory emotion versus anger and sadness as reactions to discrete events (e.g., Putnam, Ellis, & Rothbart, 2001; Putnam et al., 2008). Dyson and colleagues (2012) suggest that the age at which fear and dysphoria begin to show a structure more similar to adult negative emotionality is one important developmental question. A consideration of the underlying genetic and environmental influences found provides evidence for the greater independence of fear in middle childhood, even according to parent-report.

However, despite the fact that anger was more strongly related to sadness than fear for both genetic and common environmental reasons, anger also had substantial emotion-specific genetic variance, which might plausibly index genetic influences on approach-related aspects of temperament that are shared with high intensity positive emotions or impulsivity, but not the withdrawal-oriented emotions of sadness and fear. Previous research supports both theoretical and empirical links between anger and approach-related or reward-seeking behaviors (Carver & Harmon-Jones, 2009; He, Xu, & Degnan, 2012), impulsivity and high intensity positive emotion (Deater-Deckard et al., 2010; Lara & Akiskal, 2006; Rothbart et al., 2000), with physiological measures such as left frontal asymmetry predicting anger, exuberance and approach orientation (Davidson, 2002; E. Harmon-Jones, Vaughn-Scott, Mohr, Sigelman, & C. Harmon-Jones, 2004). In addition, one recent quantitative genetic study reported that anger does share genetic and nonshared environmental variance with parent-report of anticipatory positive affect and approach (Deater-Deckard et al., 2010), but although it is likely that this variance is independent of proneness to sadness or fear, genetic and environmental influences on these emotions have yet to be considered in the same model as anger and positive affect. Finally, at the neurobiological level, distinct systems motivating behavioral inhibition and activation may underlie differences in anger and fear through approach and avoidance behavior and motivation, whereas other systems or structures involved in reactivity (e.g., the amygdala; Blair, 2010; Coccaro, McCloskey, Fitzgerald, & Phan, 2007; LeDoux, 2000) or regulation (e.g., the ACC and mPFC; Posner et al., 2011) might be important for negative emotional reactivity on a broader level.

The lack of unique variance in parent-report of sadness in the current study was somewhat more surprising, as I expected to find evidence for independent genetic influences on each emotion, consistent with phenotypic research suggesting different patterns of neural activation associated with anger, sadness and fear (Vytal & Hamann, 2010; Damasio et al., 2000). Indeed, when observed data was considered, anger was found to be primarily heritable whereas sadness was largely influenced by the common environment, and bivariate Cholesky decompositions suggested that observed anger and sadness were also largely distinct, with a small amount of shared genetic variance explaining the modest correlation between them. In contrast, social and object fear in the current study appeared to be explained by the same common environmental component, although fear as assessed in the Scary Mask episode was largely explained by episode-specific nonshared environmental variance, whereas fear in storytelling showed a moderate common environmental component, suggesting that some of the same aspects of children's fear were evoked by both episodes, but Storytelling was more successful in tapping into systematic individual differences. It is worth noting that Scary Mask did assess fear in a social context, and by middle childhood the most frightening aspect of that episode may have been the interaction with the stranger rather than the frightening mask itself. As a consequence, it may have been that fear in Scary Mask was closer to the social fear assessed in Storytelling than the fear of risky or frightening objects and circumstances assessed by the CBQ. Unfortunately, correlations between observed fear and both observed anger and sadness were not high enough to allow an examination of genetic and environmental variance. However, the finding that observed fear is not highly related to observed anger or sadness is consistent with previous findings for observed

data in preschoolers (Durbin, Klein, Hayden, Buckley, & Moerk, 2005; Dyson et al., 2012).

Implications

Taken as a whole, the primary implication of the findings is that the genetic and environmental underpinnings of temperament may be very different depending on the level of analysis, the reporter, and the context. The idea of emotion or temperament as something trait-like or constitutionally based is not meaningless, but it is complex and situation-dependent, perhaps to a greater degree than typically assumed by researchers. In the current study, the fact that a single, common set of primarily genetic influences is able to account for the covariance of parent-reported anger, sadness, and fear does support the validity of a broad, dispositional tendency toward negative emotionality. However, there was also evidence for heritable aspects of both anger and fear that could not be accounted for by this core set of genetic influences, as well as modest but significant common environmental influences on mother-report of anger and sadness which were not shared with fear. As such, what is common to all three negative emotions is meaningful, but it is not able to encompass the full extent of systematic, biologically-based individual differences in discrete emotions. Fear, in particular, appears to be less well-represented by broad measures of negative emotionality, which may better represent anger and particularly sadness in middle childhood.

In addition, an examination of genetic and environmental influences and phenotypic convergence across reporters and contexts suggests that considering proneness to anger, sadness and fear at the trait level may overlook important contextual influences. Despite evidence for considerable shared genetic variance across mother- and

father-report for all emotions, findings suggest that measures intended to assess the same negative emotions may in some cases reflect very different aspects of children's temperament and emotional responding. Although examination of temperament at the cross-emotion or cross-reporter level is useful for arriving at a consistent, reliable measure (Rothbart & Bates, 2006), some questions might benefit from a narrower, more emotion-specific or even more context-specific approach.

For instance, there is a considerable body of literature examining relations between temperament and risk for psychopathology, with evidence for both broad and specific mechanisms (e.g., low effortful control as a risk for multiple disorders versus anger-proneness as a risk for conduct problems), but much of this research has still taken place at the level of broad negative affectivity or neuroticism (Nigg, 2006), and the maintenance of distinctions between proneness to anger, sadness or distress, and fear, may prove helpful in elucidating multiple biologically- or environmentally-influenced pathways to disorders. For example, consistent with neuroimaging research suggesting associations between reactivity to threat and reactive aggression (Blair, 2010) the combination of fear- and anger-proneness might be important for risk for high levels of reactive aggression (Siever, 2008; Nigg, 2006), to a greater degree than either fear or anger alone, whereas low levels of fear but not necessarily anger might be related to risk for instrumental aggression (Blair, Colledge, Murray, & Mitchell, 2001; Nigg, 2006). In addition, the lower representation of fear at the level of broad negative emotionality, compared to anger and sadness, is one possible reason why negative emotionality is often a stronger and more consistent predictor of externalizing than internalizing symptoms (Rothbart & Bates, 2006; Sanson et al., 2004), which are relatively more strongly related

to fear (Lemery et al., 2002; Oldehinkel et al., 2007; Muris & Ollendick, 2005). If so, measuring fear or withdrawal tendencies specifically, rather than broad negative emotionality or neuroticism, might be necessary in order to clarify conflicting findings relating temperament to risk for internalizing problems in childhood. However, in studies examining molecular genetic variants or physiological mechanisms, genes associated with systems that might be expected to influence a wide variety of emotional and behavioral responses, such serotonergic functioning (Lesch & Merschdorf, 2000; Rothbart & Bates, 2006; Spoont, 1992), might be best considered in relation to broader negative emotionality.

Limitations

The current study is subject to several limitations, which should be taken into account when interpreting findings. First, it is important to note that due to a limited number of observed episodes per emotion and relatively low correlations between episodes, it was not always possible to aggregate data across multiple episodes. Aggregation of observed data has been found to increase both reliability and convergence with parent-report (Eaton, 1994; Forman et al., 2003), and is likely to be important for arriving at a measure of temperament which is more comparable to parent-report. A related limitation was that it was not possible to create anger and sadness composites from separate episodes, leading to the possibility that covariance between observed anger and sadness was partially due to shared episode context. Finally, the use of multiple coders per episode is a potential disadvantage, as it may have reduced reliability, and the use of fewer coders for each episode would be advantageous in future research.

The generalizability of the current sample is also limited, as participants were largely Caucasian and middle class, with a majority of participants reporting at least some college education. Although the demographic characteristics of the sample are representative for Wisconsin, it is important to note that estimates of heritability are specific to the population under consideration, sensitive to relative genetic and environmental heterogeneity, and uninformative about the origin or malleability of differences in a trait across populations or generations (Lemery-Chalfant, 2010). As such, it will be important to replicate findings in more diverse samples. Two other important questions specific to twin research are the extent to which twin findings can be generalized to singletons, and whether or not the environments experienced by MZ twins are more similar to the environments experienced by DZ twins in a way that is relevant to the phenotype of interest (i.e., does the Equal Environments Assumption hold). There is evidence that infant twins do not differ from unrelated singletons on temperament according to the IBQ (Goldsmith & Campos, 1990), and adult twins do not differ from singleton siblings on personality according to the MPQ (Johnson, Krueger, Bouchard, & McGue, 2002). In addition, there is some research suggesting that the Equal Environments Assumption does hold for adult personality and temperament, as greater similarity in childhood experiences appears to be unrelated to greater similarity in personality, phenotypic differences in infant temperament relate to actual rather than perceived zygosity, and statistical tests suggest no difference in means, standard deviations, or distributions of raw difference scores between MZ and DZ twins on infant temperament (Borkenau et al., 2002; Goldsmith et al., 1999; Morris-Yates et al., 1990).

Finally, although the twin design enables researchers to parse out genetic and environmental influences to a degree that cannot be matched by non-genetically informed research, this design rests on several assumptions and holds certain intrinsic limitations. Because quantitative genetic designs do not measure genetic or environmental variance directly, but instead approximate these latent constructs by measuring phenotypic similarities and differences, violation of the assumptions of the twin design can lead to biases in the estimation of genetic or environmental variance. For instance, any environmental influences which act to increase the similarity of individuals who are more genetically related, relative to those who share a lower percentage of genes, will contribute to the additive genetic component in a quantitative genetic model. In addition, when assortative mating occurs on the basis of a genetically influenced trait, DZ twins and full siblings will share more than the expected 50% of segregating genes, leading to a violation of the twin model assumptions and an overestimation of common environmental variance. A third limitation of the classic twin design is that common environmental variance and dominant genetic variance cannot be estimated within the same model, because without the inclusion of data from other relatives, these two parameters are confounded. Unaccounted for dominance will drive up MZ correlations relative to DZ correlations, and thus will lead to an overestimation of additive genetic variance at the expense of common environmental variance. The inclusion of data from relatives other than twins (e.g. parents, cousins) would resolve this problem and allow for the simultaneous modeling of C and D components, but without such data, unaccounted for dominance in ACE models is a limitation that must be kept in mind when interpreting results.

Finally, the classic twin design assumes that phenotypic variance can be partitioned into a linear combination of independent additive genetic, dominant genetic, common environmental and unique environmental components, with negligible gene-environment interaction or correlation. In reality, both gene-environment interaction and gene-environment correlation are likely to be common (Rutter et al., 2006; Shanahan & Hofer, 2005). Gene-environment interaction with family level factors (e.g. SES) will be attributed to additive genetic variance, because this type of interaction will act to increase rather than decrease the differences among genetically non-identical individuals who share the same environment. In contrast, gene-environment interaction with factors unique to the individual will act to decrease the resemblance of both MZ and DZ twins, and thus will be attributed to unique environmental variance. Thus, while the A component does provide a relatively good estimate of the broad genetic influence on a trait, including additive and dominant influences and the effects of gene-environment interaction, small or nonsignificant estimates of C should not be taken as evidence for the irrelevance of family-level influences on child development.

Future Directions

There are several ways that future research might build on current findings. First, in light of conceptual and empirical distinctions between fear, anger and sadness, it is worth considering what aspects of these three emotions might be influenced by a single set of shared genetic factors. It may be that there is a heritable tendency towards high reactivity to aversive, threatening or surprising stimuli in general, which might manifest as anger, sadness or fear depending on the specifics of the context, as exemplified by Rothbart and colleagues' (2001) negative emotionality or a lower physiological threshold

of responsivity (Kagan & Fox, 2006). However, it is also important to consider the possible contribution of genetic influences on effortful control or other aspects of self-regulation, which might be expected to modulate children's reactive temperament across a range of emotional and behavioral tendencies (Eisenberg et al., 2010; Rothbart & Bates, 2006). Negative emotionality and effortful control are distinct dimensions of temperament, but they are negatively correlated (Rothbart et al., 2001), and might be expected to share genetic variance with anger, fear and sadness (e.g.,Gagne & Goldsmith, 2011).

However, as noted earlier, the current findings also highlight the need for a greater consideration of narrow as well as broad dimensions of temperament, perhaps especially when examining relations between temperament and psychopathology. For instance, it is possible to draw a distinction between irritability, seen as an aversive reaction to stimulation, and frustration, seen as a response to blocked resources and goals (Deater-Deckard & Wang, 2012). Although both irritability and frustration overlap considerably with the emotional response of anger and may not be entirely distinct, it may be that there are genetic influences on irritability that are shared with sadness and fear but not frustration, whereas frustration may be more related to approach orientation and impulsivity. A differentiation between approach-related anger and anger in response to threatening or overstimulating experiences may be informative for examining different pathways to risk for externalizing problems, particularly if these aspects of anger are found to be genetically or environmentally distinct.

Finally, rather than attempting to increase the similarity between parent-report and observed measures, perhaps it would be useful to take advantage of what each does

well. Parents may be more successful at generalizing their children's emotional reactivity across a variety of situations (Rothbart & Bates, 2006), whereas observational measures are able to capture a child's in-the-moment reactions embedded within a particular context which parents might not commonly observe (such as a child's response to criticism, or interactions with peers), and thus may allow for drawing distinctions (e.g. between social and nonsocial anger) that parents might not consider or notice. This is not to argue that parent-report and observational measures of temperament cannot be used to measure the same construct. There are some contexts in which parents are better predictors of children's behaviors, with some evidence for higher maternal accuracy in predicting toddlers' behavior in lab-TAB episodes that require children to explore a novel environment, relative to distress and frustration episodes (Kiel & Buss, 2006), and as noted earlier, convergence is increased when multiple observed episodes are aggregated (Eaton, 1994; Forman et al., 2003). However, the importance of the context in shaping children's observed reactions might best be seen not only as a source of discrepancies between observational and reported measures, but as a potential advantage in its own right. A more systematic exploration of context effects in future research might offer a valuable window into children's reactivity and regulation across a variety of real world situations, in which the same emotions may have very different meanings, adaptive functions, and implications for later development.

REFERENCES

Ahadi, S., Rothbart, M. K., & Ye, R. (2003). Children's temperament in the US and China: Similarities and differences. *European Journal of Personality, 7*, 359-377.

Akaike, H. (1987). Factor analysis and AIC. *Psykometrika, 52*(3), 317-332.

Alcaro, A., Huber, R., & Panksepp, J. (2007). Behavioral functions of the mesolimbic dopaminergic system: An affective neuroethological perspective. *Brain Research Reviews, 56*(2), 283-321.

Allport, G. W. (1961). *Pattern and growth in personality*. New York: Holt, Rineheart, and Winston.

Alink, L. R. A., van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., Mesman, J., Juffer, F. & Koot, H. M. (2008). Cortisol and externalizing behavior in children and adolescents: Mixed meta-analytic evidence for the inverse relation of basal cortisol and cortisol reactivity with externalizing behavior. *Developmental Psychobiology, 50*(5), 427-450.

Amstadter, A. B., Daughters, S. B., MacPherson, L., Reynolds, E. K., Danielson, C. K., Wang, F. … Lejuez, C. W. (2012). Genetic associations with performance on a behavioral measure of distress intolerance. *Journal of Psychiatric Research, 46*, 87-94.

Ashby, F. G., Isen, A. M. & Turken, A. U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review, 106*. 529-550.

Baker, J. K., Fenning, R. M., & Crnic, K. A. (2011). Emotion socialization by mothers and fathers: Coherence among behaviors and associations with parent attitudes and children's social competence. *Social Development, 20*(2), 412-430.

Baker, B. L., & Heller, T. L. (1996). Preschool children with externalizing behaviors: Experience of fathers and mothers. *Journal of Abnormal Child Psychology, 24*(4), 513-532.

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

Bates, J. E., Pettit, G. S., Dodge, K. A., & Ridge, B. (1998). The interaction of temperamental resistance to control and restrictive parenting in the development of externalizing behavior. *Developmental Psychology, 34*, 982-995.

Battaglia, M., Ogliari, A., Zanoni, A., Citterio, A., Pozzoli, U., Giorda, R. … Marino, C.

(2005). Influence of the serotonin transporter promoter gene and shyness on children's cerebral responses to facial expressions. *Archives of General Psychiatry, 62*, 85-94.

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

Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and reformulation. *Psychological Bulletin, 106*(1), 59-73.

Berridge, K. C. (2003). Comparing the emotional brains of humans and other animals. In R. J. Davidson, H. H. Goldsmith, & K. Scherer (Eds.), *Handbook of affective sciences* (pp. 25-51). Oxford: Oxford University Press.

Bezdjian, S., Baker, La. A., & Tuvblad, C. (2011). Genetic and environmental influences on impulsivity: A meta-analysis of twin, family and adoption studies. *Clinical Psychology Review, 31*, 1209-1223.

Baker, B. L., & Heller, T. L. (1996). Preschool children with externalizing behaviors: Experience of fathers and mothers. *Journal of Abnormal Child Psychology, 24*(4), 513-532.

Blaire, C., Peters, R., & Granger, D. (2004). Physiological and neuropsychological correlates of approach/withdrawal tendencies in preschool: Further examination of the behavioral inhibition system/behavioral activation system scales for young children. *Developmental Psychobiology, 45*(3), 113-124.

Boker, S. M., Neale, M. C., Maes, H. H., Wilde, M. J., Spiegel, M., Brick, T. R. … Fox, J. (2011). OpenMx: An Open Source Extended Structural Equation Modeling Framework. *Psychometrika, 76*(2), 306-317.

Boomsma, D. I., van Baal, G. C., & Koopmans, J. R. (1999). A religious upbringing reduces the influence of genetic factors on disinhibition: Evidence for interaction between genotype and environment on personality. *Twin Research, 2*, 115-125.

Borkenau, P., Riemann, R., Angleitner, A., & Spinath, F. M. (2002). Similarity of childhood experiences and personality resemblance in monozygotic and dizygotic twins: A test of the equal environments assumption. *Personality and Individual Differences, 33*, 261-269.

Braumgart, J. M., Plomin, R., DeFries, J. C., & Fulker, D. W. (1992). Genetic influence on tester-rated infant temperament as assessed by Bayley's Infant Behavior Record: Nonadoptive and adoptive siblings and twins. *Developmental Psychology, 28*(1), 40-47.

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience and Biobehavioural Review, 26*, 809-817.

Buhrmester, D., Camparo, L., Christensen, A., Gonzalez, L. S., & Hinshaw, P. (1992). Mothers and fathers interacting in dyads and triads with normal and hyperactive sons. *Developmental Psychology, 28*(3), 500-509.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*, 261-304.

Buss, A. H. & Plomin, R. (1975). *A temperament theory of personality development*. New York: Wiley.

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

Buss, K. A., Schumacher, J. R., Dolski, I., Kalin, N. H., Goldsmith, H. H., & Davidson, R. J. (2003). Right frontal brain activity, cortisol, and withdrawal behavior in 6-month old infants. *Behavioral Neuroscience, 117*(1), 11-20.

Button, T. M., Maughan, B., & Eley, T. C. (2008). Punitive discipline, negative life events and gene-environment interplay in the development of externalizing behavior. *Psychological Medicine, 38*, 29-39.

Calkins, S. D., Blandon, A. Y., Williford, A. P. & Keane, S. P. (2007). Biological, behavioral and relational levels of resilience in the context of risk for early childhood behavior problems. *Development and Psychopathology, 19*, 675-700.

Canli, T., Sivers, H., Whitfield, S. L., Gotlib, I. H., & Gabrieli, J. D. (2002). Amygdala response to happy faces as a function of extraversion. *Science, 296*, 2191.

Carter, C. S. (1986). The reproductive and adrenal systems. In M. G. Coles, E. Donchin, & S. E. Porges (Eds.), *Psychophysiology: Systems, processes and applications* (pp. 172-182). New York: Guildford Press.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin, 135*(2), 183-204.

Caspi, A., Henry, B., McGee, R. O., Moffitt, T. E., & Silva, P. A. (1995). Temperamental origins of child and adolescent behavior problems: From age three to age fifteen. *Child Development, 66*(1), 55-68.

Caspi, A. (2000). The child is the father of the man: Personality continuities from childhood to adulthood. *Journal of Personality and Social Psychology, 78*(1), 158-172.

Caspi, A. & Shiner, R. L. (2006). Personality Development. In W. Damon & N. Eisenberg (Eds.), *Handbook of Child Psychology: Vol. 3. Social, Emotional, and Personality Development. Vol. 6* (pp. 300-365). New York: Wiley.

Caspi, A., & Silva, P. A. (1995). Temperament qualities at age three predict personality traits in young adulthood: Longitudinal evidence from a birth cohort. *Child Development, 66,* 486–498.

Cassano, M., & Perry-Parrish, C. (2007). Influence of gender on parental socialization of children's sadness regulation. *Social Development, 16*(2), 210-231.

Cassano, M. C., and Zeman, J. L. (2010). Parental socialization of sadness regulation in middle childhood: The role of expectations and gender. *Developmental Psychology, 46*(5), 1214-1226.

Calzada, E.J., Eyberg, S. M., Rich, B., & Querido, J. G. (2004). Parenting disruptive preschoolers: Experiences of mothers and fathers. *Journal of Abnormal Child Psychology, 32*(2), 203-213.

Ceci, S. J., Barnett, S. M., & Kanaya, T. (2003). Developing child proclivities into adult competencies: The overlooked multiplier effect. *Psychology of Abilities, Competencies and Expertise, 1,* 70-92.

Chen, J., Lipska, B. K., Halim, N., Ma, Q. D., Matsumoto, M., Melhem, S. … Weinberger, D. R. (2004). Functional analysis of genetic variation in catechol-*O*-methyltransferase (COMT): Effects on mRNA, protein, and enzyme activity in postmortem human brain. *American Journal of Human Genetics, 75*, 807-821.

Christensen, A., Margolin, G., & Sullaway, M. (1992). Interparental agreement on child behavior problems. *Psychological Assessment, 4*(4), 419-425.

Coccaro, E. F., McCloskey, M. S., Fitzgerald, D. A., & Phan, K. L. (2007). Amygdala and orbitofrontal reactivity to social threat in individuals with impulsive aggression. *Biological Psychiatry, 62*, 168-178.

Compas, B. E., Connor-Smith, J., & Jaser, S. S. (2004). Temperament, stress-reactivity, and coping: Implications for depression in childhood and adolescence. *Journal of Clinical Child and Adolescent Psychology, 33*(1), 21-31.

Crocket, L. J., Moilanen, K. L., Raffaelli, M. & Randall, B. A. (2006). Psychological profiles and adolescent adjustment: A person-centered approach. *Development and Psychopathology, 18*, 195-214.

Da Silva Alves, F., Schmitz, N., Figee, M., Abeling, N., Hasler, G., van der Meer, J. …

van Amelsvoort, T. (2011). Dopaminergic modulation of the human reward system: A placebo-controlled dopamine depletion fMRI study. *Journal of Psychopharmacology, 25*, 538-549.

Dahl, R. E. (2009). Adolescent brain development: A period of vulnerabilities and opportunities. *Annals of the New York Academy of Sciences, 1021*, 1-22.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience, 3*(10), 1049-1056.

Davidov, M. & Grusec, J. E. (2006). Untangling the links of parental responsiveness to distress and warmth to child outcomes. *Child Development, 77*(1), 44-58.

Davidson, R. J. (2002). Anxiety and affective style: Role of prefrontal cortex and amygdala. *Biological Psychiatry, 51*, 68-80.

De Haan, M., Gunnar, M. R., Tout, K., Hart, J., & Stansbury, K. (1997). Familiar and novel contexts yield different associations between cortisol and behavior among 2-year-old children. *Developmental Psychobiology, 33*(1), 93-101.

DeYoung, C. G. & Gray, J. R. (2009). Personality neuroscience: Explaining individual differences in affect, behavior, and cognition. In P. J. Corr & G. Matthews (Eds.), *The Cambridge Handbook of Personality Psychology* (pp. 323-346). New York: Cambridge University Press.

Deater-Deckard, K., Beekman, C., Wang, Z., Kim, J., Petrill, S. A., Thompson, L. A., & DeThorne, L. S. (2010). Approach/positive anticipation, frustration/anger, and overt aggression in childhood. *Journal of Personality, 78*(3), 991-1010.

Deater-Deckard, K., Petrill, S. A., & Thompson, L. A. (2007). Anger/frustration, task persistence, and conduct problems in childhood: A behavioral genetic analysis. *Journal of Childhood Psychology and Psychiatry, 48*(1), 80-87.

Degnan, K. A., Calkins, S. D., Keane, S. P., & Hill-Soderlund, A. L. (2008). Profiles of disruptive behavior across early childhood: Contributions of frustration reactivity, physiological regulation, and maternal behavior. *Child Development, 79*(5), 1357-1376.

Deater-Deckard, K., & Wang, Z. (2012). Anger and irritability. In M. Zentner & R. L. Shiner (Eds.), *Handbook of Temperament* (pp. 124-144). New York: Guilford.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences, 22*, 491-569.

Derryberry, D., Reed, M. A., & Pilkenton, C. (2003). Temperament and coping: Advantages of an individual differences perspective. *Development and Psychopathology, 15,* 1049-1066.

Dettling, A. C., Parker, S. W., Lane, S., Sebanc, A., & Gunnar, M. R. (2000). Quality of care and temperament determine changes in cortisol concentrations over the day for young children in childcare. *Psychoneuroendocrinology, 25,* 819-836.

DiLalla, L. F., Kagan, J., & Reznick, J. S. (1994). Genetic etiology of behavioral inhibition among 2-year-old children. *Infant Behavior and Development, 17,* 405-412.

Diener, M. L., & Kim, D. Y. (2004) Maternal and child predictors of preschool children's social competence. *Applied Developmental Psychology, 25,* 3-24.

Ding, Y.C., Chi, H.C., Grady, D.L., Morishima, A., Kidd, J. R., Kidd, K. K., … Moyzis, R. K. (2002). Evidence of positive selection acting at the human dopamine receptor D4 gene locus. *Proceedings of the National Academy of Sciences, 99,* 309–314.

Donzella, B., Gunnar, M. R., Krueger, W. K., & Alwin, J. (2000). Cortisol and vagal tone responses to competitive challenge in preschoolers: Associations with temperament. *Developmental Psychobiology, 37*(4), 209-220.

Duhig, A. M., Renk, K., Epstein, M. K., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice, 7*(4), 435-453.

Durbin, E. C., Klein, D. N., Hayden, E. P., Buckley, M. E., & Moerk, K. C. (2005). Temperamental emotionality in preschoolers and parental mood disorders. *Journal of Abnormal Psychology, 114*(1), 28-37.

Durbin, C. E. & Wilson, S. (2011). Convergent validity of and bias in maternal reports of child emotion. *Psychological Assessment,* doi: 10.1037/a0026607

Dyson, M. W., Olino, T. M., Durbin, E. C., Goldsmith, H. H., & Klein, D. N. (2012). The structure of temperament in preschoolers: A two-stage factor analytic approach. *Emotion, 12*(1), 44-57.

Eaton, W. O. (1994). Temperament, development, and the Five-Factor Model: Lessons from activity level. In C. F. Halverson Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The Developing Structure of Temperament and Personality from Infancy to Adulthood* (pp. 173–187). Hillsdale, NJ: Erlbaum.

Eaves, L., Heath, A., Martin, N., Maes, H., Neale, M., Kendler, K. … Corey, L. (1999). Comparing the biological and cultural inheritance of personality and social attitudes in the Virginia 30 000 study of twins and their relatives. *Twin Research, 2*, 62-80.

Eisenberg, N., Cumberland, A., Spinrad, T., Fabes, R. A., Shepard, S. A., Reiser, M. … Guthrie, I. K. (2001). The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. *Child Development, 72*(4), 1112-1134.

Eisenberg, N., Sadovsky, A., Spinrad, T. L., Fabes, R. A., Losoya, S. H., Valiente, C. … Shepard, S. A. (2005). The relations of problem behavior status to children's negative emotionality, effortful control, and impulsivity: Concurrent relations and predictions of change. *Developmental Psychology, 41*(1), 193-211.

Eisenberg, N., Spinrad, T. L., Fabes, R. A., Reiser, M., Cumberland, A., Shepard, S. R. … Thompson, M. (2004). The relations of effortful control and impulsivity to children's resiliency and adjustment. *Child Development, 75*(1), 25-46.

Eisenberg, N., Valiente, C., Spinrad, T. L., Cumberland, A., Liew, J., Reiser, M., . . . Losoya, S. H. (2009). Longitudinal relations of children's effortful control, impulsivity, and negative emotionality to their externalizing, internalizing, and co-occurring behavior problems. *Developmental Psychology, 45*(4), 988-1008.

Eisenberg, N., Spinrad, T. L., & Eggum, N. D. (2010). Emotion-related self-regulation and its relation to children's maladjustment. *Annual Review of Clinical Psychology, 6*(1), 495-525.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*(3/4), 169-200.

Emde, R. N., Robinson, J. L., Corley, R. P., Nikkari, D., & Zahn-Waxler, C. (2001). Reactions to restraint and anger-related expressions during the second year. In R. N. Emde, & J. K. Hewitt (Eds.), *Infancy to Early Childhood: Genetic and Environmental Influences on Developmental Change* (pp. 127-140). New York: Oxford University Press.

Emde, R. N., Plomin, R., Robinson, J., Corley, R., DeFries, J., & Fulker, D. W. … Zahn-Waxler, C. (1992). Temperament, emotion, and cognition at fourteen months: The MacArthur Longitudinal Twin Study. *Child Development, 63*, 1437-1455.

Evans, D. M., Gillespie, N. A., & Martin, N. G. (2002). Biometrical genetics. *Biological Psychology, 61*, 33-51.

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U., Church, J. A., Miezin, F. M.

… Peterson, S. E. (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Computational Biology, 5*(5): e1000381. doi:10.1371/journal.pcbi.1000381

Feinberg, M. E., Button, T. M., Neiderhiser, J. M., Reiss, D., & Hetherington, M. (2007). Parenting and adolescent behavior and depression. *Archives of General Psychiatry, 64*, 457-465.

Fish, M., Stifter, C. & Belsky, J. (1991). Conditions of continuity and discontinuity in infant negative emotionality: newborn to five months. *Child Development, 62*, 1525-1537.

Fivush, R., Brotman, M. A., Buckner, J. P., & Goodman, S. H. (2000). Gender differences in parent-child emotion narratives. *Sex Roles, 42*, 233-253.

Forbes, E. E., Brown, S. M., Kimak, M., Ferrell, R. E., Manuck, S. B. & Hariri, A. R. (2009). Genetic variation in components of dopamine neurotransmission impacts ventral striatal reactivity associated with impulsivity. *Molecular Psychiatry, 14*, 60–70.

Forget-Dubois, N., Perusse, D., Turecki, G., Girard, A., Billette, J., Rouleau, G. … Tremblay, R. E. (2003). Diagnosing zygosity in infant twins: Physical similarity, genotyping, and chorionicity. *Twin Research, 6*, 479-485.

Forman, D. R., O'Hara, M. W., Larsen, K., Coy, K. C., Gorman, L. L., & Stuart, S. (2003). Infant emotionality: Observational methods and the validity of maternal reports. *Infancy, 4*, 541–565.

Fortunato, C. K., Dribin, A. E., Granger, D. A., & Buss, K. A. (2008) Salivary alpha-amylase and cortisol in toddlers: Differential relations to affective behavior. *Developmental Psychobiology, 50*(8), 807-818.

Fox, N. A., Nichols, K. E., Henderson, H. A., Rubin, K., Schmidt, L., Hamer, D., … Pine, D. S. (2005). Evidence for a gene-environment interaction in predicting behavioral inhibition in middle childhood. *Psychological Science, 16*(12), 921-926.

Frick, P. J., & Morris, A. S. (2004). Temperament and developmental pathways to conduct problems. *Journal of Child and Adolescent Psychology, 33*(1), 54-68.

Gagne, J. R., & Goldsmith, H. H. (2011). A longitudinal analysis of anger and inhibitory control in twins from 12 to 36 months of age. *Developmental Science, 14*(1), 112-124.

Gagne, J. R. & Saudino, K. J. (2010). Wait for it! A twin study of inhibitory control in

early childhood. *Behavior Genetics, 40*(3), 327-337.

Gagne, J. R., Van Hulle, C. A., Aksan, N., Essex, M. J., & Goldsmith, H. H. (2011). Deriving childhood temperament measures from emotion-eliciting behavioral episodes: Scale construction and initial validation. *Psychological Assessment, 23*(2), 337-353.

Gagne, J. R., Vendlinski, M. K., & Goldsmith, H. H. (2009). The genetics of childhood temperament. In K. Yong-Kyu (Ed.), *The Handbook of Behavior Genetics* (pp. 251-267). New York: Springer.

Ganiban, J. M., Saudino, K. J., Ulbricht, J., Neiderhiser, J. M., & Reiss, D. (2008). Stability and change in temperament during adolescence. *Journal of Personality and Social Psychology, 95*(1), 222-236.

Gao, W., Zhu, H., Giovanello, K. S., Smith, K. J., Shen, D., Gilmore, J. H., & Lin, W. (2009). Evidence on the emergence of the brain's default network from 2-week-old to 2-year-old healthy pediatric subjects. *Proceedings of the National Academy of Sciences, 106*(16), 6790-6795.

Gartstein, M. A., & Marmion, J. (2008). Fear and positive affectivity in infancy: Convergence/discrepancy between parent-report and laboratory-based indicators. *Infant Behavior and Development, 31*, 227-238.

Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the Revised Infant Behavior Questionnaire. *Infant Behavior and Development, 26*, 64-86.

Gilliom, M., & Shaw, D. S. (2004). Codevelopment of externalizing and internalizing problems in early childhood. *Development and Psychopathology, 16*(2), 313-333.

Gjone, H., & Stevenson, J. (1997). A longitudinal twin study of temperament and behavior problems: Common genetic or environmental influences? *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(10), 1448-1456.

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

Goldsmith, H. H. (1991). A zygosity questionnaire for young twins: A research note. *Behavior Genetics, 21*, 257-269.

Goldsmith, H. H., Buss, K. A., & Lemery, K. S. (1997). Toddler and childhood temperament: Expanded content, stronger genetic evidence, new evidence for the importance of environment. *Developmental Psychology, 33*, 891–905.

Goldsmith, H. H., Campos, J. J. (1986). Fundamental issues in the study of early

temperament: The Denver Twin Temperament Study. In M. E. Lamb, & A. Brown (Eds.), *Advances in Developmental Psychology* (pp. 231-283). Hillsdale, NJ: Erlbaum.

Goldsmith, H. H., & Campos, J. J. (1990). The structure of temperamental fear and pleasure in infants: A psychometric perspective. *Child Development, 61*(6), 1944-1964.

Goldsmith, H. H., & Gottesman, I. I. (1981). Origins of variation in behavioral style: A longitudinal study of temperament in young twins. *Child Development, 52*(1), 91-103.

Goldsmith, H. H., & Lemery, K. L. (2000). Linking temperamental fearfulness and anxiety symptoms: A behavior-genetic perspective. *Biological Psychiatry, 48*(12), 1199-1209.

Goldsmith, Lemery, K. S., Buss, K. A., & Campos, J. J. (1999). Genetic analyses of focal aspects of infant temperament. *Developmental Psychology, 35*(4), 972-985.

Goldsmith, H. H., Reilly, H. H., Lemery, K. L., Longley, S., & Prescott, A. (2001). *The Laboratory Temperament Assessment Battery: Middle Childhood Version.* Unpublished manuscript, University of Wisconsin-Madison.

Goldsmith, H. H., & Rothbart, M. K. (1999). *Laboratory Temperament Assessment Battery (Technical Manual).* Madison, WI: University of Wisconsin-Madison.

Gunnar, M. R., Wewerka, S., Frenn, K., Long, J. D., & Griggs, C. (2009). Developmental changes in hypothalamus-pituitary-adrenal activity over the transition to adolescence: Normative changes and associations with puberty. *Development and Psychopathology, 21*, 69-85.

Gunnar, M. R., Larson, M. C., Hertsgaard, L., Harris, M. L., & Brodersen, L. (1992). The stressfulness of separation among nine-month-old infants: Effects of social context variables and infant temperament. *Child Development, 63*, 290-303.

Gunnar, M. R., & Vazquez, D. M. (2001). Low cortisol and a flattening of expected daytime rhythm: Potential indices of risk in human development. *Development and Psychopathology, 13*, 515-538.

Hagekull, B., & Bohlin, G. (1998). Preschool temperament and environmental factors related to the five-factor model of personality in middle childhood. *Merrill-Palmer Quarterly, 44*, 194–215.

Hankin, B. L., Nederhoff, E., Oppenheimer, C. W., Jenness, J., Young, J. F., Abela, J. R., … Oldehinkel, A. J. (2011). Differential susceptibility in youth: Evidence that 5-

HTTLPR x positive parenting is associated with positive affect 'for better and worse'. *Translational Psychiatry, 1*, e44. doi:10.1038/tp.2011.44

Harmon-Jones, C., Schmeichel, B. J., Mennitt, E., & Harmon-Jones, E. (2011). The expression of determination: Similarities between anger and approach-related positive affect. *Journal of Personality and Social Psychology, 100*(1), 172-181.

Harmon-Jones, E., & Allen, J. B. (1998). Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affect valence. *Journal of Personality and Social Psychology, 74*(5), 1310-1316.

Harmon-Jones, E., Harmon-Jones, C., Abramson, L., & Peterson, C. K. (2009). PANAS positive activation is associated with anger. *Emotion, 9*(2), 183-196.

Harmon-Jones, E., Vaugn-Scott, K., Mohr, S., Sigelman, J., & Harmon-Jones, C. (2004). The effect of manipulated sympathy and anger on left and right frontal cortical activity. *Emotion, 4*(1), 95-101.

Hayden, E. P., Klein, D. N., & Durbin, C. E. (2005). Parent reports and laboratory assessments of child temperament: A comparison of their associations with risk for depression and externalizing disorders. *Journal of Psychopathology and Behavioral Assessment, 27*(2), 89-100.

Hayden, E. P., Dougherty, L. R., Maloney, B., Durbin, C. E., Olino, T. M., Nurnberger, J. I. Jr, … Klein, D. N. (2007). Temperamental fearfulness in childhood and the serotonin transporter promoter region polymorphism: A multimethod association study. *Psychiatric Genetics, 17*, 135-142.

He, J., Xu, Q., & Degnan, K. A. (2012). Anger expression and persistence in young children. *Social Development, 21*(2), 233-353.

Heath, A. C., Berg, K., Eaves, L. J., Solaas, M. H., Sundet, J., Nance, W. E. … Magnus, P. (1985). No decline in assortative mating for education level. *Behavior Genetics, 15*(4), 349-369.

Hegvik, R. L., McDevitt, S. C., & Carey, W. B. (1982). The Middle Childhood Temperament Questionnaire. *Developmental and Behavioral Pediatrics, 3*(4), 197-200.

Heils, A., Teufel, A., Petri, S., Stober, G., Riederer, P., Bengel, D. & Lesch, K. P. (1996). Allelic variation of human serotonin transporter gene expression. *Journal of Neurochemistry, 66*(6), 2621-2624.

Heisler, L. K., Pronchuk, N., Nonogaki, K., Zhou, L., Raber, J., Tung, L. … Tecott, L. H. (2007). Serotonin activates the hypothalamic-pituitary-adrenal axis via serotonin

2C receptor stimulation. *Journal of Neuroscience, 27*(26), 6956-6964.

Hoekstra, R. A., Bartels, M., Hudziak, J. J., Van Beijsterveldt, T. C., & Boomsma, D. I. (2008). Genetic and environmental influences on the stability of withdrawn behavior in children: A longitudinal. Multi-informant twin study. *Behavior Genetics, 38*, 447-461.

Holmes, A. (2008). Genetic variation in cortico-amygdala serotonin function and risk for stress-related disease. *Neuroscience and Biobehavioral Reviews, 32*, 1293-1314.

Hwang, J., & Rothbart, M. K. (2003). Behavior genetics studies of infant temperament: Findings vary across parent-report instruments. *Infant Behavior and Development, 26*, 112–114.

Ivorra, J. L., Sanjuan, J., Jover, M., Carot, J. M., de Frutos, R. & Molto, M. D. (2010). Gene-environment interaction of child temperament. *Journal of Developmental and Behavioral Pediatrics, 31*(7), 545-554.

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science, 2*(3), 260-280.

Izard, C. E. (1972). *Patterns of emotions: A new analysis of anxiety and depression.* New York: Academic Press.

Jang, K. L., Dick, D. M., Wolf, H., Livesly, W. J. & Paris, J. (2005). Psychosocial adversity and emotional instability: An application of gene-environment interaction models. *European Journal of Personality, 19*, 359-372.

Johnson, D. L., Wiebe, J. S., Gold, S. M., Andreasen, N. C., Hichwa, R. D., Watkins, G. L., & Boles Ponto, L. L. (1999). Cerebral blood flow and personality: A positron emission tomography study. *The American Journal of Psychiatry, 156*(2), 252-257.

Johnson, W., & Krueger, R. F. (2005). Genetic effects on physical health: Lower at higher income levels. *Behavioral Genetics, 35*(5), 579-590.

Johnson, W., Krueger, R. F., Bouchard, T. J., & McGue, M. (2002). The personalities of twins: Just ordinary folks. *Twin Research, 5*(2), 125-131.

Kagan, J. (1992). Yesterday's premises, tomorrow's promises. *Developmental Psychology, 28*(6), 990-997.

Kagan, J. (2007). A trio of concerns. *Perspectives on Psychological Science, 2*(4), 361-376.

Kagan, J. & Fox, N. A. (2006). Biology, culture and temperamental biases. In W. Damon & N. Eisenberg (Eds.), *Handbook of Child Psychology: Vol. 3. Social, Emotional, and Personality Development. Vol. 6* (pp. 167-223). New York: Wiley.

Kagan, J., Reznick, J. S., & Snidman, N. (1987). The physiology and psychology of behavioral inhibition in children. *Child Development, 58*(6), 1459-1473.

Kagan, J., & Snidman, N. (1999). Early childhood predictors of adult anxiety disorders. *Biological Psychiatry, 46*, 1536-1541.

Karg, K., Burmeister, M., Shedden, K. & Sen, S. (2011). The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: Evidence of genetic moderation. *Archives of General Psychiatry, 68*(5), 444-454.

Keiley, M. K., Lofthouse, N., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2003). Differential risks of covarying and pure components in mother and teacher reports of externalizing and internalizing behaviors across ages 5 to 14. *Journal of Abnormal Child Psychology, 31*(3), 267-283.

Kendler, K. S., Aggen, S. H., Jacobson, K. C., & Neale, M. C. (2003). Does the level of family dysfunction moderate the impact of genetic factors on the personality trait of neuroticism? *Psychological Medicine, 33*, 817-825.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). A test of the equal-environment assumption in twin studies of psychiatric illness. *Behavior Genetics, 23*(1), 21-27.

Kiel, E. J., & Buss, K. A. (2006). Maternal accuracy in predicting toddlers' behaviors and associations with toddlers' fearful temperament. *Child Development, 77*(2), 355-370.

Kochanska, G., Coy, K. C., Tjebkes, T. L., & Husarek, S. J. (1998). Individual differences in emotionality in infancy. *Child Development, 69*(2), 375-390.

Kochanska, G. (1991). Socialization and temperament in the development of guilt and conscience. *Child Development, 62*, 1379–1392.

Kochanska, G. (1995). Children's temperament, mothers' discipline, and security of attachment: Multiple pathways to emerging internalization. *Child Development, 66*, 597–615.

Kochanska, G., Aksan, N., & Joy, M. E. (2007). Children's fearfulness as a moderator of parenting in early socialization: Two longitudinal studies. *Developmental Psychology, 43*(1), 222-237.

Krishnan, V., & Nestler, E. J. (2008). The molecular neurobiology of depression. *Nature, 455*(7215), 894-902.

Krueger, R. F., South, S., Johnson, W. & Iacono, W. (2008). The heritability of personality is not always 50%: Gene-environment interactions and correlations between personality and parenting. *Journal of Personality, 76*(6), 1485-1522.

Lachman, H. M., Papolos, D. F., Saito, T., Yu, Y. M., Szumlanski, C. L. & Weinshilboum, R. M. (1996). Human catechol-O-methyltransferase pharmacogenetics: Description of a functional polymorphism and its potential application to neuropsychiatric disorders.*Pharmacogenetics, 6*(3), 243-250.

Lara, D. R., & Askikal, H. S. (2006). Toward an integrative model of the spectrum of mood, behavioral and personality disorders based on fear and anger traits: II. Implications for neurobiology, genetics and psychopharmacological treatment. *Journal of Affective Disorders, 94*, 89-103.

Larsen, J. T.,Berntson, G. G., Poehlmann, K. M., Ito, T. A., & Cacioppo, J. T. (2008). The psychophysiology of emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *The Handbook of Emotions* (3rd Ed., pp. 180-195). New York: Guilford.

Laucht, M., Becker, K., Blomeyer, D. & Schmidt, M. H. (2007). Novelty seeking involved in mediating the association between the dopamine D4 receptor gene exon III polymorphism and heavy drinking in male adolescents: Results from a high-risk community sample. *Biological Psychiatry, 61*(1), 87-92.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23*, 155-184.

Leerkes, E. M., & Crockenberg, S. C. (2003). The impact of maternal characteristics and sensitivity on the concordance between maternal reports and laboratory observations of infant negative emotionality. *Infancy, 4*(4), 517-539.

Lemery, K. S., Essex, M. J., & Smider, N. A. (2002). Revealing the relation between temperament and behavior 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*(1), 189-204.

Lemery-Chalfant, K. (2006). Wisconsin Twin Panel: Current directions and findings. *Twin Research and Human Genetics, 9*(6), 1030-1037.

Lemery-Chalfant, K. L. (2010). How genes and environments work together to promote resilience. In J. Reich, A. J. Zautra, & J. S. Hall (Eds.), *Handbook of Adult Resilience* (pp. 55-78) New York: Guildford.

Lemery-Chalfant, K., Doelger, L., & Goldsmith, H. H. (2008). Genetic relations between effortful and attentional control and symptoms of psychopathology in middle childhood. *Infant and Child Development, 17*, 365-385.

Lemery-Chalfant, K., Kao, K., Swann, G., & Goldsmith, H. (2013). Childhood temperament: Passive gene-environment correlation, gene-environment interaction, and the hidden importance of the family environment. *Development and Psychopathology, 25*, 51-63.

Lemogne, C., Gorwood, P., Boni, C., Pessiglioni, M., Lehericy, S. & Fossati, P. (2011). Cognitive appraisal and life stress moderate the effects of the 5-HTTLPR polymorphism on amygdala reactivity. *Human Brain Mapping, 32*(11), 1856-1867.

Lengua, L. J., & Kovacs, E. A. (2005). Bidirectional associations between temperament and parenting and prediction of adjustment problems in middle childhood. *Applied Developmental Psychology, 26*, 21-38.

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*(1), 164-181.

Lesch, K. P. (2003). Neuroticism and serotonin: A developmental genetic perspective. In R. Plomin, J. C. DeFries, Craig, I. W., & McGuffin, P. (Eds.), *Behavioral Genetics in the Postgenomic Era* (pp. 389-423). Washington, DC: American Psychological Association.

Lesch, K. P., Balling, U., Gross, J., Strauss, K., Wolozin, B. L., Murphy, D. L. & Riederer, P. (1994). Organization of the human serotonin transporter gene. *Journal of Neural Transmission, 95*(2), 157-162.

Lesch, K. P., & Merschdorf, U. (2000). Impulsivity, aggression, and serotonin: A molecular psychobiological perspective. *Behavioral Sciences and the Law, 18*, 581-604.

Leve, L. D., Kim, H. K., & Pears, K. P. (2005). Childhood temperament and family environment as predictors of internalizing and externalizing trajectories from ages 5 to 17. *Journal of Abnormal Child Psychology, 33*(5), 505-520.

Locke, R. L., Davidson, R. J., Kalin, N. H., & Goldsmith, H. H. (2009). Children's context inappropriate anger and salivary cortisol. *Developmental Psychology,*

45(5), 1284-1297.

Lucki, I. (1998). The spectrum of behaviors influenced by serotonin. *Biological Psychiatry, 44*, 151-162.

Majdandzic, M., van den Boom, D. C., & Heesbeen, D. G. (2008). Peas in a pod: Biases in the measurement of sibling temperament? *Developmental Psychology, 44*(5), 1354-1368.

Mangelsdorf, S. C., Schoppe, S. J., & Burr, H. (2000). The meaning of parental reports: A contextual approach to the study of temperament and behavior problems in childhood. In V. J. Molfese, & D. L. Molfese (Eds.), *Temperament and Personality Development Across the Lifespan* (pp. 121-140). Mahwah, NJ: Erlbaum.

Manuck, S. B., Flory, J. D., McCaffrey, J. M., Matthews, K. A., Mann, J. J., & Muldoon, M. F. (1998). Aggression, impulsivity, and central nervous system serotonergic responsivity in a nonpatient sample. *Neuropsychopharmacology, 19*(4), 287-299.

Markon, K. E., Krueger, R. F. & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88*(1), 137-157.

Matheny, A. P., Jr. (1989). Children's behavioral inhibition over age and across situations: Genetic similarity for a trait during change. *Journal of Personality, 57,* 215–235.

Matheny, A. P. (1980). Bayley's Infant Behavior Record: Behavioral components and twin analyses. *Child Development, 51*(4), 1157-1167.

McClowry S. G., Hegvik, R., & Teglasi, H. (1993). An examination of the construct validity of the middle childhood temperament questionnaire. *Merrill-Palmer Quarterly, 39*, 279–293.

McElwain, N. L., Halberstadt, A. G., & Volling, B. L. (2007). Mother- and Father-reported reactions to children's negative emotions: Relations to young children's emotional understanding and friendship quality. *Child Development, 78*(5), 1407-1425.

McEwen, B. S. (1998). Stress, adaption and disease: Allostasis and allostatic load. *Annals of the New York Academy of Science, 840*, 33-44.

McGuire, S., Clifford, J., Fink, J., Basho, S., & McDonnell, A. M. (2003). Children's reactions to the unfamiliar in middle childhood and adolescence: An observational twin/sibling study. *Personality and Individual Differences, 35*, 339-354.

Mier, D., Kirsch, P. & Meyer-Lindenberg, A. (2010). Neural substrates of pleiotropic action of genetic variation in *COMT*: a meta-analysis. *Molecular Psychiatry, 15*, 918-927.

Moilanen, K. L., Shaw, D. S., Dishion, T. J., Gardner, F., & Wilson, M. (2010). Predictors of longitudinal growth in inhibitory control in early childhood. *Social Development, 19*(2), 326-347.

Moon, M., & Hoffman, C. D. (2008). Mothers' and fathers' differential expectancies and behaviors: Parent x child gender effects. *The Journal of Genetic Psychology, 164*(3), 261-279.

Morris, A. S., Silk, J. S., Steinberg, L., Sessa, F. M., Avenevoli, S., & Essex, M. J. (2002). Temperamental vulnerability and negative parenting as interacting predictors of child adjustment. *Journal of Marriage and Family, 64*, 461–471.

Morris-Yates, A., Andrews, G., Howie, P., & Henderson, S. (1990). Twins: A test of the equal environments assumption. *Acta Psychiatrica Scandinavica, 81*(4), 322-326.

Mullineaux, P., Deater-Deckard, K., Petrill, S., Thompson, L., & DeThorne, L. (2009). Temperament in middle childhood: A behavioral genetic analysis of fathers' and mothers' reports. *Journal of Research in Personality, 43*, 737-746.

Mun, E. Y., Fitzgerald, H. E., von Eye, A., Puttler, L. I., & Zucker, R. A. (2001). Temperamental characteristics as predictors of externalizing and internalizing child behavior problems in the contexts of high and low parental psychopathology. *Infant Mental Health Journal, 22*, 393–415.

Munafo, M. R., Brown, S. M. & Hariri, A. R. (2008). Serotonin transporter (5-HTTLPR) genotype and amygdala activation: A meta-analysis. *Biological Psychiatry, 63*, 852-857.

Muris, P., & Ollendick, T. H. (2005). The role of temperament in the etiology of child psychopathology. *Clinical Child and Family Psychology Review, 8*(4), 271-289.

Murphy, B. C., Shepard, S. A., Eisenberg, N., & Fabes, R. A. (2004). Concurrent and across time prediction of young adolescents' social functioning: The role of emotionality and regulation. *Social Development, 13*(1), 56–86.

Murphy, F. C., Nimmo-Smith, I., & Lawrence, A. D. (2003). Functional neuroanatomy of emotion: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience, 3*, 207–233.

Muthén, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38*, 171-189.

Muthén, L. K., & Muthén, B. O. (1998-2011). *Mplus User's Guide (Sixth Edition)*. Los Angeles, CA: Muthén & Muthén.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht, the Netherlands: Kluwer Academic.

Neale, M. C., Røysamb, E., & Jacobsen, K. (2006). Multivariate genetic analysis of sex-limitation and G x E interaction. *Twin Research and Human Genetics, 9*(4), 481-489.

Neppl, T. K., Donnellan, M. B., Scaramella, L. V., Widaman, K. F., Spilman, S. K., Ontai, L. L., & Conger, R. D. (2010). Differential stability of temperament and personality from toddlerhood to middle childhood. *Journal of Research in Personality, 44*, 386-396.

Nesse, R. M., & Ellsworth, P. C. (2009). Evolution, Emotions, and Emotional Disorders. *American Psychologist, 64*(2), 129-139.

Nigg, J. T. (2006). Temperament and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 47*(3/4), 395-422.

Oldehinkel, A. J., Hartman, C. A., Ferdinand, R. F., Verhulst, F. C., & Ormel, J. (2007). Effortful control as moderator of the association between negative emotionality and adolescents' mental health problems. *Development and Psychopathology, 19*, 523-539.

Oldehinkel, A. J., Hartman, C. A., de Winter, A. F., Veenstra, R., & Ormel, J. (2004). Temperamental profiles associated with internalizing and externalizing problems In preadolescence. *Development and Psychopathology, 16*, 421–440.

Pacheco, J, Beevers, C. G., Benavides, C., McGeary, J., Stice E. & Schnyer, D. M. (2009). Frontal-limbic white matter pathway associations with the serotonin transporter gene promoter region (5-HTTLPR) polymorphism. *The Journal of Neuroscience, 29*(19), 6229-6233.

Panksepp, J., & Watt, D. (2011). What is basic about basic emotions? Lasting lessons from affective neuroscience. *Emotion Review, 3*(4), 387-396.

Panksepp, J. (1993). Neurochemical control of moods and emotions: Amino acids to neuropeptides. In M. Lewis&J.M. Haviland (Eds.), *Handbook of Emotions* (pp. 87–107). New York: Guilford Press.

Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *Behavioral and Brain Sciences, 5*, 407–467.

Panksepp, J. (2010). Affective neuroscience of the emotional BrainMind: evolutionary perspectives and implications for understanding depression. *Dialogues in Clinical Neuroscience, 12*(4), 533-545.

Paris, J. (2005). Neurobiological dimensional models of personality: A review of the models of Cloninger, Depue, and Diever. *Journal of Personality Disorders, 19*, 156-170.

Perez-Edgar, K., & Fox, N. A. (2005). Temperament and Anxiety Disorders. *Child and Adolescent Psychiatric Clinics of North America, 14*(4), 681-706.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage, 16*, 331-348.

Pleck, J. H., & Masciadrelli, B. P. (2004). Parental involvement in U.S. Residential fathers: Levels, sources, and consequences. In M. E. Lamb (Ed.), *The Role of the Father in Child Development* (4th Edition, pp. 222-271). New York: Wiley.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioral Genetics* (5th ed.). New York: Worth Publishers.

Porges, S. W. (2007). The polyvagal perspective. *Biological Psychology, 74*, 116-143.

Posner, M. I. & Rothbart, M. K. (2009). Toward a physical basis of attention and self-regulation. *Physics of Life Reviews, 6*(2), 103-120.

Posner, M. I., Rothbart, M. K., Sheese, B. E. & Voelker, P. (2011). Control networks and neuromodulators of early development. *Developmental Psychology,* doi: 10.1037/a0025530

Presley, R., & Martin, R. P. (1994). Toward a structure of preschool temperament: Factor structure of the temperament assessment battery for children. *Journal of Personality, 62*, 415–448.

Price, T. S., Freeman, B., Craig, I. W., Petrill, S. A., Ebersole, L., & Plomin, R. (2000). Infant zygosity can be assigned by parental report questionnaire data. *Twin Research, 3*, 129–133.

Purcell S. (2002). Variance components models for gene–environment interaction in twin analysis. *Twin Research, 5*, 554–571.

Putnam, S. P., Ellis, L. K., & Rothbart, M. K. (2001). The structure of temperament from infancy through adolescence. In A. Eliasz & A. Angleitner (Eds.), *Advances in Research in Temperament* (pp. 165-182. Germany: Pabst Science Publishers.

Putnam, S. P., Rothbart, M. K. & Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood and early childhood. *Infant and Child Development, 17*, 387-405.

Rettew, D. C., Althoff, R. R., Dumenci, L., Ayer, L. & Hudziak, J. J. (2008). Latent profiles of temperament and their relations to psychopathology and wellness. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(3), 273-281.

Rhee, S. H., Corley, R. P., Friedman, N. P., Hewitt, J. K., Hink, L. K., Johnson, D. P. … Young, S. E. (2012). The etiology of observed negative emotionality from 14 to 24 months. *Frontiers in Genetics, 3*(9), doi:10.3389/fgene.2012.00009

Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin, 112*(3), 485-499.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 3–25.

Robinson, J., Kagan, J., Reznick, J. S., & Corley, R. (1992). The heritability of inhibited and uninhibited behavior: A twin study. *Developmental Psychology, 28*(6), 1030-1037.

Roisman, G., & Fraley, R. C. (2006). The limits of genetic influence: A behavior-genetic analysis of infant-caregiver relationship quality and temperament. *Child Development, 77*(6), 1656-1667.

Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development, 52*, 569–578.

Rothbart, M. K. (2011). *Becoming Who We Are: Temperament, Personality and Development*. New York: Guilford.

Rothbart, M. K., Ahadi, S. A., Hershey, K., & Fisher, P. (2001). Investigations of temperament at 3 to 7 years: The children's behavior questionnaire. *Child Development, 72,* 1394–1408.

Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 105-176). New York: Wiley.

Rothbart, M. K. & Bates, J. E. (2006). Temperament. In W. Damon & N. Eisenberg (Eds.), *Handbook of Child Psychology: Vol. 3. Social, Emotional, and Personality*

Development. Vol. 6 (105-176). New York: Wiley.

Rothbart, M. K., Derryberry, D., & Posner, M. I. (1994). A psychobiological approach to the development of temperament. In J. E. Bates & T. D. Wachs (Eds.), *Temperament: Individual differences at the interface of biology and behavior* (pp. 83–116). Washington, DC: American Psychological Association.

Rothbart, M. K., Evans, D. E., & Ahadi, S. A. (2000). Temperament and personality: Origins and Outcomes. *Journal of Personality and Social Psychology, 78*(1), 122-135.

Rothbart, M. K., & Goldsmith, H. H. (1985). Three approaches to the study of infant temperament. *Developmental Review, 5*, 237–260.

Roussos, P., Giakoumaki, S. G., & Bitsios, P. (2009). Cognitive and emotional processing in high novelty seeking associated with the L-DRD4 genotype. *Neuropsychologia, 47*, 1654-1659.

Rowe, D. C., & Plomin, R. (1977). Temperament in early childhood. *Journal of Personality Assessment, 41*(2), 150-156.

Russell, A., Hart, C. H., Robinson, C. C., & Olsen, S. F. (2003). Children's sociable and aggressive behavior with peers: A comparison of the U.S. and Australia, and contributions of temperament and parenting styles. *International Journal of Behavioral Development, 27*, 74–86.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Rutter, M., Moffitt, T. E. & Caspi, A. (2006). Gene-environment interplay and psychopathology: Multiple varieties but real effects. *Journal of Child Psychology and Psychiatry, 47*(3/4), 226-261.

Rydell, A., Berlin, L., & Bohlin, G. (2003). Emotionality, emotion regulation, and adaptation among 5-to-8-year-old children. *Emotion, 3*, 30-47.

Sallquist, J. V., Eisenberg, N., Spinrad, T. L., Reiser, M., Hofer, C., Liew, J. … Eggum, N. (2009). Positive and negative emotionality: Trajectories across six years and relations with social competence. *Emotion, 9*(1), 15-28.

Sanson, A., Hemphill, S. A., & Smart, D. (2004). Connections between temperament and social development: A review. *Social Development, 13*, 142–170.

Sanson, A., Oberklaid, F., Pedlow, R., & Prior, M. (1991). Risk indicators: Assessment of infancy predictors of preschool behavioural maladjustment. *Journal of Child*

Psychology and Psychiatry, 32, 609–626.

Sanson, A. & Prior, M. (1999). Temperament and behavioural precursors to oppositional defiant disorder and conduct disorder. In H. C. Quay & A. E. Hogan (Eds.), *Handbook of Disruptive Behaviour Disorders* (pp. 397–417). New York: Kluwer Academic.

Sanson, A. V., Smart, D. F., Prior, M., Oberklaid, F., & Pedlow, R. (1994). The structure of temperament from 3 to 7 years: Age, sex, and sociodemographic inf luences. *Merrill-Palmer Quarterly, 40*, 233–252.

Saudino, K. J. (2003). Parent ratings of infant temperament lessons from twin studies. *Infant Behavior and Development, 26*, 100–107.

Saudino, K. J. (2005) Behavioral genetics and child temperament. *Journal of Developmental and Behavioral Pediatrics, 26*(3), 214-223.

Saudino, K. J., McGuire, S., Reiss, D., Hetherington, E. M., & Plomin, R. (1995). Parent ratings of EAS temperaments in twins, full siblings, half-siblings, and step siblings. *Journal of Personality and Social Psychology, 68*(4), 723-733.

Saudino, K. J., Plomin, R., & DeFries, J. C. (1996). Tester-rated temperament at 14, 20 and 24 months: Environmental change and genetic continuity. *British Journal of Developmental Psychology, 14*, 129-144.

Saudino, K. J., & Eaton, W. O. (1991). Infant temperament and genetics: An objective twin study of motor activity level. *Child Development, 62*, 1167-1174.

Saudino, K. J. (2009). Do different measures tap the same genetic influences? A multi-method study of activity level in young twins. *Developmental Science, 12*(4), 626-633.

Scarantino, A., & Griffiths, P. (2011). Don't give up on basic emotions. *Emotion Review, 3*(4), 444-454.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype-environment effects. *Child Development, 54*, 242–435.

Schmidt, L. A., Fox, N. A., Rubin, K. H., Sternberg, E. M., Gold, P. W., Smith, C. C., & Schulkin, J. (1997). Behavioral and neuroendocrine responses in shy children. *Developmental Psychobiology, 30*, 127–140.

Schmidt, L. A. Fox, N. A., Schulkin, J., & Gold, P. W. (1999). Behavioral and psychophysiological correlates of self-presentation in temperamentally shy children. *Developmental Psychobiology, 35*, 119–135.

Schoppe-Sullivan, S. J., Kotila, L. E., Jia, R., Lang, S. N., & Bower, D. J. (2013). Comparisons of levels and predictors of mothers' and fathers' engagement with their preschool-aged children. *Early Child Development and Care, 183*(3-4), 498-514.

Schwartz, C. E., Snidman, N., & Kagan, J. (1996). Early childhood temperament as a determinant of externalising behaviour in adolescence. *Development and Psychopathology, 8*, 527–537.

Schwartz, C. E., Snidman, N., & Kagan, J. (1999). Adolescent social anxiety as an outcome of inhibited temperament in childhood. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 1008–1015.

Seifer, R., Sameroff, A. J., Barrett, L. C., & Krafchuk, E. (1994). Infant temperament measured by multiple observations and mother report. *Child Development, 65*, 1478–1490.

Seifer, R., Sameroff, A., Dickstein, S., Schiller, M., & Hayden. (2004). Your own children are special: Clues to the sources of reporting bias in temperamental assessments. *Infant Behavior & Development, 27,* 323-341.

Shanahan, M. J., & Hofer, S. M. (2005). Social context in gene-environment interactions: Retrospect and prospect. *Journals of Gerontology, 60B*(Special Issue 1), 65-76.

Shiner, R. L., Buss, K. A., McClowry, S. G., Putnam, S. P., Saudino, K. J., & Ventner, M. (2012). What is temperament *now*? Assessing progress in temperament research on the twenty-fifth anniversary of Goldsmith et al. (1987).

Shiner. (1998). How shall we speak of children's personalities in middle childhood? A preliminary taxonomy. *Psychological Bulletin, 124*(3), 308-332.

Siever, L. J. (2008). Neurobiology of aggression and violence. *American Journal of Psychiatry, 165*, 429-442.

Silberg, J. L., San Miguel, V. F., Murelle, E. L., Prom, E., Bates, J. E., Canino, G. … Eaves, L. (2005). Genetic and environmental influences on temperament in the first year of life: The Puerto Rico Infant Twin Study (PRINTS). *Twin Research and Human Genetics, 8*(4), 328-336.

Singh, A. L. & Waldman, I. D. (2010) The etiology of associations between negative emotionality and childhood externalizing disorders. *Journal of Abnormal Psychology, 119*(2), 376-388.

South, S. C., & Krueger, R. F. (2008). Marital quality moderates genetic and

environmental influences on the internalizing spectrum. *Journal of Abnormal Psychology, 117*(4), 826-837.

Spoont, M. R. (1992). Modulatory role of serotonin in neural information processing: Implications for human psychopathology. *Psychological Bulletin, 112*, 330–350.

Stifter, C. A., Willoughby, M. T., Towe-Goodman, N., & The Family Life Project Key Investigators. (2008). Agree or agree to disagree? Assessing the convergence between parents and observers on infant temperament. *Infant and Child Development, 17*(4), 407-426.

Stocker, C. M., Richmond, M. K., Rhoades, G. K., & Kiang, L. (2007). Family emotional processes and adolescents' adjustment. *Social Development, 16*(2), 310-325.

Sugiura M, Kawashima R, Nakagawa M, Okada K, Sato T, Goto R, Sato K, Ono S, Schormann T, Zilles K, & Fukuda H. (2000). Correlation between human personality and neural activity in cerebral cortex. *Neuroimage, 11*, 541-546.

Tackett, J. L., Krueger, R. F., Iacono, W. G., & McGue, M. (2008). Personality in middle childhood: A hierarchical structure and longitudinal connections with personality in late adolescence. *Journal of Research in Personality, 42*, 1456-1462.

Tackett, J. L., Waldman, I. D., Van Hulle, C. A., & Lahey, B. (2011). Shared genetic influences on negative emotionality and major depression/conduct disorder comorbidity. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*(8), 818-827.

Tauscher, J., Bagby, M. R., Javanmard, M., Christensen, B. K., Kasper, S., & Kapur, S. (2001). Inverse relationship between serotonin 5-HT_{1A} receptor binding and anxiety: A [^{11}C]Way-100635 PET Investigation in healthy volunteers. *American Journal of Psychiatry, 158*, 1326-1328.

Thomas, A., & Chess, S. (1977). *Temperament and Development.* New York: New York University Press.

Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd Edition, pp. 114-137). New York: Guildford.

Tschann, J. M., Kaiser, P., Chesney, M. A., Alkon, A. & Boyce, W. T. (1996). Resilience and vulnerability among preschool children: Family functioning, temperament, and behavior problems. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 184-192.

Tunbridge, E. M., Harrison, P. J., & Weinberger, D. R. (2006). Catechol-*O*-methyltransferase, cognition, and psychosis: Val^{158}Met and beyond. *Biological Psychiatry, 60*, 141-151.

Turkheimer, E. (2000). Three laws of behavioral genetics and what they mean. *Current Directions in Psychological Science, 9*(5), 160-164.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science, 14*(6), 623-628.

Uher, R. & McGuffin, P. (2008). The moderation by the serotonin transporter gene of environmental adversity in the aetiology of mental illness: Review and methodological analysis. *Molecular Psychiatry, 13*, 131-146.

Uher, R. & McGuffin, P. (2010). The moderation by the serotonin transporter gene of environmental adversity in the etiology of depression: 2009 update. *Molecular Psychiatry, 15*, 18-22.

Valiente, C., Eisenberg, N., Smith, C. L., Reiser, M., Fabes, R. A., Losoya, S. … Murphy, B. C. (2003). The relations of effortful control and reactive control to children's externalizing problems: A longitudinal assessment. *Journal of Personality, 71*(6), 1171-1196.

van den Acker, A. L., Decovic, M., Prinzie, P., & Asscher, J. J. (2010). Toddlers' temperament profiles: Stability and relations to negative and positive parenting. *Journal of Abnormal Child Psychology, 38*, 485-495.

Van Tol, H. H., Wu, C. M., Guan, H., Ohara, K., Bunzow, J. R., Civelli, O., … Jovanovic, V. (1992). Multiple dopamine D4 receptor variants in the human population. *Nature, 358*, 149-152.

Volbrecht, M. M., Lemery-Chalfant, K., Aksan, N., Zahn-Waxler, C., & Goldsmith, H. H. (2007). Explaining the familial link between positive affect and empathy development in the second year. *Journal of Genetic Psychology, 168*(2), 105-129.

Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Journal of Cognitive Neuroscience, 22*(12), 2864-2885.

Walker, A. J., & McGraw, L. A. (2000). Who is responsible for responsible fathering? *Journal of Marriage and Family, 62*(2), 563-569.

Wang, Z., Deater-Deckard, K., Cutting, L., Thompson, L. A., & Petrill, S. A. (2012). Working memory and parent-rated components of attention in middle childhood: A

behavior genetic study. *Behavior Genetics, 42*, 199-208.

Watson, D., Gamez, W., & Simms, L. J. (2005). Basic dimensions of temperament and their relation to anxiety and depression: A symptom-based perspective. *Journal of Research in Personality, 39*, 44-66.

Williams, L. J., & Holahan, P. J. (1994). Parsimony-based fit indices for multiple-indicator models: Do they work? *Structural Equation Modeling, 1*(2), 161-189.

Wood, A. C., Rijsdijk, F., Saudino, K. J., Asherson, P., & Kuntsi, J. (2008). High heritability for a composite index of children's activity level measures. *Behavior Genetics, 38*(3), 266-276.

Wright, A. W., Parent, J., Forehand, R., Edwards, M. C., Connors-Burrow, N. A., & Long, N. (2012). The relation of parent and child gender to parental tolerance of child disruptive behaviors. *Journal of Child and Family Studies.* Doi: 10.1007/s10826-012-9632-y

Yavich, L., Forsberg, M. M., Karayiorgou, M., Gogos, J. A., & Mannisto, P. T. (2007). Site-specific role of Catechol-O-Methyltransferase in dopamine overflow within prefrontal cortex and dorsal striatum. *Journal of Neuroscience, 27,* 10196–10202.

Yeung, W. J., Sandberg, J. F., Davis-Kean, P. E., & Hofferth, S. L. (2001). Children's time with fathers in intact families. *Journal of Marriage and Family, 63*(1), 136-154.

Zuckerman, M. (1995). Good and bad humors: Biochemical bases of personality and its disorders. *Psychological Science, 6*, 325–332.

Table 1

Summary of Observed Composites Formed in Each Episode

All parameters were standardized using Z-Score transformations prior to composite formation	
Only parameters **specific to** discrete emotions (e.g. anger vs. sadness) were used in this study.	
Impossibly Perfect Stars	
Parameters Used in Final Composite: Anger	Notes
Latency to first bodily or expressed anger, in seconds	Windsorized and reverse coded to form speed scores
Mean intensity expressed anger over 24 epochs (0-2 scale)	
Mean intensity bodily anger over 24 epochs (0-2 scale)	Skewness > 2; Square-root transformation used
Mean intensity vocal anger over 4 1 min. time blocks (0-2 scale)	Skewness > 2; Square-root transformation used
Parameters Used in Final Composite : Sadness	
Latency to first bodily or expressed sadness, in seconds	Windsorized and reverse coded to form speed scores
Mean intensity expressed Sadness over 24 epochs (0-2 scale)	Skewness > 2; Square-root transformation used
Mean intensity bodily Sadness over 24 epochs (0-2 scale)	Skewness > 2; Square-root transformation used
Mean intensity vocal Sadness over 4 1 min. time blocks (0-2 scale)	
Not Sharing	
Parameters Used in Final Composite: Anger	Notes
Latency to first bodily or expressed anger, in seconds	Windsorized and reverse coded to form speed scores
Latency to first vocal anger, in seconds	Windsorized and reverse coded to form speed scores
Mean intensity anger (combined expressed/bodily) over18 epochs (0-3 scale)	
Mean intensity vocal anger over 4 time blocks	Each block was 1/4th of the episode; variable length Skewness > 2; Square-root transformation used
Parameters Used in Final Composite: Sadness	Notes
Latency to first bodily or expressed sadness, in seconds	Windsorized and reverse coded to form speed scores
Mean intensity sadness (combined expressed/bodily) over18 epochs (0-3 scale)	
Transparent Box	
Parameters Used in Final Composite: Anger	Notes
Rough latency to first bodily or expressed anger, in seconds	Time in seconds until first epoch with anger intensity > 0; Windsorized and reverse coded to form speed scores
Mean intensity expressed anger over 24 epochs (0-2 scale)	Skewness > 2, kurtosis > 7; Square-root transformation used
Mean intensity bodily anger over 24 epochs (0-2 scale)	Skewness > 2, kurtosis > 7; Square-root transformation used
Parameters Used in Final Composite: Sadness	Notes
No sadness variables from this episode used (low frequency)	(Over 500/726 participants showed no sadness)
Wrong Gift	

Parameters Used in Final Composite: Anger	Notes
Mean intensity of expressed anger over 29 epochs (0-2 scale)	Skewness > 2; Square-root transformation used
Mean intensity of bodily anger over 29 epochs (0-2 scale)	Skewness > 2, kurtosis > 7; Square-root transformation used

Parameters Used in Final Composite: Sadness	Notes
Mean intensity of expressed sadness over 29 epochs (0-2 scale)	
Mean intensity of bodily sadness over 29 epochs (0-2 scale)	

Scary Mask

Parameters Used in Final Composite: Fear	Notes
Latency to first expressed or bodily fear, in seconds	Windsorized and reverse coded to form speed scores
Mean intensity of expressed fear over 6 epochs (0-3 scale)	
Mean intensity of bodily fear over 6 epochs (0-3 scale)	

Storytelling

Parameters used in final composite: Fear	Notes
Latency to first expressed or bodily fear, in seconds	Windsorized and reverse coded to form speed scores
Mean intensity of bodily fear over 42 epochs (0-3 scale)	
Expressed fear was not used (low frequency; 1009/1032 participants showed no expressed fear)	
Episode length varied with time child spent speaking; 80% of the sample completed the episode by 24 epochs	

Table 2

Final Cross-episode Composites for Anger, Sadness and Fear

Mean composite of anger across three episodes: Impossibly Perfect Stars (speed, mean intensity bodily, mean intensity expressed, mean intensity vocal) Not Sharing (speed to anger (expressed/bodily), speed to vocal anger, mean intensity anger (expressed/bodily), mean intensity vocal anger)
Transparent Box (Speed to anger (expressed/bodily), mean intensity expressed, mean intensity bodily)
Mean composite of sadness across two episodes: Impossibly Perfect Stars (speed, mean intensity bodily, mean intensity expressed, mean intensity vocal)
Not Sharing (speed to first expressed/bodily sadness, mean intensity expressed/bodily)
One episode-level object fear composite:
Scary Mask (speed, mean intensity expressed, mean intensity bodily)
One episode-level social fear composite:
Storytelling (speed, mean intensity bodily)

Note. A sadness composite for Wrong Gift was computed, but not used in final analyses due to low or negative cross-episode correlations

Table 3

Zero-order Correlations Among Episode-Level Composites

	1	2	3	4	5	6	7	8	9
1. Anger (Stars)	1	.285**	.311**	.095*	.239**	.074**	-.088**	.058*	.091**
2. Anger (Not Sharing)		1	.249**	.167**	.121**	.085**	.025	-.037	.041
3. Anger (Transparent Box)			1	.152**	.108**	.032	-.063	-.094*	.002
4. Anger (Wrong Gift)				1	.073*	.087**	.229**	-.155**	-.164**
5. Sadness (Stars)					1	.174**	.07*	.096**	.069*
6. Sadness (Not Sharing)						1	.067*	.018	.075**
7. Sadness (Wrong Gift)							1	-.187**	-.098**
8. Social Fear (Storytelling)								1	.262**
9. Object Fear (Scary Mask)									1

Note. * $p < .05$, ** $p < .001$.

Table 4

Means, Standard Deviations, Ranges, Skewness and Kurtosis

Scale	N	*M*	*SD*	Min	Max	Skewness	Kurtosis
Anger (mother)	1549	4.51	0.95	1.00	7.00	-0.30	0.15
Sadness (mother)	1549	3.91	0.79	1.40	6.27	-0.20	0.04
Fear (mother)	1549	3.88	0.99	1.38	6.70	0.04	-0.30
Shyness (mother)	1548	3.60	1.12	1.00	6.80	0.17	-0.51
Negative Emotionality (mother)	1549	4.10	0.70	1.38	6.22	-0.23	0.32
Anger (father)	1260	4.44	0.88	1.40	6.80	-0.20	-0.12
Sadness (father)	1260	3.83	0.71	1.40	5.97	-0.24	0.05
Fear (father)	1260	3.83	0.92	1.30	6.59	0.07	-0.29
Shyness (father)	1259	3.59	1.02	1.00	6.60	0.07	-0.32
Negative Emotionality (father)	1260	4.03	0.64	1.74	5.94	-0.26	0.32
Anger (observed)	1402	0.01	0.76	-1.83	2.33	0.48	0.11
Sadness (observed)	1401	0.00	0.78	-1.90	2.36	0.28	-0.125
Object Fear (observed)	1344	-0.01	0.77	-2.32	2.25	-0.61	0.73
Social Fear (observed)	1382	-0.01	0.86	-1.43	2.34	0.22	-1.07

Table 5

Zero-Order Correlations for Mother- and Father-report of Negative Emotions

	1	2	3	4	5	6	7	8	9	10
1. Anger (mother)	1	.560**	.256**	.098**	.622**	.356**	.152**	.054*	.778**	.493**
2. Sadness (mother)		1	.420**	.149**	.333**	.475**	.242**	.038	.820**	.447**
3. Fear (mother)			1	.226**	.102**	.200**	.575**	.132**	.738**	.397**
4. Shyness (mother)				1	.022	.058*	.138**	.634**	.205**	.098**
5. Anger (father)					1	.58**	.205**	.062*	.452**	.777**
6. Sadness (father)						1	.354**	.140**	.430**	.810**
7. Fear (father)							1	.240**	.426**	.707**
8. Shyness (father)								1	.100**	.196**
9. Negative Emotionality (mother)									1	.573**
10. Negative Emotionality (father)										1

Note. * $p < .05$, ** $p < .001$. Mother and Father Negative Emotionality are mean composites of the lower-order scales.

Table 6: Zero-Order Correlations for In-home Observation of Negative Emotions

	1.	2.	3.	4.
1. Anger (observed)	1	.206**	.096**	-.007
2. Sadness (observed)		1	.081**	.071**
3. Object Fear (Observed)			1	.262**
4. Social Fear (Observed)				1

Note. * $p < .05$, ** $p < .001$.

Table 7

Zero-Order Correlations for Observed and Parent Reported Negative Emotions

	Mother-report					Father-report				
	Anger	Sadness	Fear	Shyness	NE	Anger	Sadness	Fear	Shyness	NE
Anger	.049	0.00	-.07*	-.227*	-.012	.087*	.035	-.035	-.213*	-.036
Sadness	.063*	.033	.01	-.098*	.046	.073*	.079*	.05	-.065*	.085*
Object Fear	-.011	.065*	.04	.048	.035	.022	.028	.05	.091*	.044
Social Fear	-.05	.005	-.02	.012	-.03	-.04	-.039	-.022	.003	-.042

Note. * $p < .05$, ** $p < .001$.

Table 8:

Twin Intraclass Correlations

	Total		Male		Female		
Scale	MZ	DZ	MZ	DZ	MZ	DZ	OS DZ
Anger M	0.693	0.507	0.709	0.475	0.678	0.489	0.529
Sadness M	0.726	0.468	0.697	0.461	0.748	0.624	0.404
Fear M	0.756	0.309	0.77	0.234	0.743	0.411	0.313
Shyness M	0.674	0.02	0.657	-0.102	0.664	0.161	0.008
Anger F	0.709	0.366	0.772	0.219	0.646	0.447	0.404
Sadness F	0.727	0.438	0.727	0.398	0.723	0.566	0.396
Fear F	0.805	0.37	0.862	0.459	0.74	0.408	0.307
Shyness F	0.697	0.041	0.592	-0.103	0.745	0.071	0.075
NE M	0.773	0.527	0.774	0.508	0.778	0.554	0.53
NE F	0.809	0.449	0.874	0.447	0.751	0.492	0.431
Anger	0.473	0.253	0.488	0.213	0.419	0.412	0.212
Sadness	0.264	0.237	0.214	0.354	0.314	0.316	0.12
Object Fear	0.182	0.175	0.194	0.28	0.197	0.044	0.128
Social fear	0.484	0.386	0.439	0.348	0.505	0.377	0.419

Note. OS = Opposite Sex.

Table 9

Univariate AC/DE Results (Parent-report)

Scale	Model	-2LL	df	Δ -2LL	Δ df	p	AIC	A	C	D	E
					Mother-Report						
Anger	**ACE**	**3200.83**	**1296**	–	–	–	**608.83**	**0.45**	**0.26**	–	**0.29**
	AE	3212.3	1297	11.47	1	< .001	618.3				
	CE	3224.38	1297	23.55	1	< .001	630.38				
	E	3467.2	1298	266.37	2	< .001	871.2				
Sadness	**ACE**	**2735.37**	**1296**	–	–	–	**143.37**	**0.45**	**0.27**	–	**0.28**
	AE	2746.55	1297	11.18	1	< .001	152.55				
	CE	2759.46	1297	24.09	1	< .001	165.46				
	E	3013.49	1298	278.12	2	< .001	417.49				
Fear	ACE	3342.11	1296	–	–	–	750.11	0.74	.00	–	0.26
	ADE	3338.77	1296	–	–	–	746.77	0.44	–	0.31	0.25
	AE	**3342.11**	**1297**	**3.34**	**1**	**0.07**	**748.11**	**0.74**	–	–	**0.26**
	E	3561.64	1298	222.87	2	< .001	965.64				
Shyness	**ADE**	**3859.43**	**1296**	–	–	–	**1267.43**	**.00**	–	**0.68**	**0.32**
	AE	3897.59	1297	38.15	1	< .001	1303.59				
	E	3974.39	1298	114.95	2	< .001	1378.39				
					Father-Report						
Anger	ACE	2589.21	1088	–	–	–	413.21	0.73	.00	–	0.27
	ADE	2589.2	1088	–	–	–	413.2	0.71	–	0.02	0.27
	AE	**2589.21**	**1089**	**0.01**	**1**	**0.92**	**411.21**	**0.73**	–	–	**0.27**
	E	2770.14	1090	180.94	2	< .001	590.14				
Sadness	**ACE**	**2101.4**	**1088**	–	–	–	**-74.6**	**0.53**	**0.19**	–	**0.28**
	AE	2105.87	1089	4.47	1	0.03	-72.13				
	CE	2128.32	1089	26.92	1	< .001	-49.68				
	E	2321.43	1090	220.03	2	< .001	141.43				
Fear	ACE	2624.42	1088	–	–	–	448.42	0.8	.00	–	0.2
	ADE	2624.4	1088	–	–	–	448.4	0.77	–	0.03	0.2
	AE	2624.42	1089	0.02	1	0.89	446.42	0.8			0.2
	E	2873.52	1090	249.12	2	< .001	693.52				
Shyness	**ADE**	**3040.89**	**1088**	–	–	–	**864.89**	**.00**	–	**0.71**	**0.29**
	AE	3072.73	1089	31.84	1	< .001	894.73				
	E	3155.21	1089	114.33	2	< .001	975.21				

Note. The most parsimonious final model is indicated in bold. When both ACE and ADE models are fit, results for reduced models are reported in comparison to the better fitting full model.

Table 10

Univariate AC/DE Results (Observed)

Scale	Model	-2LL	df	Δ -2LL	Δ df	p	AIC	A	C	D	E
Anger	ACE	2652.21	1231	–	–	–	190.21	0.39	0.06	–	0.55
	AE	**2652.58**	**1232**	**0.37**	**1**	**0.54**	**188.58**	**0.47**	–	–	**0.53**
	CE	2660.1	1232	7.89	1	< .001	196.1				
	E	2730.44	1233	78.23	2	< .001	264.44				
Sadness	ACE	2831.78	1230	–	–	–	371.78	0.06	0.18		0.76
	AE	2834.29	1231	2.51	1	0.11	372.29				
	CE	**2831.94**	**1231**	**0.16**	**1**	**0.69**	**369.94**	–	**0.22**	–	**0.78**
	E	2862.82	1232	31.04	2	< .001	398.82				
Object Fear	ACE	3306.17	1178	–	–	–	950.17	0	0.17	–	0.83
	AE	3308.42	1179	2.25	1	0.13	950.42				
	CE	**3306.17**	**1179**	**0**	**1**	**1**	**948.17**	**0**	**0.17**	–	**0.83**
	E	3322.7	1180	16.52	2	< .001	962.7				
Social Fear	ACE	3290.52	1211	–	–	–	868.52	0.15	0.30	–	0.55
	AE	3299.6	1212	9.08	1	< .001	875.6				
	CE	**3291.86**	**1212**	**1.34**	**1**	**0.25**	**867.86**	**0**	**0.40**	–	**0.6**
	E	3398.78	1213	108.26	2	< .001	972.78				

Note. The most parsimonious final model is indicated in bold.

Table 11

Genetic and Environmental Contributions to Variance and Covariance Across Reporter

Model	-2LL	df	Δ -2LL	Δ df	p	AIC
Anger (Mother- and Father):						
Full	5351.79	2383	–	–	–	585.79
Final:	5355.71	2386	3.92	3	.27	583.71
Sadness (Mother- and Father):						
Full	4613.91	2383	–	–	–	-152.09
Final:	4618.45	2385	4.54	2	.10	-151.55
Fear (Mother- and Father): Full	5637.46	2383	–	–	–	871.46
Final:	5641.43	2386	3.97	3	.26	869.43
Shyness (Mother- and Father):						
Full	6409.16	2383	–	–	–	1643.14
Final:	6414.61	2385	5.47	2	.06	1644.61

Model	Scale	A1	C1	D1	E1	A2	C2	D2	E2
Full	Anger (M)	0.45	0.27	–	0.28				
	Anger(F)	0.52	0	–	0.05	0.22	0	–	0.21
Final	Anger (M)	**0.36**	**0.33**	**–**	**0.31**				
	Anger(F)	**0.74**	**–**	**–**	**0.03**	**–**	**–**	**–**	**0.23**
Full	Sadness (M)	.50	.22	–	.28				
	Sadness (F)	.29	.00	–	.03	.27	.16	–	.25
Final	Sadness (M)	**0.47**	**0.24**	**–**	**0.28**				
	Sadness (F)	**0.32**	**–**	**–**	**0.03**	**0.41**	**–**	**–**	**0.24**
Full	Fear (M)	.43	**–**	.32	.24				
	Fear (F)	.17	**–**	.16	.02	.47	**–**	.00	.18
Final	Fear (M)	**0.75**	**–**	**–**	**0.25**				
	Fear (F)	**0.33**	**–**	**–**	**0.02**	**0.47**	**–**	**–**	**0.18**
Full	Shyness (M)	.00	–	.69	.31				
	Shyness (F)	.00	–	.36	.06	.07	–	.29	.22
Final	Shyness (M)	**0.02**	**–**	**0.67**	**0.31**				
	Shyness (F)	**–**	**–**	**0.39**	**0.05**	**0.31**	**–**	**–**	**0.25**

Note. The most parsimonious final model is indicated in bold. AC/DE components of variance for each phenotype are standardized according to the total variance of that phenotype. AC/DE components of covariance are standardized according to the total variance of the second phenotype.

Table 12

Fit of Full and Most Reduced Cholesky Decompositions Across Emotion

Model		-2LL	df	Δ -2LL	Δ df	p	AIC
Trivariate Anger, Sadness and Fear	Full:	8700.5	3882	–	–	–	936.5
Mother-report	Final	8709.62	3887	9.11	5	0.1	935.62
Trivariate Anger, Sadness and Fear	Full:	6873.51	6873.51	–	–	–	357.51
Father-report	Final	6885.82	6885.82	12.31	8	0.14	353.82
Four-variable Anger, Sadness, Fear, Shyness	Full:	12539.12	5170	–	–	–	2199.12
Mother-report	Final	12556.36	5182	17.24	12	0.14	2192.36
Four-variable Anger, Sadness, Fear, Shyness	Full:	9857.14	4388	–	–	–	1181.14
Father-report	Final	9877.12	4353	19.98	15	0.17	1171.12
Bivariate Anger and Sadness	Full:	5444.5	2460	–	–	–	524.5
Observed	Final	5445.51	2464	1.17	4	0.88	517.67
Bivariate Object and Social Fear	Full:	6494.64	2388	–	–	–	1718.64
Observed	Final	6495.03	2393	0.39	5	1	1709.03

Table 13

Genetic and Environmental Contributions to Variance and Covariance Across Emotion

Model	Scale	A1	C1	E1	A2	C2	E2	A3	C3	E3
Full	Anger (M)	0.41	0.3	0.29						
	Sadness (M)	0.11	0.26	0.02	0.34	.00	0.27			
	Fear (M)	0.02	0.05	.00	0.08	.00	0.03	0.59	.00	0.23
Final	**Anger (M)**	**0.47**	**0.24**	**0.29**						
	Sadness (M)	**0.17**	**0.18**	**0.01**	**0.36**	**–**	**0.28**			
	Fear (M)	**0.09**	**–**	**–**	**0.1**	**–**	**0.02**	**0.55**	**–**	**0.24**
Full	Anger (F)	0.65	0.06	0.28						
	Sadness (F)	0.24	0.17	0.02	0.31	.00	0.26			
	Fear (F)	0.04	0.02	.00	0.08	.00	0.01	0.66	.00	0.19
Final	**Anger (F)**	**0.72**	**–**	**0.28**						
	Sadness (F)	**0.35**	**–**	**0.02**	**0.38**	**–**	**0.25**			
	Fear (F)	**0.05**	**–**	**–**	**0.12**	**–**	**–**	**0.63**	**–**	**0.2**
Full	Anger (O)	.41	.05	.53						
	Sadness (O)	.03	.10	.00	.02	.09	.76			
Final	**Anger (O)**	**0.47**	**–**	**0.53**						
	Sadness (O)	**0.09**	**–**	**–**	**–**	**0.16**	**0.75**			
Full	Object Fear (O)	.00	0.18	0.82						
	Social Fear (O)	0.07	0.37	.00	.00	.00	0.56			
Final	**Object Fear (O)**	**–**	**0.18**	**0.82**						
	Social Fear (O)	**–**	**0.41**	**–**	**–**	**–**	**0.59**			

Note. The most parsimonious final model is indicated in bold. AC/DE components of variance for each phenotype are standardized according to total variance of that phenotype. AC/DE components of covariance are standardized according to the total variance of the second phenotype.

Table 14

Genetic and Environmental Contributions to Variance and Covariance Across Emotion (Four-Variable Models)

Model	Scale	A1	C1	E1	A2	C2	E2	A3	C3	E3	A4	C4	E4
Full	Anger (M)	0.42	0.29	0.29									
	Sadness (M)	0.12	0.25	0.02	0.34	0	0.27						
	Fear (M)	0.02	0.05	0	0.08	0	0.03	0.59	0	0.23			
	Shyness (M)	0	0.01	0.01	0	0	0.02	0.03	0	0.01	0.54	0	0.38
Final	**Anger (M)**	**0.45**	**0.25**	**0.3**									
	Sadness (M)	**0.17**	**0.18**	**0.01**	**0.36**	**–**	**0.28**						
	Fear (M)	**0.09**	**–**	**–**	**0.1**	**–**	**0.02**	**0.55**	**–**	**0.24**			
	Shyness (M)	**–**	**0.01**	**–**	**–**	**–**	**0**	**0.1**	**–**	**–**	**0.51**	**–**	**0.4**
Model	Scale	A1	D1	E1	A2	D2	E2	A3	D3	E3	A4	D4	E4
Full	Anger (F)	0.71	0.01	0.28									
	Sadness (F)	0.35	0	0.01	0.35	0.02	0.25						
	Fear (F)	0.05	0	0	0.03	0.18	0.01	0.53	0	0.19			
	Shyness (F)	0	0.63	0.01	0.03	0	0	0.05	0	0.01	0	0	0.27
Final	**Anger (F)**	**0.71**	**0.01**	**0.28**									
	Sadness (F)	**0.36**	**–**	**0.01**	**0.38**	**–**	**0.25**						
	Fear (F)	**0.05**	**–**	**–**	**0.12**	**–**	**–**	**0.63**	**–**	**0.2**			
	Shyness (F)	**–**	**0.61**	**–**	**0.07**	**–**	**–**	**0.03**	**–**	**–**	**–**	**–**	**0.29**

Note. The most parsimonious final model is indicated in bold. AC/DE components of variance for each phenotype are standardized according to total variance of that phenotype. AC/DE components of covariance are standardized according to the total variance of the second phenotype.

Table 15

Change in Fit from Cholesky to Independent Pathway Model

Mother-Report				
Model	-2LL	Δ df	AIC	Δ AIC
Cholesky (full)	8700.50	–	936.5	–
Cholesky (final)	8709.62	5	935.62	-0.88
Independent Pathway (Full)	8700.64	0	936.64	0.14
Independent Pathway (Final)	8709.93	6	933.93	-2.57
Father-Report				
Model	-2LL	Δ df	AIC	Δ AIC
Cholesky (full)	6873.51	–	357.51	–
Cholesky (final)	6885.82	8	353.82	-3.69
Independent Pathway (Full)	6876.27	0	360.27	2.76
Independent Pathway (Final)	6889.34	9	355.34	-2.17

Note. Change in AIC is relative to the full Cholesky decomposition. Smaller AIC values indicate a better-fitting model. Negative AIC values indicate that the reduced model had the smaller AIC.

Table 16

Model Fit and Parameter Estimates for Independent Pathway and Common Pathway Model

Mother-report														
Model	-2LL	df	Δ -2LL	Δ df	p	AIC	Scale	Ac	Cc	Ec	Au	Cu	Eu	λ
IPM (Full)	8700.64	3882	–	–	–	936.64	Anger (M)	.10	.30	.02	.26	.00	.24	
							Sadness (M)	.45	.26	.29	.00	.00	.00	
							Fear (M)	.10	.04	.03	.59	.00	.23	
IPM (Final)	**8709.93**	**3888**	**9.29**	**6**	**0.16**	**933.93**	**Anger (M)**	**.17**	**.21**	**.01**	**.34**	**–**	**.27**	
							Sadness (M)	**.53**	**.17**	**.29**	**–**	**–**	**–**	
							Fear (M)	**.18**	**–**	**.03**	**.56**	**–**	**.24**	
CPM (Full)	8722.56	3886	21.92	4	< .001	950.56	NE (F)	.52	.33	.15				
							Anger (F)				.4	.00	0.24	.5511
							Sadness (F)				.00	.00	0.16	.7109
							Fear (F)				.58	.00	0.22	.4535
CPM (Final)	.8722.56	3890	21.92	8	0.01	942.56	NE (F)	0.52	0.33	0.15				
							Anger (F)				.4	–	.24	.5511
							Sadness (F)				–	–	.16	.7109
							Fear (F)				.58	–	.22	.4535
Father-report														
Model	-2LL	df	Δ -2LL	Δ df	p	AIC	Scale	Ac	Cc	Ec	Au	Cu	Eu	λ
IPM (Full)	6876.27	3258	–	–	–	360.27	Anger (F)	0.27	0.07	0.02	0.37	.00	0.27	
							Sadness (F)	0.57	0.16	0.27	.00	.00	.00	
							Fear (F)	0.1	0.03	0.01	0.67	.00	0.19	
IPM (Final)	6889.34	3267	13.07	9	0.16	355.34	Anger (F)	.33	–	.02	.38	–	.27	
							Sadness (F)	.74	–	.26	–	–	–	
							Fear (F)	.15	–	–	.65	–	.2	
CPM (Full)	6881.93	3262	5.67	4	0.23	357.93	NE (F)	.69	.18	.13				
							Anger (F)				.38	.00	.24	.5376
							Sadness (F)				.00	.00	.15	.6452
							Fear (F)				.67	.00	.19	.336
CPM (Final)	**6885.13**	**3267**	**8.87**	**9**	**0.45**	**351.13**	**NE (F)**	**.88**	**–**	**.12**				
							Anger (F)				**.38**	**–**	**.25**	**.5307**
							Sadness (F)				**–**	**–**	**.16**	**.6397**
							Fear (F)				**.68**	**–**	**.19**	**.332**

Note. The most parsimonious final model is indicated in bold. In the common pathway model, **λ** indicates factor loadings on the common negative emotionality factor

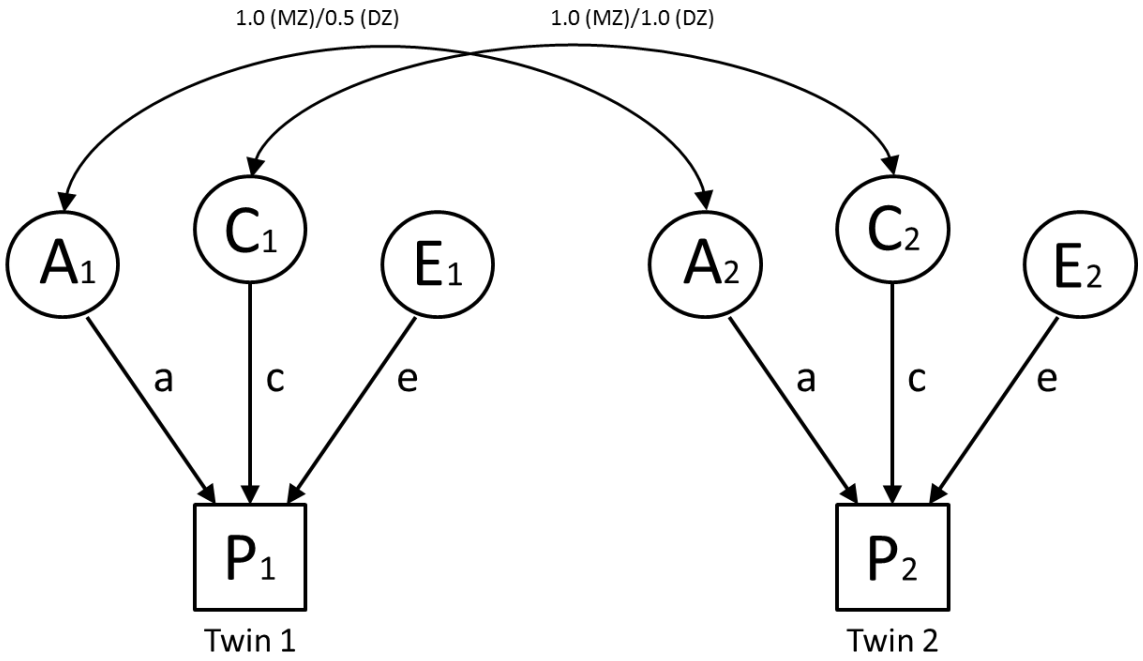

Figure 1. Example Univariate ACE model including both twins. P_1 and P_2 represent the scores of twin 1 and twin 2, respectively, on a measured phenotype of interest. Correlations between latent additive genetic factors (A_1 and A_2) are set to 1.0 for MZ twins and 0.5 for DZ twins. Correlations between latent common environmental factors (C_1 and C_2) are set to 1.0 for both MZ and DZ twins. Latent unique environmental factors (E_1 and E_2) are uncorrelated between twins.

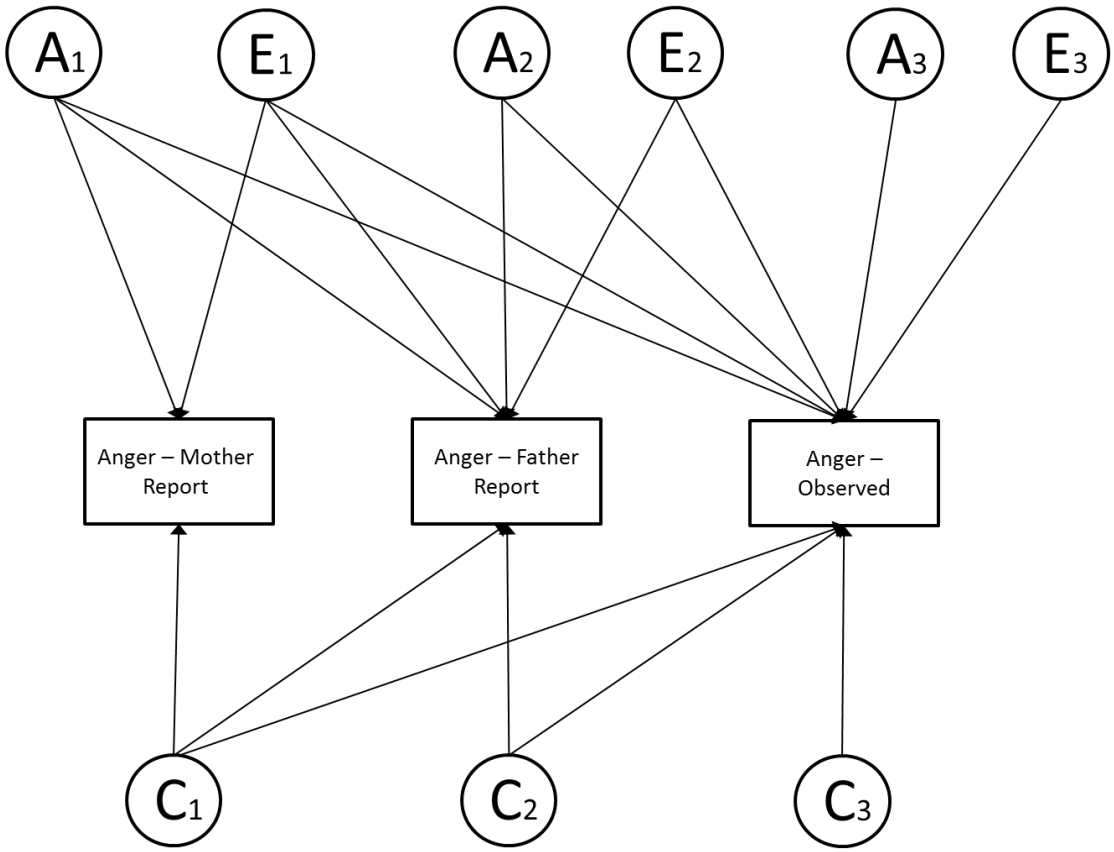

Figure 2. Example trivariate Cholesky decomposition, showing one twin only for simplicity. Latent A_1, C_1, and E_1 factors represent additive genetic, common environmental, and unique environmental influences on the first phenotype (here, mother-report of anger) , which may also be shared with the second and third phenotypes (father-report and in-home observation of anger). Latent A_2, C_2, and E_2 factors represent additive genetic, common environmental, and unique environmental influences on father-report of anger that are independent of mother-report but may be shared with in-home observation, and Latent A_3, C_3, and E_3 factors represent influences unique to observed anger.

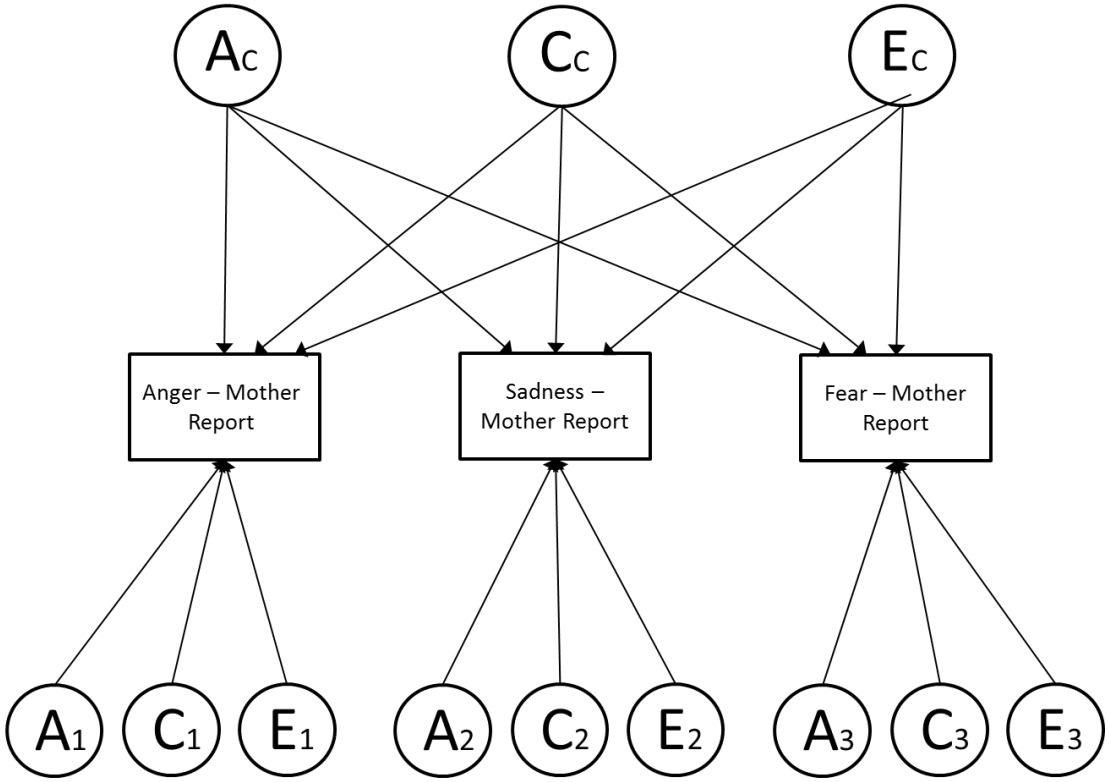

Figure 3. Example Independent pathway model, showing one twin only for simplicity. This model tests the assumption that covariance among a set of phenotypes (here, mother-report of anger, sadness and fear) can be fully accounted for by a single set of shared additive genetic, common environmental, and unique environmental factors, but allows the relative genetic and environmental influence on each phenotype to vary (e.g. fear might load more strongly on A_C and anger on C_C). Additive genetic, common environmental, and unique environmental influences unique to each phenotype are also estimated (e.g. A_1, C_1 and E_1 are unique to mother-report of anger).

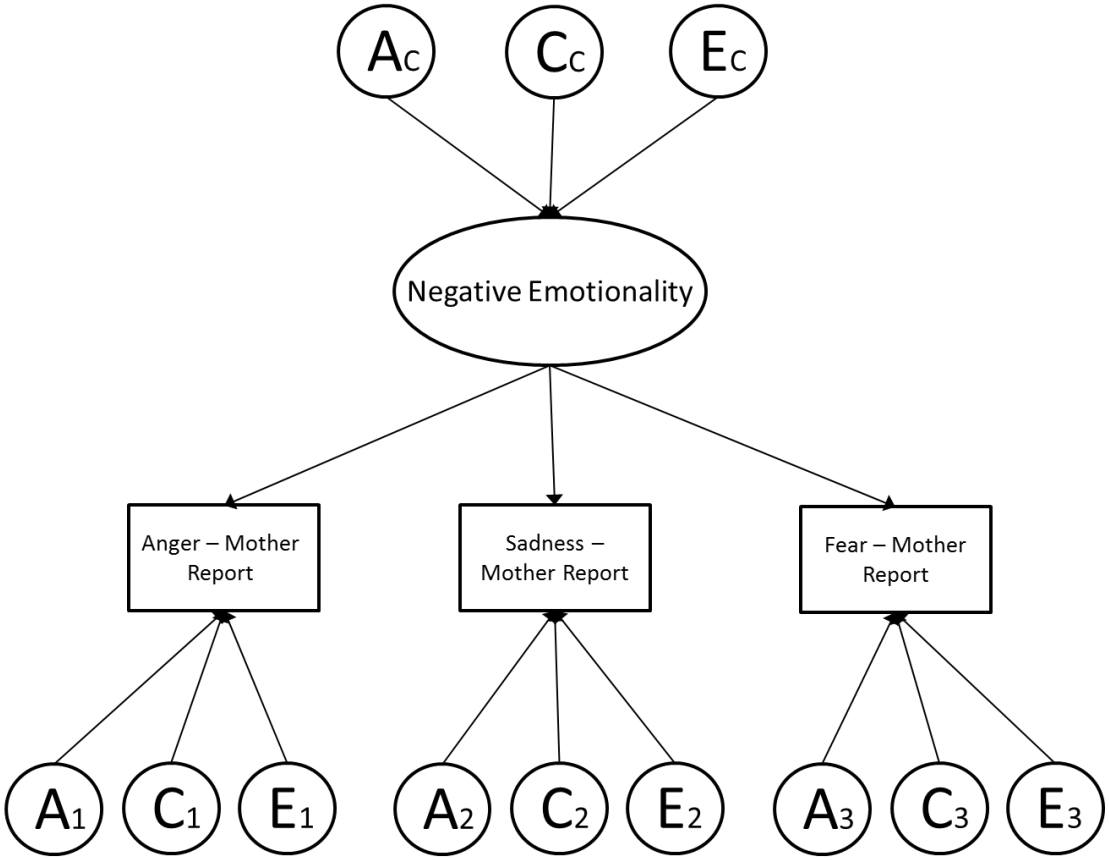

Figure 4. Example Common pathway model, showing one twin only for simplicity. This model tests the assumption that covariance among a set of phenotypes (here, mother-report of anger, sadness, and fear) can be fully accounted for by a single common phenotypic factor (e.g. negative emotionality), and decomposes the variance in the common factor into additive genetic, common environmental, and unique environmental factors. Additive genetic, common environmental, and unique environmental influences unique to each phenotype are also estimated (e.g. A_1, C_1 and E_1 are unique to mother-report of anger).

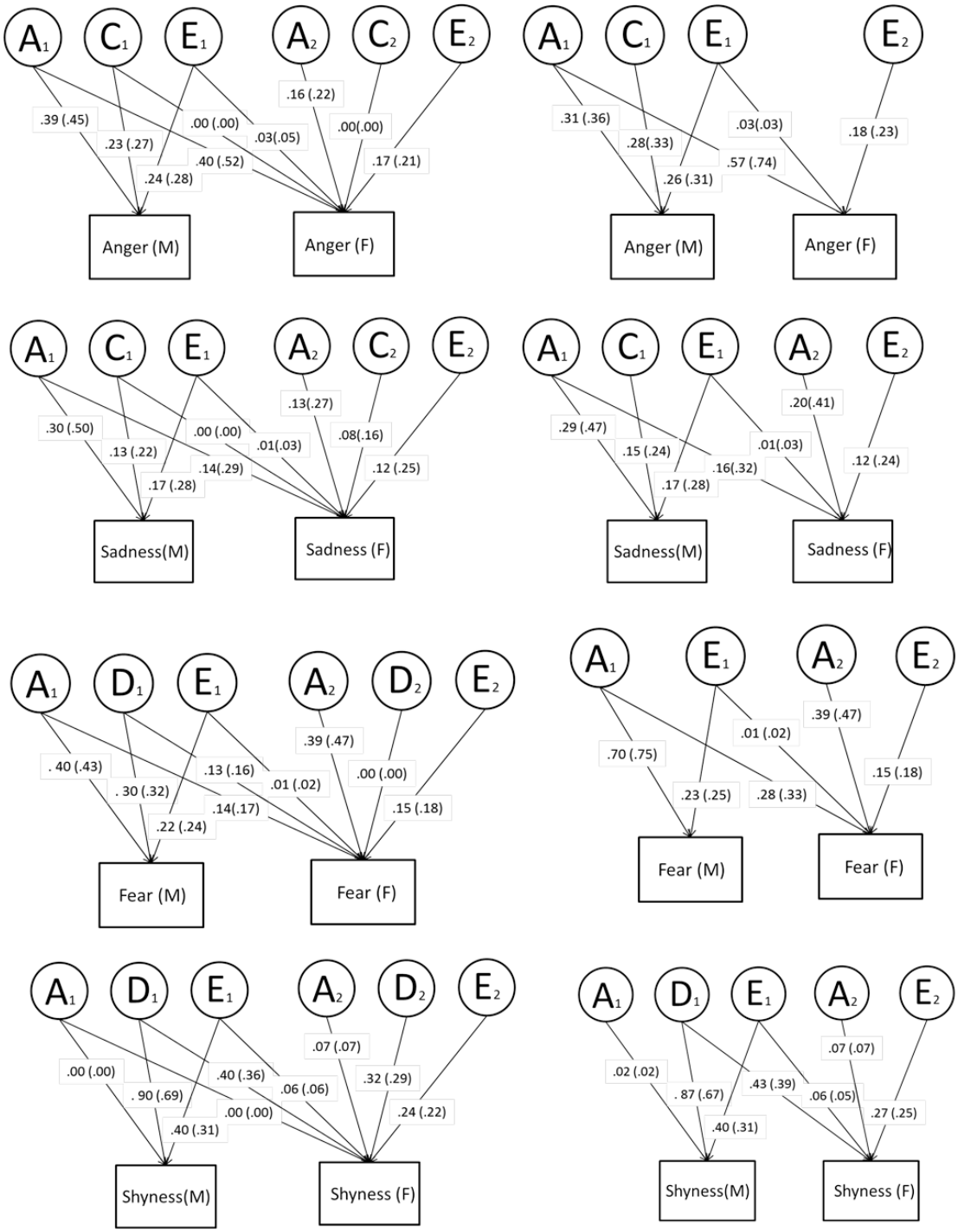

Figure 5. Full and final bivariate Cholesky decompositions across reporter for anger, sadness, fear and shyness. The full model is depicted on the right, and the final model is depicted on the left. Standardized A, C and E components are in parentheses.

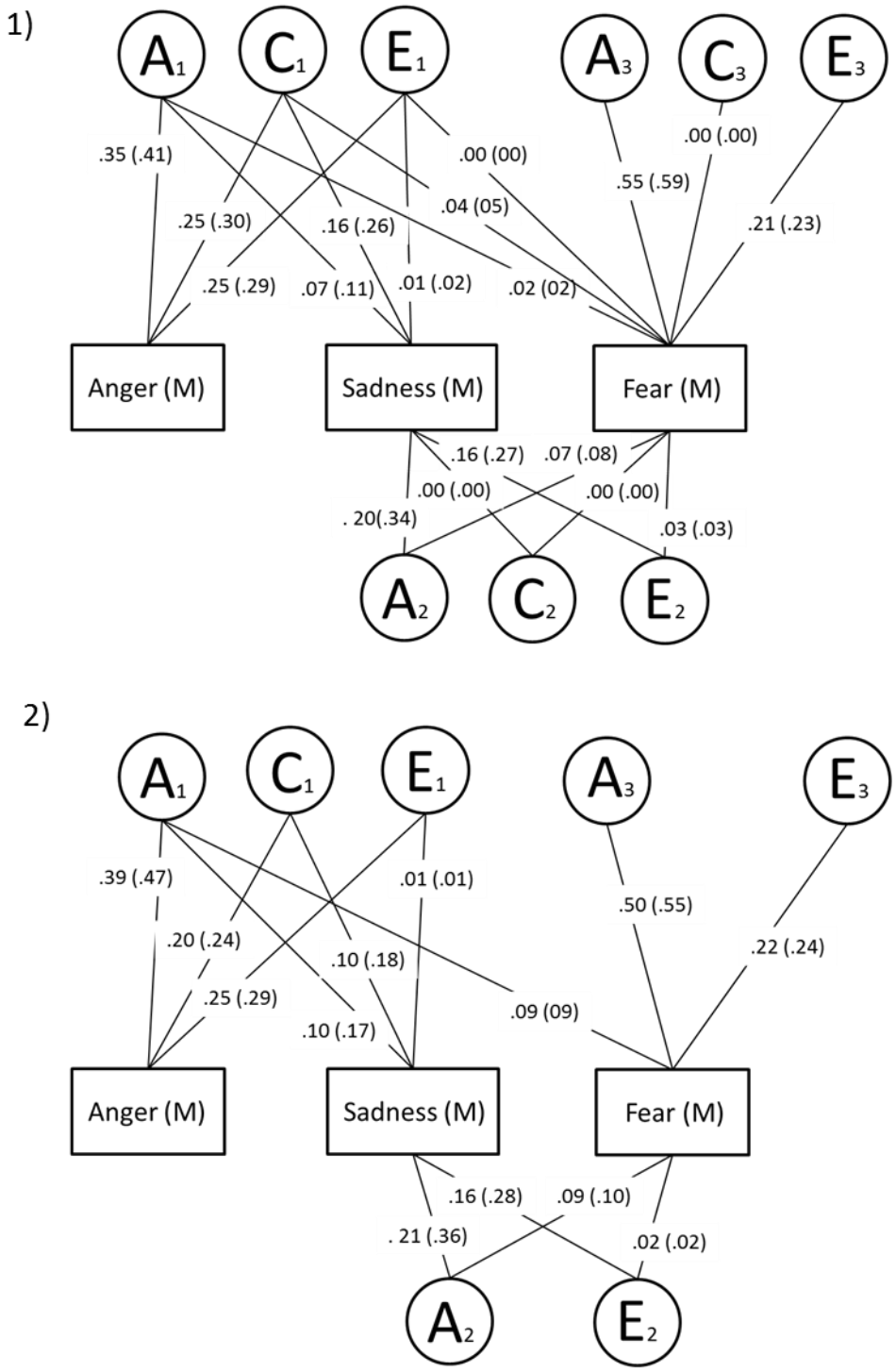

Figure 6. Full and final trivariate Cholesky decompositions for mother-report of anger, sadness and fear. Standardized A, C and E components are in parentheses.

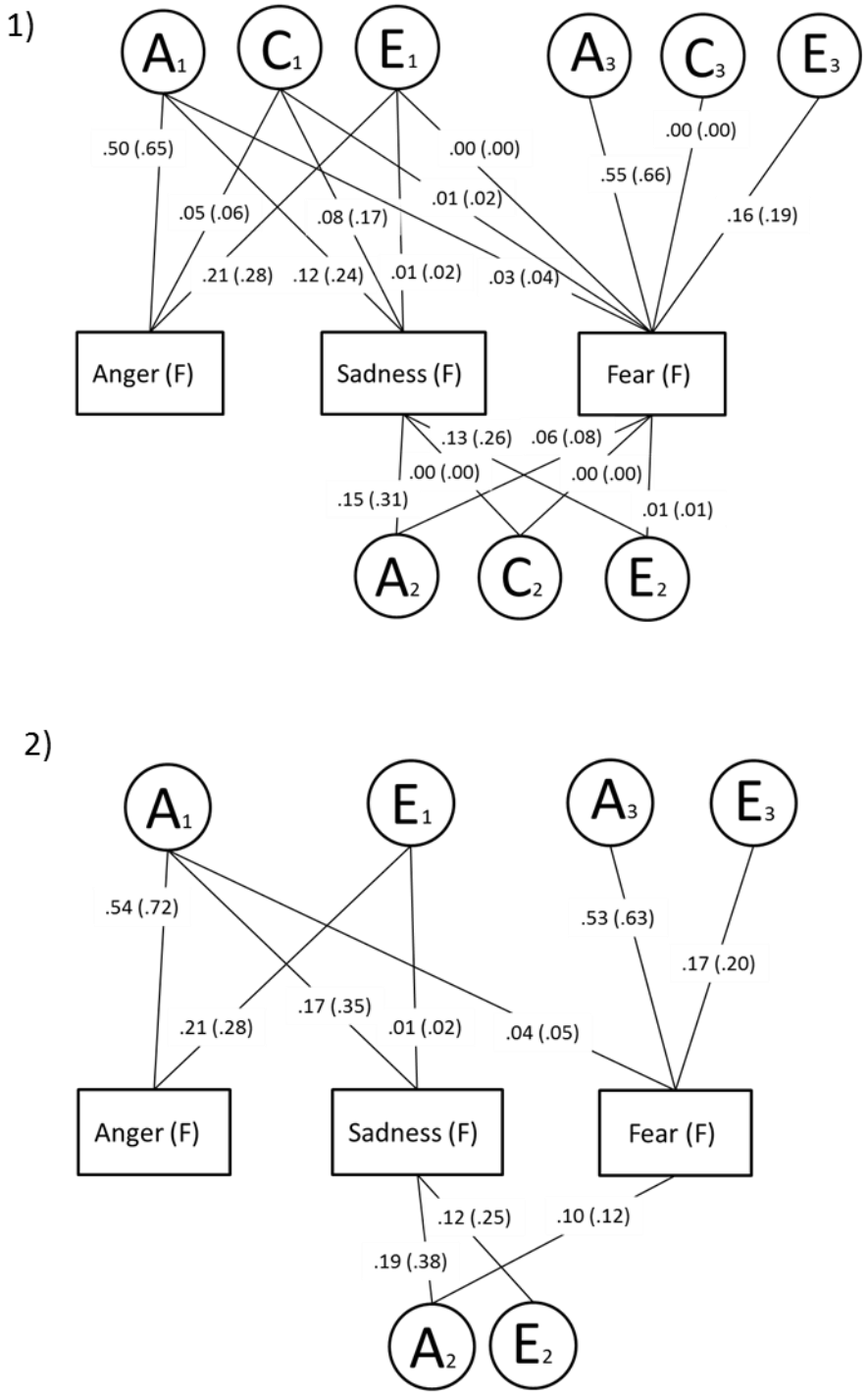

Figure 7. Full and final trivariate Cholesky decompositions for father-report of anger, sadness and fear. Standardized A, C and E components are in parentheses.

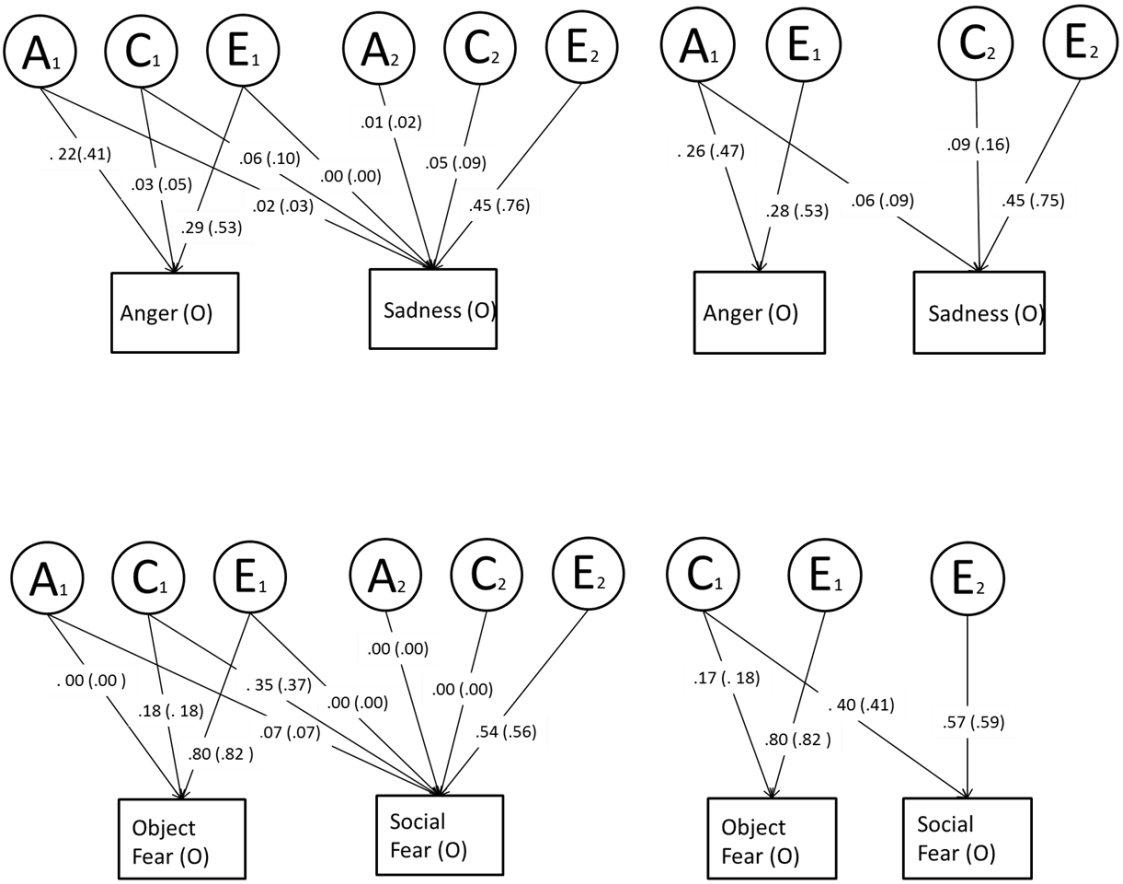

Figure 8. Full and final bivariate Cholesky decompositions for in-home observation of anger and sadness, and for in-home observation of object and social fear. The full model is depicted on the left, and the final, reduced model on the right. Standardized A, C and E components are in parentheses.

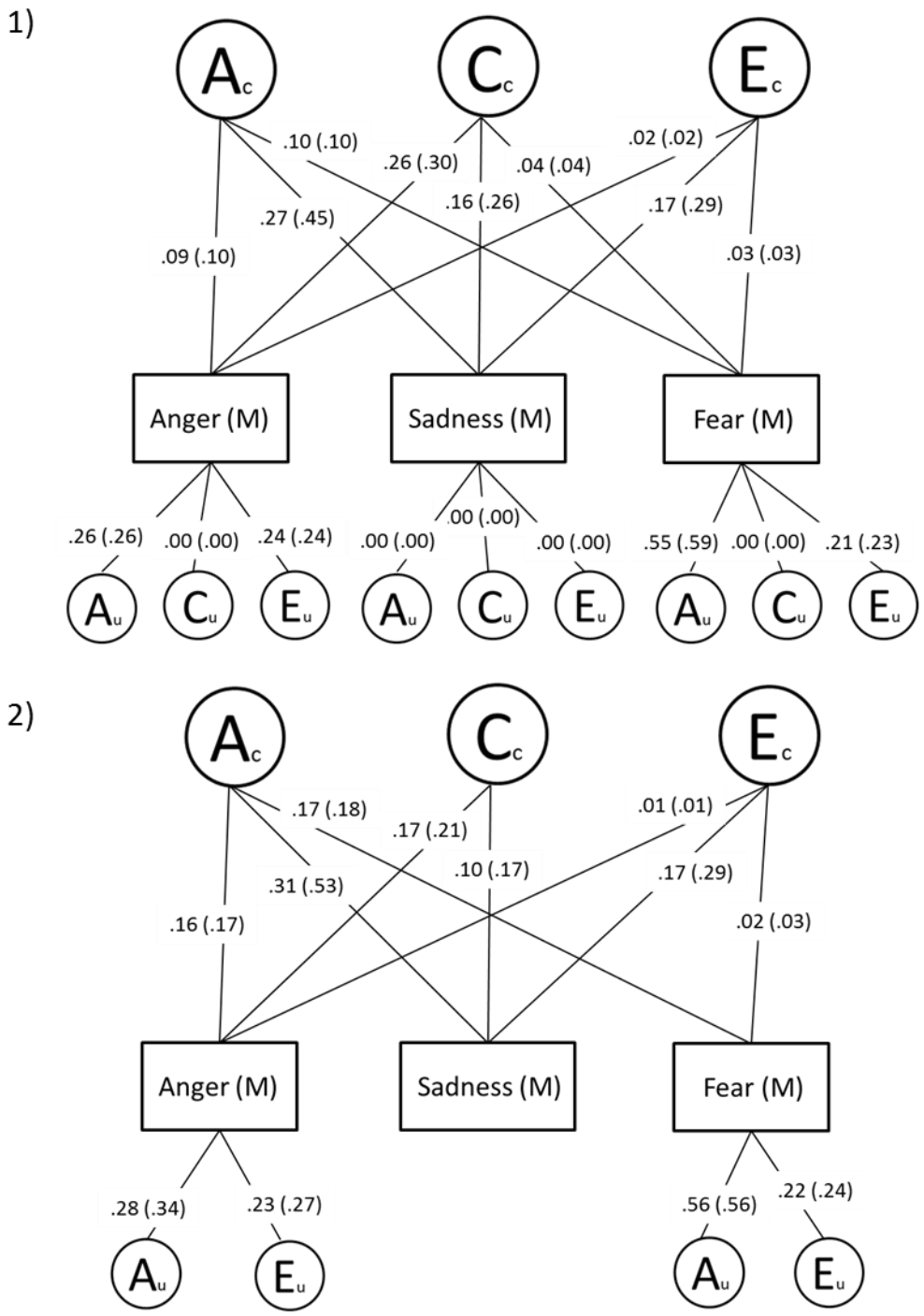

Figure 9. Full and final independent pathway models for mother-report of anger, sadness and fear. Standardized A, C and E components are in parentheses.

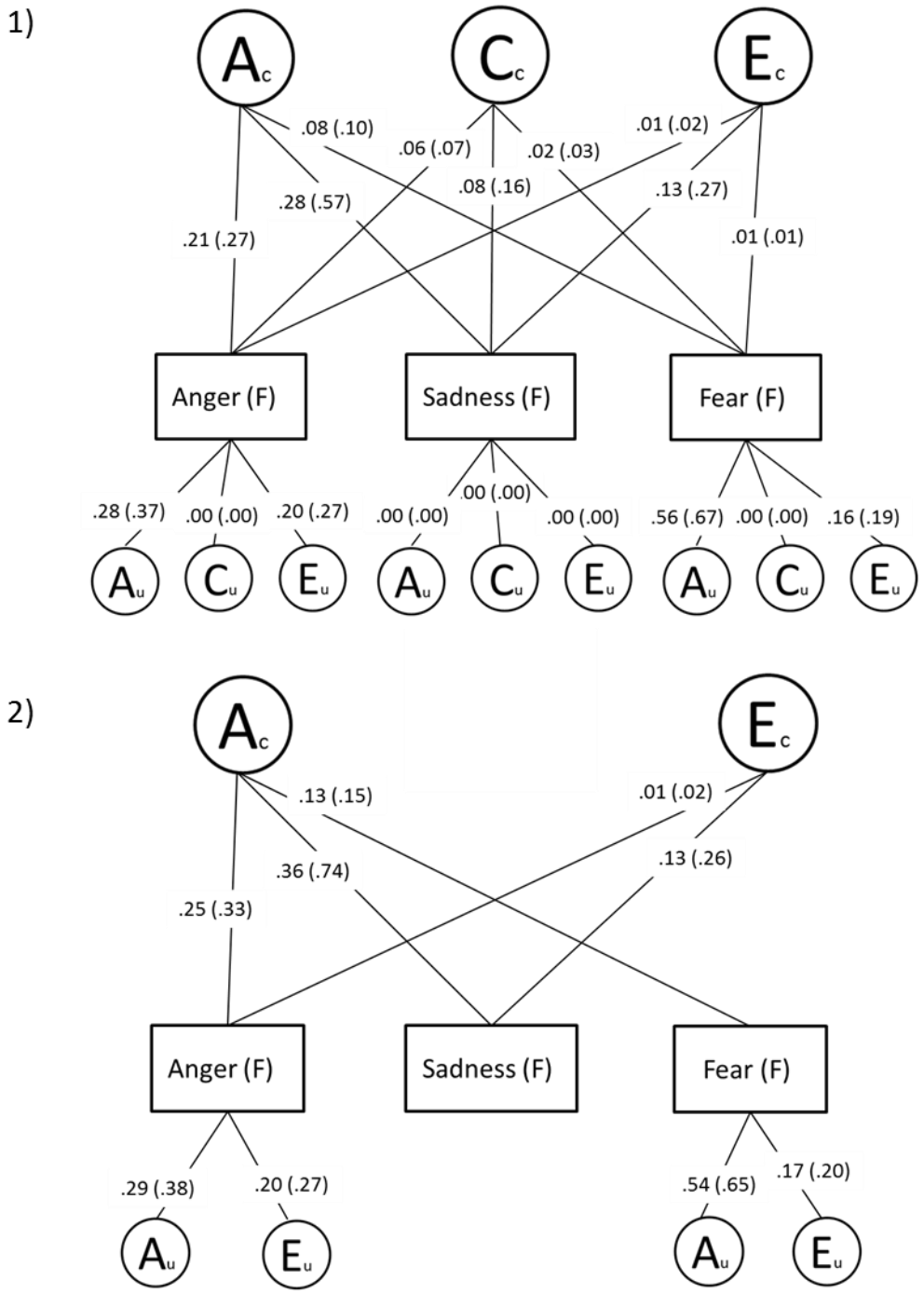

Figure 10. Full and final independent pathway models for father-report of anger, sadness and fear. Standardized A, C and E components are in parentheses.

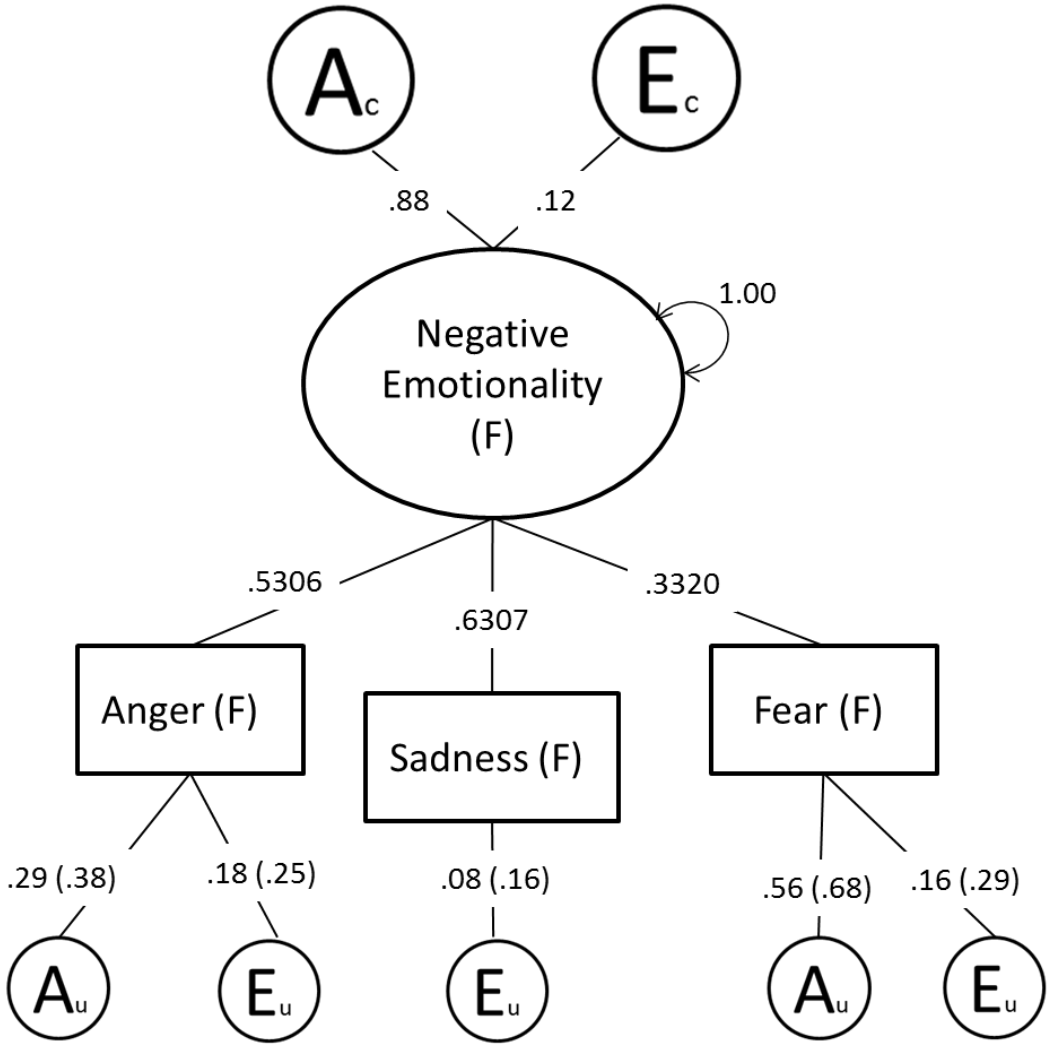

Figure 11. Final common pathway models for father-report of anger, sadness and fear. Standardized A, C and E components are in parentheses.

APPENDIX A

CHILDREN'S BEHAVIOR QUESTIONNAIRE

SELECTED SCALES

Children's Behavior Questionnaire

Version l

Subject No. _____

Today's Date _____

Sex of Child _____

Date of Child's Birth:

_____ _____ _____
Month Day Year

Age of Child _____ _____
Years months

Instructions: Please read carefully before starting:

On the next pages you will see a set of statements that describe children's reactions to a number of situations. We would like you to tell us what your child's reaction is likely to be in those situations. There are of course no "correct" ways of reacting; children differ widely in their reactions, and it is these differences we are trying to learn about. Please read each statement and decide whether it is a "true" or "untrue" description of your child's reaction within the past six months. Use the following scale to indicate how well a statement describes your child:

Circle #	If the statement is:
l	extremely untrue of your child
2	quite untrue of your child
3	slightly untrue of your child
4	neither true nor false of your child
5	slightly true of your child
6	quite true of your child
7	extremely true of your child

If you cannot answer one of the items because you have never seen the child in that situation, for example, if the statement is about the child's reaction to your singing and you have never sung to your child, then circle NA (not applicable).

Please be sure to circle a number or NA for every item.

Anger (AN) Scale (10 items) Amount of negative affect related to interruption of ongoing tasks or goal blocking.

7. Rarely gets irritated when s/he makes a mistake. {REVERSED SCORE]

1 2 3 4 5 6 7 NA

12. Has temper tantrums when s/he doesn't get what s/he wants.

1 2 3 4 5 6 7 NA

24. Gets quite frustrated when prevented from doing something s/he wants to do.

1 2 3 4 5 6 7 NA

30. Gets angry when s/he can't find something s/he wants to play with.

1 2 3 4 5 6 7 NA

46. Rarely gets upset when told s/he has to go to bed. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

50. Becomes easily frustrated when tired.

1 2 3 4 5 6 7 NA

61. Rarely protests when another child takes his/her toy away. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

68. Easily gets irritated when s/he has trouble with some task (e.g., building, drawing, dressing).

1 2 3 4 5 6 7 NA

71. Gets angry when called in from play before s/he is ready to quit.

1 2 3 4 5 6 7 NA

77. Gets mad when provoked by other children.

1 2 3 4 5 6 7 NA

Sadness (SD) Scale (10 items) Amount of negative affect and lowered mood and energy related to exposure to suffering, disappointment and object loss

6. Cries sadly when a favorite toy gets lost or broken.

1 2 3 4 5 6 7 NA

16. Tends to feel "down" at the end of an exciting day.

1 2 3 4 5 6 7 NA

25. Becomes upset when loved relatives or friends are getting ready to leave following a visit.

1 2 3 4 5 6 7 NA

27. Does not usually become tearful when tired. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

32. Her/his feelings are easily hurt by what parents say.

1 2 3 4 5 6 7 NA

36. Becomes tearful when told to do something s/he does not want to do.

1 2 3 4 5 6 7 NA

41. Rarely cries when s/he hears a sad story. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

42. Rarely becomes upset when watching a sad event in a TV show. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

49. Sometimes appears downcast for no reason.

1 2 3 4 5 6 7 NA

60. Rarely becomes discouraged when s/he has trouble making something work. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

Fear (FE) Scale (10 items) Amount of negative affect, including unease, worry nervousness related to anticipated pain or distress and/or potentially threatening situations.

5. Is not afraid of large dogs and/or other animals. [REVERSED SCORE]

1	2	3	4	5	6	7	NA

21. Is afraid of loud noises.

1	2	3	4	5	6	7	NA

23. Doesn't worry about injections by the doctor. [REVERSED SCORE]

1	2	3	4	5	6	7	NA

26. Is not afraid of the dark. [REVERSED SCORE]

1	2	3	4	5	6	7	NA

31. Is afraid of fire.

1	2	3	4	5	6	7	NA

34. Is very frightened by nightmares.

1	2	3	4	5	6	7	NA

52. Is afraid of the dark.

1	2	3	4	5	6	7	NA

55. Is rarely frightened by "monsters" seen on TV or at movies. [REVERSED SCORE]

1	2	3	4	5	6	7	NA

63. Is not afraid of heights. [REVERSED SCORE]

1	2	3	4	5	6	7	NA

70. Is rarely afraid of sleeping alone in a room. [REVERSED SCORE]

1	2	3	4	5	6	7	NA

Shyness (SH) Scale (10 items) Slow or inhibited approach in situations involving novelty or uncertainty.

3. Sometimes prefers to watch rather than join other children playing.

1 2 3 4 5 6 7 NA

8. Seems to be at ease with almost any person. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

14. Gets embarrassed when strangers pay a lot of attention to her/him.

1 2 3 4 5 6 7 NA

18. Acts very friendly and outgoing with new children. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

22. Joins others quickly and comfortably, even when they are strangers. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

28. Is sometimes shy even around people s/he has known a long time.

1 2 3 4 5 6 7 NA

33. Sometimes seems nervous when talking to adults s/he has just met.

1 2 3 4 5 6 7 NA

39. Acts shy around new people.

1 2 3 4 5 6 7 NA

45. Is comfortable asking other children to play. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

51. Talks easily to new people. [REVERSED SCORE]

1 2 3 4 5 6 7 NA

APPENDIX B

LABORATORY TEMPERAMENT ASSESSMENT BATTERY

SELECTED EPISODES

I'm Not Sharing

A. RATIONALE

This task targets the child's feelings of being treated unjustly.

B. PHYSICAL SETTING

The child is sitting at a table and chair, facing the camera, and the child tester is sitting to child's right. The parent should not be present during this episode.

C. STIMULI

Eighteen pieces of individually wrapped candies in a small plastic bag and two plastic containers with narrow tops (to prevent child from holding/unwrapping candy during episode).

D. PROCEDURE

The episode begins when child tester gives child the first piece of candy and the candy hits the bottom of the container and ends when the child tester finishes saying, "let's each pick just two pieces of candy."

E. CAMERA INSTRUCTIONS

This shot requires focusing on the child's face and upper body, as well as the candy and containers.

F. SCORING

The episode is divided into 10-second epochs and can vary in length depending on the individual child and child tester.

Variables to be scored:

a. Baseline state (coded by child tester)
b. Intensity of positivity
c. Presence of negativity

Anger variables

d. Intensity of anger expression
e. Intensity of bodily anger/ frustration
f. Latency to first anger response
g. Presence of resistance

Sadness variables

h. Intensity of sadness expression
i. Intensity of bodily sadness
j. Presence of resignation
h. Latency to first sadness response

Language coding

i. Anger vocalizations
j. Sadness vocalizations
k. Positive vocalizations

Definitions of variables:

a. Baseline state: The child's state prior to the beginning of the episode is coded by the child tester using the following scale:
   1 = alert and calm
   2 = alert and active
   3 = tired or inattentive
   4 = whiny, complaining or distressed
   5 = resistant

b. Intensity of positivity at any time during the epoch is noted, so long as it is obvious. Examples may include positive verbalizations or positive facial or bodily expressions.
   0 = no positivity
   1 = Tense, small, restrained smiles or small giggles
   2= "True," joyful smiles and unrestrained laughter

c. Presence of negativity at any time during the epoch is noted, so long as it is obvious. Examples may include negative verbalizations or negative facial or bodily expressions.
   0 = no
   1 = yes

Anger variables

d. Intensity of anger expression: Peak intensity of facial anger is coded in each epoch using AFFEX (See Appendix for definitions) and is rated on the following scale:

0 = No facial region shows codeable anger movement.
1 = Only one facial region shows codeable anger movement, identifying low intensity anger, or expression is ambiguous.
2 = Two facial regions show codeable anger movement, or expression in one region (*e.g.*, brows) is definite.
3 = All three facial regions show codeable anger movement, or coder otherwise has impression of strong anger.

e. Intensity of bodily anger/ frustration: Peak intensity of bodily anger/ frustration is coded in each epoch and is rated on the following scale:
0 = No detectable bodily anger/ frustration.
1 = Low bodily anger/ frustration: Slight bodily tensing or mild frustration behavior (*e.g.*, gripping the container tightly, sighing mildly).
2 = Moderate bodily anger/ frustration: Moderate bodily tensing (*e.g.*, balling the fists, swinging the legs) or moderate frustration behavior (*e.g.*, pushing away from the table, sighing heavily, pushing the container away).
3 = High bodily anger/ frustration: High bodily tensing (*e.g.*, trembling) or high frustration behavior (*e.g.*, throwing the container).

f. Latency to first anger response: Interval, in seconds, from the beginning of the episode to the first codeable anger response (*e.g.*, facial, vocal, bodily).
* If the child is already displaying anger before the episode starts, the code = 1.
* If the child never displays anger, the code = 241.

g. Presence of Resistance: Whether or not the child attempts to get the candy is coded within each epoch.
0=The child does not resist or attempt to get the candy in any way.
1=The child actively attempts to obtain the candy.

Sadness variables

h. Intensity of sadness expression: Peak intensity of facial sadness is coded in each epoch using AFFEX (see Appendix for definitions) and is rated on the following scale:
0 = No facial region shows codeable sadness movement.
1 = Only one facial region shows codeable sadness movement, identifying low intensity sadness, or expression is ambiguous.
2 = Two facial regions show codeable sadness movement, or expression in one region is definite (*e.g.*, brows).
3 = All three facial regions show codeable sadness movement, or coder otherwise has impression of strong facial sadness.

i. Intensity of bodily sadness: Peak intensity of bodily sadness is coded in each epoch and is rated on the following scale:
0 = No detectable bodily sadness.
1 = Mild bodily sadness (*e.g.*, downcast head).
2 = Moderate bodily sadness (*e.g.*, slumped shoulders).
3 = High bodily sadness (*e.g.*, head in arms or hands).

j. Presence of resignation: Presence of resignation is noted in each epoch.
0 = No detectable resignation
1 = Attempts to obtain the candy have been reduced; child is accepting the fact that s/he will receive less candy.

h. Latency to first sadness response: Interval, in seconds, from the beginning of the episode to the first codeable sadness response (*e.g.*, facial, vocal, bodily).
* If the child is already displaying sadness before the episode starts, the code = 1.
* If the child never displays sadness, the code = 241.

Language coding

i. Anger vocalizations: Peak intensity of anger vocalizations is recorded for each of the four time blocks.
0 = No evidence of anger vocalizations or ambiguous.
1 = Mild/moderate anger vocalizations.
2 = Moderate/high anger vocalizations.
3 = Extreme anger vocalizations.

m. Sadness vocalizations: Peak intensity of sadness vocalizations is recorded for each of the four time blocks.
0 = No evidence of sadness vocalizations or ambiguous.
1 = Mild/moderate sadness vocalizations.
2 = Moderate/high sadness vocalizations.
3 = Extreme sadness vocalizations.

n. Positive vocalizations: Peak intensity of positive vocalizations is recorded for each of the four time blocks.
0 = No evidence of positive vocalizations or ambiguous.
1 = Mild/moderate positive vocalizations.

2 = Moderate/high positive vocalizations.

I'm Not Sharing

I.D. # _____ Child's name and last initial: _____ Date of birth: _____
Scorer: _____ Date scored: _____ Date of visit: _____
Tape # _____ Counter # _____ Episode order: _____

Child tester: _____

Baseline (coded by child tester) _____

Latency to first anger response _____ secs.

Latency to first sadness response _____ secs.

All epochs may not be used (Code unused epochs as "8s")

Minute 1

10-second epochs	1	2	3	4	5	6
Time (Begin)						
Intensity of anger expression (0 - 3)						
Intensity of bodily anger/ frustration (0 - 3)						
Intensity of sadness expression (0 - 3)						
Intensity of bodily sadness (0 -3)						
Presence of resignation (0 = no; 1 = yes)						
Presence of positivity (0 = no; 1 = yes)						
Presence of negativity (0 = no; 1 = yes)						

I'm Not Sharing Language Coding

Latency to first anger vocalization (in secs.) _____
Latency to first sadness vocalization (in secs.) _____

Time Block 1

Counters (Begin-end) _____

Anger vocalizations: Please rank the peak score only (0-3). _____

Sadness vocalizations: Please rank the peak score only (0-3). _____

Positive vocalizations: Please rank the peak score only (0-2). _____

Impossibly Perfect Stars

RATIONALE

Katherine Korner in "Some Aspects of Hostility in Young Children" cites an experiment in which the children who participated were asked to draw the human figure repeatedly. The outcome of this experiment was that after completion of many figures the children became bored and angry.

B. PHYSICAL SETTING

The child is sitting at a table and chair, facing the camera, and the child tester is sitting to child's right. The parent should not be present during this episode.

C. STIMULI

Washable marker (Crayola® brand) and a sheet of white paper

D. PROCEDURE

The episode begins when the child begins drawing his/her first star and ends four minutes later or when child tester says to child, "That one looks pretty good."

E. CAMERA INSTRUCTIONS

This shot requires focusing on the child's face and upper body, as well as the paper and marker.

F. SCORING

This episode lasts for four minutes and is divided into 10-second epochs.

Variables to be scored:

a. Baseline (coded by child tester)
b. Presence of positivity
c. Presence of negativity
d. Number of critiques given by child tester

Anger variables

e. Intensity of anger expression
f. Intensity of bodily anger/ frustration
g. Latency to first anger response

Sadness variables

j. Intensity of sadness expression
k. Intensity of bodily sadness
l. Latency to first sadness response

Language coding

k. Anger vocalizations
l. Sadness vocalizations
m. Positive vocalizations

Definitions of variables:

a. Baseline state: The child's state prior to the beginning of the episode is coded by the child tester using the following scale:
   1 = alert and calm
   2 = alert and active
   3 = tired or inattentive
   4 = whiny, complaining or distressed
   5 = resistant

b. Presence of positivity at any time during the epoch is noted, so long as it is obvious. Examples may include positive facial or bodily expressions.
   0 = no
   1 = yes

c. Presence of negativity at any time during the epoch is noted, so long as it is obvious. Examples may include negative facial or bodily expressions.
   0 = no
   1 = yes

d. Number of critiques given by child tester.

Anger variables

e. Intensity of anger expression: Peak intensity of facial anger is coded in each epoch using AFFEX (See Appendix for definitions) and is rated on the following scale:
   0 = No facial region shows codeable anger movement.
   1 = Only one facial region shows codeable anger movement, identifying low intensity anger.
   2 = Two facial regions show codeable anger movement, or expression in one region (*e.g.*, brows) is definite.
   3 = All three facial regions show codeable anger movement, or coder otherwise has impression of strong anger.

f. Intensity of bodily anger/ frustration: Peak intensity of bodily anger/ frustration is coded in each epoch and is rated on the following scale:
   0 = No detectable bodily anger/ frustration.
   1 = Low bodily anger/ frustration: Slight bodily tensing or mild frustration behavior (*e.g.*, pressing harder on the paper, sighing with frustration).
   2 = Moderate bodily anger/ frustration: Moderate bodily tensing (*e.g.*, balling the fists, swinging the legs) or moderate frustration behavior (*e.g.*, pushing away from the table, sighing heavily, pushing the paper away).
   3 = High bodily anger/ frustration: High bodily tensing (*e.g.*, trembling) or high frustration behavior (*e.g.*, throwing the paper or marker).

g. Latency to first anger response: Interval, in seconds, from the beginning of the episode to the first codeable anger response (*e.g.*, facial, bodily).
   * If the child is already displaying anger before the episode starts, the code = 1.
   * If the child never displays anger, the code = 241.

Sadness variables

l. Intensity of sadness expression: Peak intensity of facial sadness is coded in each epoch using AFFEX (see Appendix for definitions) and is rated on the following scale:
   0 = No facial region shows codeable sadness movement.
   1 = Only one facial region shows codeable sadness movement, identifying low intensity sadness.
   2 = Two facial regions show codeable sadness movement, or expression in one region is definite (*e.g.*, brows).
   3 = All three facial regions show codeable sadness movement, or coder otherwise has impression of strong facial sadness.

m. Intensity of bodily sadness: Peak intensity of bodily sadness is coded in each epoch and is rated on the following scale:
   0 = No detectable bodily sadness.
   1 = Mild bodily sadness (*e.g.*, downcast head).
   2 = Moderate bodily sadness (*e.g.*, slumped shoulders).
   3 = High bodily sadness (*e.g.*, buries head in arms).

n. Latency to first sadness response: Interval, in seconds, from the beginning of the episode to the first codeable sadness response (*e.g.*, facial, bodily).
   * If the child is already displaying sadness before the episode starts, the code = 1.
   * If the child never displays sadness, the code = 241.

Language coding

k. Anger vocalizations: Peak intensity of anger vocalizations is recorded for each of the four time blocks.
   0 = No evidence of anger vocalizations or ambiguous.
   1 = Mild/moderate anger vocalizations.
   2 = Moderate/high anger vocalizations.
   3 = Extreme anger vocalizations.

l. Sadness vocalizations: Peak intensity of sadness vocalizations is recorded for each of the four time blocks.
   0 = No evidence of sadness vocalizations or ambiguous.
   1 = Mild/moderate sadness vocalizations.
   2 = Moderate/high sadness vocalizations.
   3 = Extreme sadness vocalizations.

m. Positive vocalizations: Peak intensity of positive vocalizations is recorded for each of the four time blocks.
   0 = No evidence of positive vocalizations or ambiguous.
   1 = Mild/moderate positive vocalizations.
   2 = Moderate/high positive vocalizations.

Impossibly Perfect Stars

I.D. # _____ Child's name and last initial: _____ Date of birth: _____
Scorer: _____ Date scored: _____ Date of visit: _____
Tape # _____ Counter # _____ Episode order: _____

Child tester: _____

Baseline (coded by child tester) _____
Latency to first anger response _____ secs.
Latency to first sadness response _____ secs.

Minute 1

10-second epochs	1	2	3	4	5	6
Time (Begin)						
Intensity of anger expression (0-3)						
Intensity of bodily anger/ frustration (0-3)						
Intensity of sadness expression (0-3)						
Intensity of bodily sadness (0-3)						
Presence of positivity (0 = no; 1 = yes)						
Presence of negativity (0 = no; 1 = yes)						
How many critiques given by CT?						

Impossibly Perfect Stars Language Coding

Latency to first anger vocalization (in secs.) _____
Latency to first sadness vocalization (in secs.) _____

Time Block 1 (Time from when the episode starts until the beginning of the first prompt by child tester)

Counters (Begin-end) _____

Anger vocalizations: Please rank the peak score only (0-3). _____

Sadness vocalizations: Please rank the peak score only (0-3). _____

Positive vocalizations: Please rank the peak score only (0-2). _____

Transparent Box

A. RATIONALE

This task is designed to evoke frustration or anger by preventing the child from playing with the selected toy. The child will be able to see the object of desire (in this case a toy) through a clear plastic box but be unable to attain it because the box will be locked and the keys will not open it.

B. PHYSICAL SETTING

The child should be sitting at a table facing the camera in a room empty of other toys/props. The parent should not be present during this episode.

C. STIMULI

Large transparent box, remote control robot, Gameboy®, Barbie® on a horse, padlock, ring of faulty keys, and correct key.

D. PROCEDURE

The episode begins after the child tester says, "I can't remember which key opens the lock. You can just try them, and when you open the box, you can play with the toy you want," and it ends four minutes later when the child tester returns to the room and says, "Didn't the lock open?"

E. CAMERA INSTRUCTIONS

It is important to capture a tight shot of the child's face and upper body, also include the locked box in the shot.

F. SCORING

This episode lasts for four minutes and is divided into 10-second epochs.

Variables to be scored:

a. Baseline state (coded by child tester)
b. Latency to first stop or to be off-task
c. Presence of positivity
d. Presence of negativity
e. Language valence
f. Language content

Anger variables

g. Intensity of anger expression
h. Intensity of bodily anger/ frustration
i. Latency to first anger response

Sadness variables

j. Intensity of sadness expression
k. Intensity of bodily sadness
l. Presence of resignation
m. Latency to first sadness response

The following variable is not coded within epochs

n. Duration of on-task work

Definitions of variables:

a. Baseline state: The child's state prior to the beginning of the episode is coded by the child tester using the following scale:
   1 = alert and calm
   2 = alert and active
   3 = tired or inattentive
   4 = whiny, complaining or distressed
   5 = resistant

b. Latency to first stop or to be off-task: The amount of time, in seconds, it takes the child to first either be off-task or give up on the getting the toys completely from the time s/he **first begins** working on getting the toys. The child's attention must shift to something other than getting the toys.
   * Off-task is defined as not purposefully getting the toys and disengaged from the task. The child must display this behavior for at least 5 seconds in order for him/her to be coded as off-task.
   * If getting the toys is given up on, the child must decide to quit work on getting the toys in order for it to be coded as a first quit.
   * If child never stops working on task, code = 241.

c. Presence of positivity at any time during the epoch is noted, so long as it is obvious. Examples may include positive verbalizations or positive facial or bodily expressions.
   0 = no
   1 = yes

d. Presence of negativity at any time during the epoch is noted, so long as it is obvious. Examples may include negative verbalizations or negative facial or bodily expressions.
   0 = no
   1 = yes

e. Language valence: The **tone** of the child's voice is coded for verbalizations. The peak intensity is coded in each epoch using the following scale:
   0 = No verbalizations or positive valence (*e.g.*, laughter).
   1 = Neutral valence.
   2 = Low negative valence.
   3 = High negative valence.

f. Language content: The **content** of the child's verbalizations is coded. The peak intensity is coded in each epoch using the following scale:
   0 = No verbalizations or positive content.
   1 = Neutral content.
   2 = Low negative content: Indirect refusal to continue (*e.g.*, "I can't open it.").
   3 = High negative content: Direct refusal to continue (*e.g.*, "I quit!").

Anger variables

g. Intensity of anger expression: Peak intensity of facial anger is coded in each epoch using AFFEX (See Appendix for definitions) and is rated on the following scale:
   0 = No facial region shows codeable anger movement.
   1 = Only one facial region shows codeable anger movement, identifying low intensity anger, or expression is ambiguous.
   2 = Two facial regions show codeable anger movement, or expression in one region (*e.g.*, brows) is definite.
   3 = All three facial regions show codeable anger movement, or coder otherwise has impression of strong anger.

h. Intensity of bodily anger/ frustration: Peak intensity of bodily anger/ frustration is coded in each epoch and is rated on the following scale:
   0 = No detectable bodily anger/ frustration.
   1 = Low bodily anger/ frustration: Slight bodily tensing or mild frustration behavior (*e.g.*, gripping the keys tightly, knitting the brows together, sighing mildly).
   2 = Moderate bodily anger/ frustration: Moderate bodily tensing (*e.g.*, balling the fists, swinging the legs) or moderate frustration behavior (*e.g.*, pushing away from the table, sighing heavily, pushing the box away).
   3 = High bodily anger/ frustration: High bodily tensing (*e.g.*, trembling) or high frustration behavior (*e.g.*, throwing the keys or the box).

i. Latency to first anger response: Interval, in seconds, from the beginning of the episode to the first codeable anger response (*e.g.*, facial, vocal, bodily).
   * If the child is already displaying anger before the episode starts, the code = 1.
   * If the child never displays anger, the code = 241.

Sadness variables

j. Intensity of sadness expression: Peak intensity of facial sadness is coded in each epoch using AFFEX (see Appendix for definitions) and is rated on the following scale:
   0 = No facial region shows codeable sadness movement.
   1 = Only one facial region shows codeable sadness movement, identifying low intensity sadness, or expression is ambiguous.
   2 = Two facial regions show codeable sadness movement, or expression in one region is definite (*e.g.*, brows).
   3 = All three facial regions show codeable sadness movement, or coder otherwise has impression of strong facial sadness.

k. Intensity of bodily sadness: Peak intensity of bodily sadness is coded in each epoch and is rated on the following scale:
   0 = No detectable bodily sadness.
   1 = Mild bodily sadness (*e.g.*, downcast head).
   2 = Moderate bodily sadness (*e.g.*, slumped shoulders).
   3 = High bodily sadness (*e.g.*, head in arms or hands).

l. Presence of resignation: Presence of resignation is coded in each epoch.
   0 = No detectable resignation.
   1 = Attempts to obtain the toys have been **reduced**; child is accepting of the fact that s/he will not get the toys.

m. Latency to first sadness response: Interval, in seconds, from the beginning of the episode to the first codeable sadness response (*e.g.*, facial, vocal, bodily).
   * If the child is already displaying sadness before the episode starts, the code = 1.
   * If the child never displays sadness, the code = 241.

The following variable is not coded within epochs

n. Duration of on-task work: The duration of on-task work is scored for each instance of on-task work on the last page. The total number of seconds that the on-task work lasted is recorded.

Transparent Box

I.D. # _____ Child's name and last initial: _____ Date of birth: _____
Scorer: _____ Date scored: _____ Date of visit: _____
Tape # _____ Counter # _____ Episode order: _____

Child tester: _____

Baseline (coded by child tester) _____

Latency to first stop _____ secs.

10-second epochs	1	2	3	4	5	6
Time (Begin)						
Intensity of facial anger (0 - 3)						
Intensity of bodily anger/ frustration (0-3)						
Intensity of facial sadness (0 - 3)						
Intensity of bodily sadness (0 - 3)						
Presence of resignation (0 = no; 1 = yes)						
Language valence (0 - 3)						
Language content (0 - 3)						
Presence of positivity (0 = no; 1 = yes)						
Presence of negativity (0 = no; 1 = yes)						

Latency to first anger response _____ secs.

Latency to first sadness response _____ secs.

Wrong Gift

A. RATIONALE

This episode will be used to assess vulnerability to negative affect and regulation of negative affect when frustrated.

B. PHYSICAL SETTING

The episode takes place in a room empty of other toys/props. The child sits at a table and chair, facing the camera, and the friendly stranger sits at the table to child's right. After the friendly stranger leaves, the child tester also sits to the child's right. The parent should not be present during this episode.

C. STIMULI

A box with five compartments (each section has a number on the bottom for ranking prizes and prize placement) and five prizes for ranking (slinky, yo-yo, top, rubber dinosaur, and a worn, broken white crayon).

D. PROCEDURE

The episode begins when the child tester says, "You know **(CHILD'S NAME)**, you've done such a great job with all these games so far today that I'm going to go tell **(FRIENDLY STRANGER'S NAME)** to get your prize!" and ends 30 seconds after the child tester and friendly stranger leave the room to let the child pick his/her prize.

.

E. CAMERA INSTRUCTIONS

The camera shot should be a fairly tight face shot, but include the child's upper body. The shot should clearly include the child with the gift.

F. SCORING

This episode is divided into five intervals.

The first interval consists of six 5-second epochs. It begins when the child finds out that s/he will get a prize and ends when the child sees the wrong gift for the first time.

The second interval consists of twelve 5-second epochs. It begins when the child sees the wrong gift for the first time and ends when the friendly stranger gets up to leave.

The third interval consists of six 5-second epochs. It begins when the friendly stranger gets up to leave and ends when the child tester starts asking questions.

The fourth interval can vary in length depending on the questions asked and the individual differences in the children and is divided into 5-second epochs. It begins when the child tester starts asking questions and ends when the friendly stranger and child tester leave.

The fifth interval is made up of six 5-second epochs. It begins when the friendly stranger and child tester leave and ends 30 seconds later.

Variables to be scored:

a. Baseline (coded by child tester)

Interval 1

b. Latency to first positivity
c. Intensity of anticipatory positive affect
d. Presence of positivity
e. Presence of negativity

Intervals 2, 3, and 4

f. Intensity of anger expression
g. Intensity of bodily anger/ frustration
h. Latency to first anger response
i. Intensity of sadness expression
j. Intensity of bodily sadness
k. Latency to first sadness response
l. Presence of positivity
m. Presence of negativity
n. Language valence
o. Language content

Interval 5

p. Presence of positivity
q. Presence of negativity

Definitions of variables:

a. Baseline state: The child's state prior to the beginning of the episode is coded by the child tester using the following scale:
   1 = alert and calm
   2 = alert and active
   3 = tired or inattentive
   4 = whiny, complaining or distressed
   5 = resistant

Interval 1 (Waiting for gift)

b. Latency to first anticipatory positive affect: The amount of time, in seconds, it takes the child to first show positive affect (*e.g.*, smiling, wiggling…) from the time s/he finds out that s/he will get a prize.
* If child is already showing positive affect when the interval begins, the code = 1.
* If child never shows positive affect during the interval, the code = 31.

c. Intensity of anticipatory positive affect: Peak intensity of anticipatory positive affect is coded in each epoch and is rated on the following scale:
0 = No anticipatory positive affect: Child doesn't seem to care about receiving the gift.
1 = Slight/ambiguous positive anticipation: Child may show a small to moderate smile.
2 = Moderate positive anticipation: Child may eagerly wait for gift, may giggle, or calmly talk about receiving the gift; there will almost definitely be a smile.
3 = High positive anticipation: Child may excitedly wait for gift, possible continuous laughter, bouncing in chair, and/or excited verbalization about the gift.

d. Presence of positivity at any time during the epoch is noted, so long as it is obvious. Examples may include positive verbalizations or positive facial or bodily expressions.
0 = no
1 = yes

e. Presence of negativity at any time during the epoch is noted, so long as it is obvious. Examples may include negative verbalizations or negative facial or bodily expressions.
0 = no
1 = yes

Intervals 2, 3, and 4

f. Intensity of anger expression: Peak intensity of facial anger is coded in each epoch using AFFEX (See Appendix for definitions) and is rated on the following scale:
0 = No facial region shows codeable anger movement.
1 = Only one facial region shows codeable anger movement, identifying low intensity anger, or expression is ambiguous.
2 = Two facial regions show codeable anger movement, or expression in one region (*e.g.*, brows) is definite.
3 = All three facial regions show codeable anger movement, or coder otherwise has impression of strong anger.

g. Intensity of bodily anger/ frustration: Peak intensity of bodily anger/ frustration is coded in each epoch and is rated on the following scale:
0 = No detectable bodily anger/ frustration.
1 = Low bodily anger/ frustration: Slight bodily tensing or mild frustration behavior (*e.g.*, dropping the prize, pushing it lightly, knitting the brows together, sighing mildly).
2 = Moderate bodily anger/ frustration: Moderate bodily tensing (*e.g.*, balling the fists, swinging the legs) or moderate frustration behavior (*e.g.*, pushing away from the table, sighing heavily, pushing the prize away).
3 = High bodily anger/ frustration: High bodily tensing (*e.g.*, trembling) or high frustration behavior (*e.g.*, throwing the prize).

h. Latency to first anger response: Interval, in seconds, from the beginning of the episode to the first codeable anger response (*e.g.*, facial, vocal, bodily).
* If the child is already displaying anger before the episode starts, the code = 1.
* If the child never displays anger, the code = 181.

i. Intensity of sadness expression: Peak intensity of facial sadness is coded in each epoch using AFFEX (see Appendix for definitions) and is rated on the following scale:
0 = No facial region shows codeable sadness movement.
1 = Only one facial region shows codeable sadness movement, identifying low intensity sadness, or expression is ambiguous.
2 = Two facial regions show codeable sadness movement, or expression in one region is definite (*e.g.*, brows).
3 = All three facial regions show codeable sadness movement, or coder otherwise has impression of strong facial sadness.

j. Intensity of bodily sadness: Peak intensity of bodily sadness is coded in each epoch and is rated on the following scale:
0 = No detectable bodily sadness.
1 = Mild bodily sadness (*e.g.*, downcast head).
2 = Moderate bodily sadness (*e.g.*, slumped shoulders).
3 = High bodily sadness (*e.g.*, head in arms or hands).

k. Latency to first fear response: Interval, in seconds, from the beginning of the episode to the first codeable fear response (*e.g.*, facial, vocal, bodily).

* If the child is already displaying fear before the episode starts, the code = 1.
* If the child never displays fear, the code = 181.

l. Presence of positivity at any time during the epoch is noted, so long as it is obvious. Examples may include positive verbalizations or positive facial or bodily expressions.
0 = no
1 = yes

m. Presence of negativity at any time during the epoch is noted, so long as it is obvious. Examples may include negative verbalizations or negative facial or bodily expressions.
0 = no
1 = yes

n. Language valence: The **tone** of the child's voice is coded for verbalizations. The peak intensity is coded in each epoch using the following scale:
0 = No verbalizations or positive valence (*e.g.*, laughter).
1 = Neutral valence.
2 = Low negative valence.
3 = High negative valence.

o. Language content: The **content** of the child's verbalizations is coded. The peak intensity is coded in each epoch using the following scale:
0 = No verbalizations or positive content.
1 = Neutral content.
2 = Low negative content: Indirect refusal to continue (*e.g.*, "This isn't the one I wanted.").
3 = High negative content: Direct refusal to continue (*e.g.*, "I don't want this prize!").

Interval 5 (Child picks out gift)

p. Presence of positivity at any time during the epoch is noted, so long as it is obvious. Examples may include positive verbalizations or positive facial or bodily expressions.
0 = no
1 = yes

q. Presence of negativity at any time during the epoch is noted, so long as it is obvious. Examples may include negative verbalizations or negative facial or bodily expressions.
0 = no
1 = yes

Wrong Gift

I.D. # _____ Child's name and last initial: _____ Date of birth: _____

Scorer: _____ Date scored: _____ Date of visit: _____

Tape # _____ Counter # _____ Episode order: _____

Child tester: _____

Baseline (coded by child tester) _____

Interval 1
Waiting for gift

Latency to first positivity _____secs.

5-second epochs	1	2	3	4	5	6
Time (Begin)						
Intensity of anticipatory positive affect (0 - 3)						
Presence of positivity (0 = no; 1 = yes)						
Presence of negativity (0 = no; 1 = yes)						

Interval 2

Friendly stranger brings in gift and waits

Latency to first anger response (For Intervals 2, 3, and 4) _____ secs.
Latency to first sadness response (For Intervals 2, 3, and 4) _____ secs.

5-second epochs	7	8	9 33	10	11	12
Time (Begin)						
Intensity of anger expression (0 - 3)						
Intensity of bodily anger/frustration (0 - 3)						
Intensity of sadness expression (0 - 3)						
Intensity of bodily sadness (0 - 3)						
Language valence (0 - 3)						
Language content (0 - 3)						
Presence of positivity (0 = no; 1 = yes)						
Presence of negativity (0 = no; 1 = yes)						

Scary Mask

A. RATIONALE

A friendly stranger puts on a scary mask. The incongruity, and the fact that it is a "scary stranger," will scare some children.

B. PHYSICAL SETTING

The child is standing on a mat, facing the camera. The friendly stranger is facing the wall digging through a bag so that the child cannot see the mask. The child tester and the parent should not be present during this episode. The child should believe they are alone.

C. STIMULI

Black hooded sweatshirt, a rubber mask that resembles a disfigured face, and one 17" x 17" floor mat

D. PROCEDURE

The episode begins when the child sees the friendly stranger's face (usually when the friendly stranger coughs) for the first time. It ends after the friendly stranger says, "Well, it was nice to meet you. I have to go now."

E. CAMERA INSTRUCTIONS

The camera should be focused on the child from his/her knees to the top of his/her head. It is crucial to get as tight a shot as possible to aid in facial coding.

F. SCORING

The episode is divided into six epochs:

The first epoch begins when the friendly stranger coughs the first time (or when the child first sees the friendly stranger's face) even if child is not looking towards the friendly stranger, and ends when the friendly stranger coughs the second time. This is about 15 seconds.

The second epoch begins when the friendly stranger coughs the second time and ends as the friendly stranger starts to say, "Hi, my name is Jamie." This is about 15 seconds.

The third epoch begins when the friendly stranger says, "Hi, my name is Jamie," and ends as the friendly stranger takes off the mask (or when the friendly stranger begins saying, "Hi! It's just me. I was just wearing this mask,"). This is about 15 seconds.

The fourth epoch begins as the friendly stranger takes off the mask (or when the friendly stranger begins saying, "Hi! It's just me. I was just wearing this mask,"), and ends as the friendly stranger asks the child to touch the eyes and mouth on the mask.

The fifth epoch begins as the friendly stranger asks the child to touch the eyes and mouth on the mask, and ends as the friendly stranger asks the child to put the mask on. If the child puts the mask on stop coding as soon as you can no longer see the child's face.

The sixth epoch begins as the friendly stranger asks the child to put the mask on, and ends after the friendly stranger says, "Well, it was nice to meet you. I have to go now."

Variables to be scored

a. Baseline (coded by the child tester)
b. Gender of friendly stranger
c. Friendly stranger's behavior
d. Latency to first clear fear response
e. Intensity of fear expression
f. Intensity of vocal distress
g. Intensity of bodily fear
h. Presence of startle
i. Intensity of avoidance
j. Intensity of approach
k. Duration of gaze aversion
l. Intensity of fidgeting
m. Presence of negativity
n. Intensity of positivity
o. Cooperation or refusal

Definitions of variables:

a. Baseline state: The child's state prior to the beginning of the episode is coded by the child tester using the following scale:
   1 = alert and calm
   2 = alert and active
   3 = tired or inattentive
   4 = whiny, complaining or distressed
   5 = resistant

b. Gender of friendly stranger: The gender of the friendly stranger is noted.
   1 = female
   2 = male

c. Friendly stranger's behavior is coded using the following scale:

0 = Completely neutral
1 = Slightly too friendly or mostly friendly
2 = Overly friendly or "over the top" friendly (ex. Deviating from script in a reassuring and comforting way)

d. Latency to first clear fear response: Interval, in seconds, from the beginning of the episode to the first codeable fear response. Include facial fear, bodily fear (only if greater than 1), startle, vocal distress, and avoidance
* If the child is already displaying fear before the episode starts, the code = 1.
* If the child never displays fear, the code = 181.

e. Intensity of fear expression: Peak intensity of facial fear is coded in each epoch using AFFEX (see Appendix) and is rated on the following scale:
0 = No facial region shows codeable fear movement.
1 = One facial region shows codeable fear movement, identifying low intensity fear. Can be ambiguous, corners of mouth go straight back, eyebrows go up and in, eyes look tense, usually short lived
2 = Two or three facial regions show codeable fear movement. Should be obvious.

f. Intensity of vocal distress: Peak intensity of vocal distress is coded in each epoch and is rated on the following scale:
0 = No vocal distress. Either do not talk or say something in a neutral or positive tone.
1 = Low or moderate vocal distress: Mild vocalizations that indicate some fear (*e.g.*, "Is that a mask"). Is typically subtle and is often a question. Sarcasm often falls into this code as well.
2 = High vocal distress: Vocalizations that indicate definite, very fearful overtones (*e.g.*, "Don't come closer."). Is typically more obvious. Child may scream, cry or use the word "scared".

g. Intensity of bodily fear: Peak intensity of bodily fear is coded in each epoch and is rated on the following scale (scale is pretty sensitive):
0 = No signs of bodily fear.
1 = Low bodily fear: Child may show a sudden decrease in activity. This decrease in activity cannot be related to visual orienting to the friendly stranger.
2 = Moderate bodily fear: Child may show bodily tensing (visible tensing of the muscles such as tightening and raising of shoulders). For example, the child may raise his/her shoulders to sigh or control breathing or ball hands into tight fists. Fidgeting does not count. (ex. Hands on hips, crossing their arms, odd/tense arm or neck position lasting for a couple seconds or more.)
3 = High bodily fear: Child may exhibit freezing (tensing entire body with very little motion) and/or trembling. This code is rather uncommon and should be obvious. Running out of room does not count for this code.

h. Presence of startle: Presence of startle response during the **first epoch after cough** is coded. A startle response must not include a blink response. A jolt or jump by the child that is easily noticed.
0 = no
1 = yes

i. Intensity of avoidance: Peak intensity of avoidance behaviors. These can be such things as increasing the distance between child to friendly stranger (e.g., leaning away), putting something between the child and the friendly stranger (e.g., putting hands over face), or subtle bodily behaviors (e.g., closing eyes). They are coded in each epoch and are rated on the following scale:
If the child made a move to approach but then returned to his/her original spot, it is not considered avoidance. It is avoidance only if the initial distance between child and the friendly stranger is increased.
0 = No avoidance: Child may stand in place.
1 = Low avoidance: Child may turn close to 90°, lean away and hold it, or close his/her eyes.
2 = Moderate avoidance: Child may take one or two steps away (even if they are small steps) from friendly stranger. If the child covers their face it is also scored.
3 = High avoidance: Child may display more than two of the above examples or may be as far away from the friendly stranger as possible (e.g., in a far corner, hiding behind something).

j. Intensity of approach: Peak intensity of approach behaviors (behaviors initiated by child to decrease the distance from child to the friendly stranger) are coded in each epoch and are rated on the following scale:
If the child made a move to avoid but then returned to his/her original spot, it is not considered an approach. It is an approach only if the initial distance between the child and the friendly stranger is decreased.
0 = No approach: Child may stand in place; does not approach friendly stranger.
1 = Low approach: Child may lean toward friendly stranger. Reaching for the mask before being prompted
2 = Moderate approach: Child may take one or two steps (even if they are small steps) toward the friendly stranger.
3 = High approach: Child may take more than two steps toward the friendly stranger. This is an uncommon code.

k. Duration of gaze aversion: Duration of gaze aversion is scored for each instance of gaze aversion on the last page. The gaze aversions must be **longer than two seconds**. The total number of seconds that the gaze aversion lasted is recorded. Do not code duration of gaze aversion for epochs that are missing.

l. Intensity of nervous fidgeting: Intensity of nervous fidgeting is noted during each epoch. Nervous fidgeting is defined as movement without a purpose that is induced by the situation (e.g., rocking behavior). Picking up an object and fiddling with it can be scored as fidgeting.
   0 = No nervous fidgeting.
   1 = Mild to Moderate nervous fidgeting. (ex. Close to body movements, playing with hair, scratching, etc.)
   2 = High nervous fidgeting. (ex. Big movements, obvious rocking, flapping arms, etc.)

m. Presence of negativity at any time during the epoch is noted, so long as it is obvious. Examples may include negative verbalizations or negative facial or bodily expressions. If any facial fear, negative vocalizations or avoidance are coded it is coded a 1 for negativity. Bodily fear must be a 2 or higher to count.
   0 = no
   1 = yes

n. Intensity of positivity at any time during the epoch is noted, so long as it is obvious. Examples may include positive verbalizations or positive facial or bodily expressions.
   0 = No positivity.
   1 = Tense, small, restrained smiles or small giggles. Child is at least a little uncomfortable
   2 = "True," joyful smiles and unrestrained laughter. Child is comfortable

   - Often times a closed mouth smile is a 1 while a larger open mouth smile is a 2

o. Cooperation or refusal: The friendly stranger asks the child a series of questions, and his/her cooperation or refusal is scored.
   0 = Child refuses.
   1 = Child cooperates.

Scary Mask

I.D. # _____ Child's name and last initial: _____ Date of birth: _____
Scorer: _____ Date scored: _____ Date of visit: _____
Tape # _____ Counter # _____ Episode order: _____

Child tester: _____

Baseline (coded by the child tester) _____

Gender of friendly stranger: (1 = female; 2 = male) _____

Friendly stranger's behavior (0 – 2) _____

Latency to first fear response _____ secs.

Interval	1	2	3	4	5	6
	First cough until second cough (~15 secs.)	Second cough until FS speaks (~15 secs.)	FS speaks until mask is off (~15 sec)	Touch the mask?	Touch the eyes and nose?	Put the mask on?
Time (Begin/End)						
Intensity of facial fear (0 - 2)						
Intensity of vocal distress (0 – 2)						
Intensity of bodily fear (0 - 3)						
Presence of startle (0 = no; 1 = yes)						
Intensity of avoidance (0 - 3)						

Intensity of approach (0 - 3)						
Presence of gaze aversion (0 = no; 1 = yes)				[redacted]		
Intensity of fidgeting (0 - 2)						
Presence of negativity (0 = no; 1 = yes)						
Intensity of positivity (0 –2)						
Cooperation (1) or refusal (0)?	[redacted]					

Verbalizations (1 = yes; 0 = no) _____

Scary Mask I.D. #

_____

Duration of gaze aversion

	1	2	3	4	5
Begin time					
End time					
Duration (in seconds)					

Storytelling

A. RATIONALE

This episode will assess social inhibition and shyness.

B. SCORING

This episode is divided into 10-second epochs and can be of various lengths depending on the individual child. There must be at least 6 seconds in the last epoch to be counted as an epoch. There is always at least one prompt by the CT, but depending on the individual child there could be numerous prompts, and the episode could last anywhere from slightly over a minute to over twelve minutes in length. If the episode is longer than 6 minutes, only code the first five minutes and the last minute (total of 36 epochs). The episode begins when the CT finishes giving the instructions. The episode ends when the CT begins corresponding verbally in conversation with the child. If audio is missing for the entire episode, then it is deemed uncodeable. Fill in the coding sheet with 9993 as missing values.

Variables to be scored:

a. Baseline state (scored by the CT)
b. Latency to first fear response
c. Latency to first fidgeting
d. Latency to begin speaking
e. Percent time speaking
f. Presence of partially-voiced/whispered speech
g. Number of disfluencies/hesitations
h. Intensity of facial fear
i. Intensity of nervous fidgeting
j. Intensity of bodily fear
k. Intensity of avoidance behavior
l. Presence of negativity
m. Number of prompts by CT (regarding speech)
n. Number of prompts by CT (regarding standing)
o. Presence of smiling

NOTE: DETERMINE BASELINE BODILY AND FACIAL CHARACTERISTICS BEFORE CODING.

Definitions of variables:

a. Baseline state: The child's state prior to the beginning of the episode is coded once prior to each episode for each child by the CT using the following scale:
   1 = alert and calm
   2 = alert and active
   3 = tired or inattentive
   4 = whiny, complaining or distressed
   5 = resistant

b. Latency to first fear response (excluding fidgeting): Interval, in seconds, from the time when the CT finishes giving the instructions to the first definite fear response (**only** include facial fear, bodily fear, avoidance behavior). If child never shows a fear response during the entire episode code 8888. If the child begins to show a fear response immediately code with a 1.

c. Latency to first fidgeting: Interval, in seconds, from the time when the CT finishes giving the instructions to the first fidgeting behavior. If child never fidgets during the entire episode code as 8888. If the child begins to fidget immediately code with a 1.

d. Latency to begin speaking: Interval, in seconds, from the time CT finishes giving the instructions to when child begins speaking. If child never speaks during the entire episode code as 8888. If the child begins to speak immediately code with a 1. **Disfluencies/hesitations are considered words.**

e. Percent of time speaking is the percentage of time the child spends speaking compared to the total time the child "has the floor." The amount of time the child "has the floor" starts when the CT finishes the introduction and ends when the CT concludes the child's portion of the episode. This value cannot exceed 360.

f. Presence of partially-voiced/whispered speech: Note presence of partially-voiced/whispered speech. Remember that just because the volume of the tape is low does not mean that the child was speaking softly/whispering. Do not include drops in speech at end of sentence (trailing off). Do not code if drop in volume is due to body positioning (e.g., turning away of face or body). Code as 9998 if child says nothing during the epoch.
   0 = Normal volume of speech
   1 = Presence of partially-voiced/whispered speech

g. Number of disfluencies/hesitations is the count of disfluencies/hesitations said by the child during an epoch (e.g. Um, Well, Ah, Hmmm, Uh, repetitions, pauses, filler words such as and, then, so that are carried out in duration). Disfluencies/hesitations that are pauses require more than 2 seconds and less than a 10 second break in speech. So, pauses of 3 through 9 seconds in duration are coded as a disfluency/hesitation. *If a disfluency/hesitation separates a pause (e.g., pause, disfluency/hesitation, pause), the pauses are treated as two separate pauses.* Do not code pauses that occur during questioning by the CT. Pauses of 10 seconds or more are considered normal breaks in speech, not disfluencies. If a pause occurs across two epochs, count it as one disfluency/hesitation during the first epoch it occurs in. Repetitions occurring between epochs should be coded in the first epoch with which they occur. Words/phrases that are said more than once in a row (i.e. "Well…well…well…") should be coded as one repetition. Code as 9998 if child says nothing during the epoch.

h. Intensity of facial fear: Peak intensity of fear or fear blends are noted in each epoch using AFFEX, (See Appendix for definitions), and rated on the following scale:
0 = No facial region shows codeable fear movement
1 = Only one facial region shows definite codeable movement or mild movement in two regions, identifying a low intensity fear.
2 = Two facial regions show definite codeable movement.
Note: If mouth or brow is blocked (half-face exposed), try to determine the movement as best you can. Please only code if you are confident of the code. Code as 0 if face is not entirely visible and note on the coding sheet that you can only see a portion of the face and whether you could or could not determine what is happening in the blocked region.

i. Intensity of Nervous Fidgeting: Intensity of nervous fidgeting is noted during each epoch. Nervous fidgeting is defined as repetitive movement without a purpose that is induced by the situation (e.g., rocking behavior). Do not include behaviors that seem related to balancing/readjusting posture. To determine intensity you should consider both the extent of the body used in the fidgeting behavior and the intensity with which the fidgeting occurs. Consider the amount of space surrounding the movement and the speed at which the movement is occurring. Levels of either of these can determine the level of the fidgeting behavior.
0 = No nervous fidgeting.
1 = Mild to Moderate nervous fidgeting (e.g., small hand movements, rocking, up on toes)
2 = High nervous fidgeting (e.g., kicking, big arm swinging).
Note: This variable is considered missing if the entire body is off screen the entire epic (9995).

j. Intensity of Bodily Fear: Peak intensity of bodily fear is noted in each epoch and rated on the following scale:
0 = No sign of bodily fear.
2 = Moderate bodily fear. Child may show bodily tensing: visible tensing of the muscles such as tightening and rising of shoulders. For example, the child may raise his/her shoulders to sigh or control breathing or balling hands into tight fists. *The shoulders need to remain raised for 2 seconds or more to code bodily tension.* Another example is the child may have tense neck/shoulders where the child doesn't move the head to look around and only eyes move. Another example is the child may clench hands into tight fists.
3 = High bodily fear. Child may exhibit freezing (tensing entire body with very little motion) and/or trembling.
Note: Bodily tensing or freezing must occur for at least 2 seconds to be coded as Bodily Fear. Exceptions are clear shrug gestures.
Note: Tension required to hold a position is not considered bodily fear. However, if it is possible to hold the position without tension, code this as bodily fear.
Note: This variable is considered missing if the entire body is off screen the entire epic (9995).

k. Intensity of Avoidance Behavior: Peak intensity of avoidance behaviors (behaviors initiated by C to maintain or increase distance from C to CT and Camera Person) are noted in each epoch and rated on the following scale: If C made a move to approach or avoid but then returned to his/her original spot, *the return to the original spot* is not considered avoidance. It is avoidance only if the initial distance between C and the CT and Camera Person is increased or decreased. Leaning towards the CT and Camera Person is not considered avoidance behavior, but leaning to the side or away is.
0 = No avoidance behavior. Stands in place.
1 = Mild avoidance behavior. Examples may include turning away 90 degrees from CT and Camera Person, covering the mouth, looking down (need to see top of head), tucking face into the chest/shoulder, crossing arms protectively (across chest), or mildly leaning away.
2 = Moderate avoidance behavior. Examples include taking one or two steps away from, towards, or to the side of CT and Camera Person, covering face, moderately leaning away, turning more than 90 degrees.
3 = High avoidance behavior. Examples include sitting down, taking more than two steps away from, towards, or to the side of CT and Camera Person, being off the mat, walking out of the camera's view or combinations of the above behaviors.

Note: When several avoidance behaviors are present during a single epoch, use these formulas:
- Any combination of behaviors that are considered mild (1) should be coded as a 1.
- Combinations involving several mild behaviors (1) along with a single moderate behavior (2) should be coded as a 2.

- Combinations involving more than one moderate behavior (2) should be coded as a 3.

Note: If head turn is less than 2 seconds, verify that it isn't just a distracter that has attracted attention (e.g., dog barking in next room). It should be clear, however, that there is a distracter present. Such cases should not be coded as avoidance behavior. When a distracter is present, code as 9987.
Note: Head turns and head down need to last at least a second in duration.
Note: This variable is considered missing if the entire body is out of view for the entire epic (9995).

l. Presence of Negativity at any time during the epoch is noted, so long as it is obvious. Examples may include negative verbalizations or negative facial or bodily expressions. Nervous Fidgeting is **not** considered as Negativity. Bodily fear, facial fear, avoidance behavior, and any sadness or anger facial, vocal, or bodily expression is considered negativity.
0 = No
1 = Yes

m. Number of prompts regarding speech given by the CT. This would include instances such as the CT says, "Is there anything else you would like to tell me about yesterday?" or "Is there anything else?"

n. Number of prompts regarding standing on mat given by the CT. This would include instances such as CT says, "Please stand up." or "Could you stand on the mat please?"

o. Presence of smiling: Presence of smiling using AFFEX (See Appendix for definitions - **smiling with both corners of mouth raised**) is scored for each instance of smiling on the last page.
Note: Smiling during speech is coded if words don't require upward movement of corners of mouth.
Note: Unilateral smiles (one corner raised) are **not** coded as smiles.
Note: If hand or other object is blocking the mouth, only code smiling if you are confident that the child is smiling and you can see enough to determine this. If you are unsure, code as 0. If the hand or object seems to be pressing up on the mouth, then don't code smiling since it may be just movement caused by the hand/object.

Storytelling

I.D. # WTP_____ A B Date scored: _____
Scorer: _____ Counter # _____
Tape # _____ Episode order: _____
Child tester: _____

Baseline (coded by child tester) _____
Latency to first fear response ____ secs
Latency to first fidgeting behavior ____ secs
Latency to begin speaking after CT introduction in seconds _____ secs
Percent of time speaking _____/_____ = _____ %
Episode End Time _____

10 second epochs	1	2	3	4	5	6
Time (Begin)						
Presence of Partially-voiced Speech (0 = no; 1 = yes)						
Number of Disfluencies &Hesitations (0 if none)						
Intensity of Facial Fear (0-2)						
Intensity of Nervous Fidgeting (0-2)						
Intensity of Bodily Fear (0-3)						
Intensity of Avoidance (0-3)						
Presence of Negativity (0 = no; 1 = yes)						
Number of prompts by CT (speech-related) (0 if none)						
Number of prompts by CT (standing-related) (0 if none)						
Presence of Smiling (0 = no; 1 = yes)						