THE MEASUREMENT OF EMOTION REGULATION:

A CONFIRMATORY ANALYSIS

by

DEBORAH JEAN ETTEL

A DISSERTATION

Presented to the Department of Special Education
and Clinical Sciences
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2009

University of Oregon Graduate School

Confirmation of Approval and Acceptance of Dissertation prepared by:

Deborah Ettel

Title:

"The Measurement of Emotion Regulation: A Confirmatory Analysis"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Special Education and Clinical Sciences by:

Kenneth Merrell, Chairperson, Special Education and Clinical Sciences
Cynthia Anderson, Member, Special Education and Clinical Sciences
Joe Stevens, Member, Educational Leadership
John Seeley, Member, Not from U of O
Jean Stockard, Outside Member, Planning Public Policy & Mgmt

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

June 13, 2009

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.

An Abstract of the Dissertation of

Deborah Jean Ettel for the degree of Doctor of Philosophy

in the Department of Special Education and Clinical Sciences

to be taken June 2009

Title: THE MEASUREMENT OF EMOTION REGULATION: A CONFIRMATORY ANALYSIS

Approved: _____
Kenneth Merrell, Ph.D.

The increasing incidence of depression worldwide has led the World Health Organization to predict that depression will be the second leading global burden of disease by 2020. Since depression is often characterized by suboptimal emotion regulation, one of the potential pathways for understanding the transmission of depression risk is through the examination of early emotion regulation development, specifically in a known at-risk group: offspring of depressed parents. A substantial body of literature underscores the myriad ways in which offspring of ever-depressed parents differ from offspring of never-depressed parents, particularly in their development of emotion regulation, and level of risk for affective disorders. Emotion regulation was defined, along with its putative component dimensions, within the context of several well developed temperament models.

This study examined emotion regulation in toddlers through data from the Infant Development Study, a longitudinal study of infant development which included parents from the Oregon Adolescent Depression Project and their offspring. A measurement model of emotion regulation based upon mother reports of toddler behavior was developed and tested as a first step in exploring this putative risk pathway. Confirmatory factor analysis was used to test three measurement models for absolute and comparative fit. A three factor model with dimensions of Negative Affectivity, Surgency, and Effortful Control, was the best fitted model of those tested. Following this aspect of the study, structural models with outcomes of problem behavior were also tested in order to examine the concurrent and predictive validity of the measure. The best fitting model was found to be significantly associated with concurrent toddler problem behavior and predictive of later toddler problem behavior, including internalizing, externalizing, and aggressive behaviors. Recommendations are presented for future study of emotion regulation as a risk transmission pathway.

CURRICULUM VITAE

NAME OF AUTHOR: Deborah Jean Ettel

PLACE OF BIRTH: Eugene, Oregon

DATE OF BIRTH: April 8, 1958

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon

DEGREES AWARDED:

Doctor of Philosophy, School Psychology, 2009, University of Oregon

Master of Science, Educational Psychology, 1982, University of Oregon

Bachelor of Science, Psychology, 1980, University of Oregon

AREAS OF SPECIAL INTEREST:

Research Methodology and Statistics

Depression, Affective Disorders, Neuropsychology

PROFESSIONAL EXPERIENCE:

Research Methodologist, Teaching Research Institute Eugene
A Division of Western Oregon University

Adjunct Instructor, University of Oregon, 1 year
Research Methods, Tests & Measurement in Education

Graduate Teaching Fellow, University of Oregon, 3 years
Statistics I & II, Tests and Measurement in Education

School Psychologist, Eugene School District #4J, 12 years

GRANTS, AWARDS AND HONORS:

Recognition Incentive Grant: *Eligibility Requirements for Talented and Gifted Educational Programs, Eugene School District #4J, 1993*

Recognition Incentive Grant: *Integration of Students with Disabilities into Mainstream Dramatic Activities in Elementary School*

Phi Beta Kappa, University of Oregon chapter, member since 1981

Oregon Scholar Award, 1975

Dean's List, College of Arts and Sciences, University of Oregon

PUBLICATIONS:

Ettel-Hetrick, D., (1995). Small woman, big power: Tae Kwon Do, *Black Belt Magazine.* Cruz Bay Publishing, Inc., Active Interest Media: El Segundo, CA

Severson, H., Pickett, M., & Hetrick, D. (1985). Comparing preservice, elementary, and junior high teachers' perceptions of school psychologists: Two decades later. *Psychology in the Schools, 22,* 2, 179-186.

ACKNOWLEDGMENTS

I wish to express sincere appreciation to Dr. Ken Merrell and Dr. John Seeley for their support of this study. In addition, special thanks to Dr. Laura Backen Jones, Dr. Lisa Sheeber, Dr. Nick Allen, Mr. Derek Kosty, and Mr. Matthew Sweeney for their expert assistance. Thanks are especially due to my very patient family, including my husband (Mark), children (David and Anna), and parents (Ralph and Jean) for all they have given over these many years. Their support has been invaluable.

To my family

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

Figure Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

One of the potential pathways for the transmission of familial depression is through the development of emotion regulation. That is, since depressive disorders involve the suboptimal regulation of emotion, perhaps this suboptimal regulation is passed from parent to offspring through some mechanism that is amenable to intervention. Emotion regulation, then, and its measurement, is at the heart of this study. First, however, the seriousness, nature, and scope of depression will be described in order to provide the context into which the examination of emotion regulation fits. Familial aggregation of depression is associated with a subtype of Major Depressive Disorder that is defined by early onset, severity, increased likelihood of recurrence, and treatment resistance (Mondimore & Potash, 2006; Sullivan, Neale, & Kendler, 2000; Weissman, Warner & Wickramaratne, 2005). Offspring of depressed parents have about a three-fold risk of developing an affective disorder, compared with offspring of never-depressed parents (Garstein & Fagot, 1998; Hammen & Brennan, 2003; Sullivan, Neale & Kendler, 2000; Weissman, et al., 2005). Depression rates are rising worldwide (World Health Organization, 2004), and offspring of depressed individuals comprise a known risk group for developing affective disorders (Downey & Coyne, 1990; Marmorstein, Malone, & Iacono, 2004; Merikangas, Dierker, & Szamari, 1998).

The offspring of depressed parents provide a unique opportunity for early intervention and prevention for a known at-risk group. Early intervention is particularly important because of the progressive nature of depression (Downey & Coyne, 1990; Lewinsohn, Rohde, Seeley, Klein, & Gotlib, 2000; Post, Rubinow, & Ballenger, 1986) and its increasing toll on humanity (World Health Organization, 2004; Substance Abuse and Mental Health Services Administration, [SAMHSA] 2004). In the familial subtype of depression, it is hypothesized that both a genetic predisposition (Holmans, Weissman, Zubenko, Scheftner, Crowe, DePaulo, et al., 2007; Mondimore & Potash, 2006; Propper & Moore, 2006; VanMeenen & Wigfield, 2005) and environmental context (including specific parenting behaviors) contribute to offspring risk (Calkins & Hill, 2007; Champagne & Meeney, 2001; Garstein & Fagot, 2003; Moore & Calkins, 2004). Although the results are mixed, suboptimal emotion regulation during infancy ("colic", fussiness/difficulty, difficulty soothing, and withdrawal) has been associated with later affective difficulties (Garstein & Fagot, 2003; Moore & Calkins, 2004; Porges, 1992).

This study seeks to complete one of the foundational steps in testing the hypothesis that, in large part, it is suboptimal emotion regulation that is the precursor to affective difficulties passed from parent to child: a reliable, valid measurement model of emotion regulation. The putative pathway to development of affective disorders is as follows: the bidirectional influence of genetic predisposition, environmental context (including parenting behaviors), and modeling associated with depression history predicts suboptimal emotion regulation, which, in turn, predicts increased risk for affective disorders.

Before embarking on this complex journey, a valid method for measuring emotion regulation in infants must be created. From there, the relations between infant emotion regulation, context, parental characteristics, and later behavioral or affective difficulties experienced by the child may be examined.

Emotion regulation has been defined as set of emotional, cognitive, behavioral and interpersonal skills which regulate and moderate the experience and expression of human emotions (Eisenberg & Morris, 2002; Forbes & Dahl, 2005; Posner & Rothbart, 2002; Rothbart & Sheese, 2007). In those individuals who do not develop these regulation skills normally, emotional reactions and feelings are experienced as more intense and intractable than those experienced by emotionally healthy individuals. These deficiencies in the ability to regulate ones' own emotional state contribute significantly to the experience of depression, and may originate in genetic predisposition and types of parenting behaviors associated with depressed individuals (Posner et al., 2002). Developmental and behavioral psychology research has shown that depressed parents tend to parent their children differently from non-depressed parents in ways that affect the development of emotion regulation in the child (Burt, Van Dulmen, Carlivati, Egeland, Sroufe, Forman, Appleyard, & Carson, 2005; Downey & Coyne, 1990; Lovejoy, Graczyk, O'Hare, & Neuman, 2000).

Because emotion regulation is one of the critical developmental tasks of infancy and early childhood, difficulty obtaining self-regulatory skills has been associated with suboptimal affective and behavioral health (Calkins, 1994; Cicchetti, Ackerman, & Izard, 2002). Some studies show that even when previously-depressed parents no longer

demonstrate symptoms of depression, their offspring show poorer outcomes than offspring of never-depressed (Garber & Flynn, 2001). Thus, current parenting behaviors, previous parenting behaviors, and genetic history may individually, and in combination, lead to the increased offspring risk.

To test the premise that risk for affective disorder is increased for offspring through the suboptimal development of emotion regulation, a reliable, valid model of emotion regulation is required (Cole, Martin, & Dennis, 2004). Validation of such a model is long overdue, in part, because measurement of infant development is fraught with conceptual and practical difficulties. There are widely differing conceptualizations of the construct of emotion regulation – a subjective change in subjective emotional state occurring dynamically (micro-momentarily) in infants who can't talk or describe how they are feeling. Emotion regulation must be distinguished successfully from emotion or emotional reactivity, and self-regulation from the effects of the regulatory behavior of others (Cole, Martin, & Dennis, 2004; Sroufe, 2000). Finally, the essential components of emotion regulation must be defined and measured accurately.

This study made use of an extant data set of measurements of infant affect in the Infant Development Study (IDS) which includes participants who are offspring of parents previously diagnosed as depressed during adolescence (Lewinsohn, Rohde, Klein, & Seeley, 1999). Through analysis of these IDS data, primarily using the conceptual framework of Posner & Rothbart (2003), this study examined the construct of emotion regulation as the putative interplay between three important dimensions: Negative Affectivity, Effortful Control, and Surgency (Posner et al., 2003; Rothbart & Hwang,

2003). By attempting to measure these three constructs with selected indicator variables, the IDS data were used to fit a measurement model based upon mothers' report of infant behavior. This model provided an operationalized picture of emotion regulation in offspring, a first step in examining the hypothesized early roots of affective disorders. One distinct advantage of this study over other studies of infant emotion regulation was that along with the measurement of emotion regulation, there was a wealth of detailed, longitudinal data about each infant's family, including: clinical diagnoses of parental depression or other mental disorders, familial aggregation of disorders, marital discord, and perceived caregiver support, among other measures.

First, separate theoretical models were tested for goodness-of-fit and for relative fit among models. A well fit model was finally specified, but only with the *post hoc* allowance of two sets of correlated indicant residuals, meaning that the original models did *not* fit well as originally specified. A discussion of potential reasons for model misfit followed, along with an explanation of the theoretical and analytical consequences of allowing residuals to correlate to improve model fit. Using the three factor model (with the correlated residuals), additional analyses examined the model's ability to predict toddler outcome in internalizing behavior problems. It was hoped that in a future study, a well-fit model could be used to identify specific parental or contextual factors associated with optimal and suboptimal development of emotion regulation. Ultimately, it may be helpful in development of appropriate interventions to prevent transmission of depression risk.

The goals of this study were congruent with the original stated goals of the Infant Development Study, in that they sought to provide a mechanism for examining emotion regulation in infants/toddlers as a putative transmission pathway for depression risk among families. Essentially, this study attempted to provide the first, necessary step in this process of discovery.

In the Infant Development Study, about one third of the sample participants had a family or parental history of depression, one third had a history of other mental health diagnoses, and about one third had no depression or other mental health problem history. Both parents and their offspring were evaluated extensively over time. The value of early intervention with at-risk families should not be underestimated, as there is evidence that improved skills in emotion regulation can potentially buffer a child against familial predisposition to mood disorders (Forbes & Dahl, 2005; Silk, Shaw, Forbes, Lane, & Kovacs, 2006), and that emotional-behavioral experiences may alter development of critical brain substrates involved in the experience and expression of emotionality (Brody, Saxena, Silverman, Alborzian, Fairbanks, Phelps, Huang, et al., 1999; Silk, et al., 2003). The merging of the disciplines of cognitive and developmental psychology, behaviorism, and neuroscience has allowed the consideration of cross-contextual interactions – not possible with the perspective of only one scientific discipline. The study of emotion, it appears, is necessarily interdisciplinary, requiring the inclusive investigation of nature *and* nurture, as well as the complex bidirectional influences of the two.

Why this Study? Why Now?

If there is ample evidence that offspring of depressed parents are at greater risk for poor outcomes, and that emotion regulation may be involved in those outcomes, why is this study necessary? What limitations and issues does it address? First, the studies linking parental depression to infant emotionality have had equivocal results. Some found significant relations (Forbes, Cohn, Allen, & Lewinsohn, 2004; Garstein & Fagot, 2003; Silk, Shaw, Forbes, Lane & Kovacs, 2006), while others found no significant relation between maternal depression and infant temperament (Dawson, Klinger, Panagiotides, Hill, & Spieker, 1992; Whiffen & Gotlib, 1989). It is an understatement to say that developmental research with infants/toddlers can be difficult. From disagreement in defining developmental constructs (Cole, Martin, & Dennis, 2004), attempting to track behaviors that are constantly in flux, to the formidable task of simply managing a longitudinal study with parents and infants, there has been too little evaluation of this population.

Second, this study was unique in that it made use of longitudinal data gathered since early adolescence (from the Oregon Adolescent Depression Project) for this fairly large sample of depressed parents. In addition to the frequent measurement occasions for the infants, the data set contains a wealth of psychological, social, and behavioral history collected for these parents, from their teen years up through young adulthood and parenthood. The data provide more than a cross-sectional snapshot; rather, they represent more of a moving picture of the course of depression across generation, and its putative impact on infant affective development. This is a story that unfolds over time.

While others have similarly measured emotion regulation (Garstein & Rothbart, 2003; Propper & Moore, 2006; Silk, Shaw, Forbes, Lane, & Kovacs, 2006; Whittle, Allen, Lubman, & Yucel, 2006), few studies offer the same wealth of family information available on probands, including longitudinal data on parental mental health diagnoses and personality type, family density of affective and other disorders, marital satisfaction, maternal health, perceived perinatal support, and laboratory dyadic observational data.

Third, the global community is currently in the midst of an economic crisis the magnitude of which the world has not experienced in decades at the same time that the incidence of depression is increasing at an alarming rate. Unfortunately, at a time when the world's focus is on increasing financial uncertainty, critical (and costly) mental health issues such as rising depression rates may be pushed into the background. This is exactly the time when at least *some* of our global attention should be focused on understanding and preventing further emotional suffering.

Finally, the discussion of how and why we regulate our emotional state has been going on throughout human history and is inarguably important to the survival of our human society. Shakespeare's Hamlet, in the play of the same name, wisely noted: "There is nothing either good or bad, but thinking makes it so" (*Hamlet*, Act II, scene ii). Shakespeare may have been echoing the sentiments of Roman poet/philosopher Seneca, who, long before Shakespeare, warned, "He is most powerful who has power over himself" (Davie, 2007, p. 28). The notions that the way we think of something makes it "good or bad" and that a powerful person is one who has "power over himself" both express the desirability of regulating ones' emotions well.

This fits within the context of Cartesian reasoning, that is, the rational mind (thought) ought to take control of the often unwilling body (unregulated emotion).

The wrestling match between emotion and reason is not only a western preoccupation. An ancient Hindu proverb advises: "Conquer your passions and you conquer the world." Further, Hsun Tzu, in the 3rd century B.C.E. offered a similar admonishment, "To yield to man's emotions will assuredly lead to strife and disorderliness . . . "(DeBary, Chan, & Watson, 1960, p.118).

Of the opposing view, philosopher David Hume, asserted that "…reason is, and ought only to be, the slave of the passions" (1711 – 1776). Much later, author D.H. Lawrence (1924 - 1964) updated the assertion that emotion should rule over reason by suggesting, "When genuine passion moves you, say what you've got to say, and say it hot." From either vantage point, it is clear that human society has long been wrestling with the issue of emotion vs. reason writ large, and that the regulation of emotion is a topic well worth our continued examination in light of its value in the prevention of human suffering. In fact, Posner and Rothbart (2000), stated: "We believe that the understanding of self-regulation is the single most crucial goal for advancing an understanding of development and psychopathology." (Posner & Rothbart, 2000, p. 427).

Public Health Benefits

The increasingly staggering emotional and financial toll taken by rising global rates of depression really constitutes a global public health crisis. In turn, awareness and understanding of the mechanism of its transmission could reap great public health

benefits. It was hypothesized that examining emotion regulation skills of the offspring of depressed parents will provide a foundation for study of familial transmission of depression risk and inform development of targeted interventions to reduce that risk to offspring. Parents with depression histories may well benefit from insight into how they can guide emotion regulation development in their children, and how improvement of their own self-regulatory skills may, in turn, reduce passing on risk. By creating and testing measurement models, the hypothetical pathways of risk transmission can be translated into operationalized, testable questions by asking: Which of our theoretical measurement models provides the best representation of the data? Does the model provide convergent and discriminant validity for the measures used? Once validated, does use of the model provide evidence to support emotion regulation as a predictor of later affective or behavioral difficulties? Still further investigation could include examination of those parental and/or environmental characteristics which are most influential in buffering offspring against increased risk.

Research Goal

The goal of this study was to develop and validate a measurement model of emotion regulation in infants/toddlers using extant data, and, subsequent to achieving adequate model fit, attempting to further validate the best-fitting measurement model through other construct validation methods (concurrent and predictive validity). This study involved comparative model fit testing of a one-factor, two-factor, and three-factor model of emotion regulation for use with toddlers.

Beyond this study, longer term goals are to evaluate the potential influence of parental affective disorders on offspring affective difficulties. To that end, the study will consist of development of a measurement model (confirmatory factor analysis; CFA) based upon empirically supported research about the measurement of emotion regulation.

Research questions for this study included:

1. How do we best measure emotion regulation in infants/toddlers ages 12 months to 36 months? What combination of observed items or indicants best represent the construct of emotion regulation?
2. What combination latent constructs best explains individual differences in emotion regulation during this early developmental period (12 months – 36 months)?
3. Is emotion regulation in offspring best explained by a one factor model of Emotion Regulation, a two factor model of Negative and Surgency/Positive Affectivity, or a three factor model of Negative Affectivity, Effortful Control, and Surgency?
3. Does suboptimal emotion regulation at age 24 months predict internalizing problem behaviors at age 36 months?

CHAPTER II

REVIEW OF LITERATURE

The research questions were addressed within this study, using the rich, longitudinal data set from the Infant Development Study begun in 1996. The methods section of this document details the specifics of how the study was enacted, and the results and discussion sections detail the findings and their implications. First, however, several critical dimensions in the familial transmission of depression must be examined, which lay the foundation for the selection and use of the theoretical constructs. The following section—a review of the relevant literature—describes the necessity for research in familial transmission of depression, due to the magnitude of risk for offspring, the types of poor outcomes associated with depression symptoms, and the importance of early intervention and prevention. To this end, the construct of emotion regulation and its place within the general construct of temperament will be examined, along with an examination of specific patterns of emotion regulation (optimal and suboptimal) which have been linked with affect and behavior. Finally, the way in which the operational definitions of key constructs fit within the chosen theoretical framework will be described.

Due to the deleterious effects of Major Depressive Disorder, much study has been devoted to examining familial transmission of the psychopathology (Birmaher, Ryan, Williamson, & Brent, 1996; Downey & Coyne, 1990; Lovejoy, Graczyk, O'Hare, &

Neuman, 2000; Marmorstein, Malone, & Iacono, 2004; Merikangas, Dierker, & Szamari, 1998). Recently, various scientific disciplines have combined efforts to examine the early roots of affective disorders, through neurobiological (Weissman, Warner, Wickramarant Moreau & Olfson, 1997) psychological, behavioral, and developmental perspectives providing a multi-dimensional approach to the study of affective psychopathology (Davidson, Fox, & Kalin, 2007; Lovejoy, Graczyk, O'Hare & Neuman, 2000; Mondimore & Potash, 2006; and Rothbart, 1981). In fact, it is more unusual today to find a study of affective psychopathology solely from one theoretical perspective than it is to find one of combined perspectives. The strongest evidence comes from those studies in which the psychological and developmental theories are congruent with the empirical data coming from neuroscience and behavioral science (Cole, Martin, & Dennis, 2004; Gross, 1998).

The research in depression prevention and treatment is no trivial pursuit. In 2004, the World Health Organization predicted that *by 2020 depression will be the #2 global burden of disease worldwide* (WHO, 2004). Estimates from the Global Burden of Disease study GBD 2000 indicate that unipolar depressive disorders make-up 4.4% of the global disease burden (65 million disability adjusted life years [DALYs] lost in total), in the same range as the total burden attributable to ischaemic heart disease, diarrheal diseases, or the combined impact of asthma and chronic obstructive pulmonary disease (World Health Organization, 2002). Thus, as author William Styron so ably put it, "It is hopelessness even more than pain that crushes the soul." (Styron, 1990, p.56).

The fallout from Major Depressive Disorder can be widespread, and includes increased risk of suicide, diminished school performance, relationship problems, behavior and conduct problems, increased risk of other psychiatric disorders, and increased risk of drug use (Angold & Costello, 1993; Beck, 1987; Clark & Beck, 1999; Hankin & Fraley, 2005; Monroe & Harkness, 2005; Klein, Lewinsohn, Seeley, Rohde, 2001, Lewinsohn, Rohde, Seeley, Klein, & Gotlib, 2000; Rohde, Lewinsohn, & Seeley, 1991). Further, depression is frequently comorbid with other adverse conditions, such as substance abuse and anxiety, conduct, and attention deficit disorders (Lewinsohn, Rohde, Klein, & Seeley, 1999; Lewinsohn, Rohde, & Seeley, 1995). Sadly, depression carries a 30-fold risk of completed suicide (Brent, Perper, Moritz, Liotus, Schweers, Bablch, & Roth, 1994; Martin & Cohen, 2000).

Depression is often a disorder with an early onset, making it the most common psychiatric disorder of adolescence. When onset is quite early in adolescence, individuals frequently demonstrate higher rates of recurrence, progression into chronicity, and continuity into adult forms of mood disorders (Lewinsohn, Clarke, Seeley, & Rohde, 1994; Lewinsohn, Rohde, Klein, & Seeley, 1999; Mrazek & Haggerty, 1994). In fact, one third of all individuals who have had Major Depressive Disorder, say they experienced their first episode before age 21 years (Andrews, Lewinsohn, Hops, & Roberts, 1993).

In 2003, nine percent of adolescents (2.2 million adolescents ages 12 to 17) experienced a Major Depressive Episode (Substance Abuse and Mental Health Services Administration; SAMHSA; 2004). For persons ages 15 to 45 years Major Depressive

Disorder has a 9% fatality rate (Chisholm, Sanderson, Ayuso-Mateos, & Saxena, 2004). Clearly, the negative effects can begin early and become progressively worse without treatment.

Some would argue that the current prevalence reports may underestimate the adverse effects of depression, because even those who do not meet diagnostic criteria for clinical depression are negatively impacted by experiencing subsyndromal symptoms of the disorder. For adolescents, those experiencing subthreshold symptoms, often have prognoses almost as poor as those who *do* meet the diagnostic criteria (Gotlib, Lewinsohn, & Seeley, 1995; Lewinsohn, Solomon, Seeley & Zeiss, 2000; Sadek & Bona, 2000; Steinhausen & Metzke, 2000). The significant psychosocial impairment associated with depression, and the chronicity of its course make subsyndromal symptomatic depression a serious subject for further research (Sadek & Bona, 2000).

Familial Subtype of Depression

Tolstoy's *Anna Karenina* (1877) opened with the famous observation that, "Happy families are all alike; every unhappy family is unhappy in its own way." Many "unhappy families" do seem to carry on this unfortunate legacy, a subtype of Major Depressive Disorder known colloquially as familial depression or alternately, endogenous depression.

Given the globally-destructive nature of depression, even a small reduction in risk is a worthwhile goal, and focusing those known to be at higher risk at birth is one method of helping those most in need of early intervention. The following section discusses

various putative transmission pathways, including parenting practices, genetic influences, and idiosyncratic physiological and temperamental differences in offspring. It provides explanation of and evidence for this subtype of Major Depressive Disorder that appears to "run in families", thus serving as a theoretical foundation for this study's focus on offspring of depressed mothers. Generational or family studies of depression have shown that children of depressed parents are at increased risk for developing psychopathology in general, and affective disorders, in particular (Beardslee, Schultz, & Selman, 1987; Billings & Moos, 1985; Downey & Coyne, 1990; Goodman & Gotlib, 1999; Hammen, 1991; Keller et al., 1986; Marmorstein, Malone, & Iacono, 2004; Orvaschel, Walsh-Allis, & Weijai, 1988; Weissman, Warner, Wickramaratne, 1987).

In addition to being at higher risk for affective disorders, those with family histories of depression are at higher risk for mood and psychiatric disorders, reduced attunement with caregivers, and cognitive delays which may continue into childhood and beyond. In sum, individuals with family history of depression are at risk for poorer outcomes (Weissman, Wickramaratne, Nomura, Warner, Pilowsky & Verdeli, 2006). Maternal depression, especially, has long been associated with poor outcomes in offspring including: (a) specific cognitive impairments and developmental delay (Beach, Henry, Stowe, & Newport, 2005); (b) deficits in infant affective and cognitive behavior (withdrawal, diminished positive affect, increased negative affect, difficulty sustaining attention, failing to persist at tasks, poor mother-child attachment and attunement (Trevarthen, 1994); and (c) differences in psycho-physiological systems – vagal tone, the

parasympathetic nervous system that inhibits heart rate, and asymmetry of anterior EEG (Davidson & Irwin, 1999; Santucci et al. 2008).

Differences in offspring of ever-depressed parents are associated in particular with emotion regulation. The infant's development of emotion regulation is affected by many environmental events or stimuli, including observation, modeling, and social referencing of the depressed parent (Morris, Silk, Steinberg, Myers, & Robinson, 2007). Since depressive symptoms often negatively impact parenting style, attachment relationship (Trevarthen, 1994), and familial emotional climate, this milieu may create the "perfect storm" context for the development of suboptimal emotion regulation. For a parent with difficulty regulating his/her own mood states, encouraging the infant to gain independent regulatory control may be a formidable task.

Even though it seems logical that emotion regulation *should* affect depression risk, equivocal results have been found. Some studies have found no differences in infant temperament between those with depressed & non-depressed mothers (Dawson, Klinger, Panagiotides, Hill, & Spieker, 1992; Mullins, Siegel & Hodges, 1984; Pauli-Pott et al., 2000; Whiffen & Gotlib, 1989). Others found significant differences in offspring temperament dependent upon maternal characteristics (Ayissi & Hubin-Gayle, 2006; McGrath, Records, & Rice, 2008; Moore, Cohn, & Campbell, 1995). Recent related studies of infant emotion regulation showed that early maternal sensitivity rather than infant temperament predicted child emotion regulation after an emotion challenge (Conway & McDonough, 2006). Infants of depressed mothers were rated as more tense, less content, and more likely to become distressed during administration of infant

development scales such as the Bayley Scales of Infant Development (Moore et al., 1995), and depressed mothers and their partners rated their offspring as more "temperamental" than did non-depressed couples (Edborg, Matthiesen, Lundh, & Widstrom, 2005).

Some studies have found significant physiological differences between offspring of depressed and non-depressed, including higher levels of physiological arousal (e.g., heart rate, salivary cortisol) during mother-infant interactions, and lower vagal tone, the parasympathetic nervous system that inhibits heart rate in offspring of depressed mothers (Pickens & Field, 1995; Porges, Doussard-Roosevelt, Portales, & Greenspan, 1996). Vagal reactivity has been described as the "brake" that slows the heart rate after physiological arousal, in order to return the system to homeostasis (Porges, 1992; Porges, et al., 1996). Research in the relation of vagal regulation and emotion regulation has shown interactions with temperament and age. Newborns with high vagal tone have been described as highly reactive, more irritable, and initially less able to soothe themselves; however, when measured again at three months these same infants were better able to soothe themselves than those with lower vagal tone. It may be that at the neonatal age, reactivity was highly adaptive for infants unable to self-regulate. Perhaps in the very early months, a high level of reactivity helped to draw the attention and help of others in regulating the infant's mood state. Porges (1992) found that low vagal tone has been associated with greater emotional reactivity and expressivity in individual. Some indicate that those with high vagal tone may have improved attentional ability, which could serve a protective function in dealing with frustration.

Porges et al. (1996) also found that low vagal tone and difficult regulation also predicted significantly more behavioral problems at three years of age. Clearly, some observable behavioral responses can be correlated with physiological arousal states and transitions. In this case, vagal tone, which involves control over the "braking and accelerating" aspect of heart rate, represents one type of physiological expression of emotion in the body.

While there is ample evidence that certain physiological aspects of human behavior (including emotion regulation) are at least partially genetically-driven, there is also evidence that behavioral and genetic factors impact each other, and that benefits of optimal emotion regulation can be observed in neurophysiological outcomes. Silk, Shaw, Forbes, Lane, & Kovacs (2004) looked at the relation between affect regulation and sleep quality, and found that better emotion regulation related to better transitions into sleep, and that in turn, better sleep promoted improved ability to regulate one's moods. Silk and colleagues (2004) asserted that better sleep regulation could be one of the protective factors that may increase emotional resilience and decrease risk of affective problems. Further, they found significant evidence of cross-contextual mediation, suggesting that the neurological characteristics responsible for emotion processing may be influenced by social context and cognition.

Modern neuropsychological research has demonstrated that brain structures themselves are malleable; that is, neurobiological systems and substrates may be altered by influences from the environment (e.g., exposure to parenting behaviors that encourage infant self-regulation may change activation patterns or brain structures).

Neurophysiological differences have also been found in brain functional activity between offspring of depressed and non-depressed as well. Dawson (1992) and Henriques & Davidson (1990, 1991) found reduced left frontal hemispheric brain activity and increased right frontal hemispheric activity during play in offspring of depressed mothers, a pattern which is associated with depressive symptoms in adults (Davidson & Fox, 1989).

Davidson & Fox (1989), among others, have posited that asymmetry of resting frontal activation may be related to infant temperament, of which emotion regulation is a key dimension. Other neurophysiological differences such as decreased hippocampal volume and amygdalar changes related to depression symptoms and emotion regulation have been noted as well (Davidson 1994), helping to establish a relation between behavioral expression of affect and observable neurophysiological processes.

Goldapple, Segal, Garson, Lau, Bieling, Kennedy & Mayberg (2004) also found evidence of directional changes in the frontal cortex and anterior cingulate gyrus, and changes in hippocampus activation in response to cognitive behavioral therapy. Similarly, Brody, Saxena, Mandelkern, Fairbanks, Ho, & Baxter (2001) found brain metabolic changes associated with observations and reports of symptom improvement in depressed individuals. Again, interdisciplinary research has provided the link between the observed behaviors and the neurological or physiological measures. Both behavioral patterns in infant expression of emotion (negativity, fussiness, difficulty soothing, or state-matching) and neurophysiological responses (such as low vagal tone and stereotypic depressive

patterns of brain activation), appear to link offspring risk of affective difficulties to their development of emotion regulation.

In addition to physiological and neurological correlates, there is a body of evidence supporting the role of genetics in the transmission of emotion reactivity and emotion regulation (Propper & Moore, 2006). It appears that the effects of stressful life experiences may depend in part upon an individual's genetic differences. Recently the 5-HTTLPR serotonin transporter gene with a "short" allele has been associated with mood disorders. Young adult males who had short allele (s/s or s/l candidate gene) were more likely to experience depression symptoms, suicidality, and diagnosable depression following a stressful life event than those without short allele (Caspi, Sugden, Moffitt, Taylor, Craig, Harrington, McClay, et al., 2003). Stanford University's recent Genetics of Recurrent Early Onset Depression (GenRED) project provided much of the recent support for the moderate heritability of Major Depressive Disorder (Levinson, 2005). Most of the work in genetic association with depression has involved the functional polymorphisms or DNA sequence variations that alter the gene expression or functioning of the gene product in the serotonin transporter (SLC6A4), serotonin 2A receptor (5HTR2A) tyrosine hydroxylase (TH; related to dopamine synthesis) or catechol-o-methuyltransferase (COMT; dopamine catabolism). Even though there is some evidence that these foci are associated with mood disorders, results appear mixed, and indicate that depression is probably the result of many gene polymorphisms rather than only one (Levinson, 2005).

Rothbart & Hwang (2003), on the other hand, found mixed results in studies of behavior x genetics interaction, depending upon the parent-report instruments used.

One of the key theoretical assumptions of this study is that further investigation is needed regarding the bidirectional influence of parenting behaviors and genetic influences. As mentioned, one important functional aspect of this line of study is the evidence that just as "brain influences behavior" – behavior also influences brain (Fernald, 2003). This complex evidence demonstrating that changes in social context, behaviors, and cognitions can result in significant brain structural changes (Fernald, 2003; Neville, 1984) is, in fact, an empirical basis for much hope. Silk, Steinberg, & Morris (2003) assert that even if a brain structure is resistant to modification, the potential effects of familial/genetic vulnerability to mood disorders may be *reduced* by increasing an individual's exposure to positive experiences in the environment.
This work supports the hypotheses that early emotion regulation (influenced both by behavior and biology) is related to later affective disorders. Better understanding of the bidirectional influences may help in the development of proactive interventions, promoting resilience to depression through both social context and biological context, as they appear to be mutually strengthening. The first step in evaluating this potential connection is creating a working definition of emotion regulation, within the context of temperament.

Emotion Regulation within Temperament Models

The construct of emotion regulation is one dimension of the broader construct of temperament, a putatively stable and primarily biologically based construct. As Gartstein and Rothbart (2003) described it, temperament consists of "constitutionally based individual differences in reactivity and self-regulation, with constitutional referring to the relatively enduring biological make-up of the individual, influenced by heredity, maturation, and experience." (Gartstein & Rothbart, 2003, p. 6). Temperament is believed to be relatively stable construct according to much empirical study (Buss & Plomin,1984; Janson & Mathiesen, 2008; Thomas, Chess, & Birch, 1968; Posner & Rothbart, 2002). Most of the early work in the field of temperament grew from the New York Longitudinal Study (NYLS; Thomas & Chess, 1977) which, through following and measuring children's behaviors over time, identified the basic dimensions of temperament of activity level, threshold, mood, rhythmicity, approach/withdrawal, intensity, adaptability, distractibility, and attention span/persistence. There is further support for these empirically derived childhood temperament characteristics, in that they are generally congruent with the dimensions found in the "Big Five" studies of adult personality (Cicchetti, Ackerman, & Izard, 1995). These are some of the core temperament dimensions on which the construct of emotion regulation is based.

The idea of exploring temperament, however, is certainly not a 20th century phenomenon. Throughout recorded human history we have been concerned with temperament. To the ancient Greeks and Romans, one's temperament or *corporis habitus* could best be explained in terms of the relative proportions of the four energies one

embodied, i.e., the four humours (Jacques, 1879). Hippocrates (460 – 370 BCE), Aristotle (384 – 322 BCE), and Plato (427 – 348 BCE), for example, all contributed to this philosophy of health (mental and physical) as defined by one's humours. While in general, good health resulted from these four humours being in balance (*temperaterum temperatum*), in particular, the relative strength of the four humours was also believed to define one's innate temperament.

For example, those with the Sanguine Temperament (Blood & Air) were primarily defined by the forces of "blood", and therefore were vital, innately healthy, attractive and full of positive force. Alternately, the Phlegmatic Temperament was predominated by the lymphatic system, in which an excess in lymphatic fluids was believed to cause sluggish circulation, clogging of vital machinery, and generally weak muscles and brain strength as well. These humours reportedly were associated with specific body and personality types that characterized subgroups of the population. The existence of individuals whose physical or mental nature did not fit neatly into one of the four distinct dimensions led to the development of compound humours, ever more complex combinations of humours used to define more sophisticated personality types. As late as 1879 psychology texts were instructing future generations of therapists to apply specific prescriptions for health based upon a person's predominant humour, (Jacques, 1879). Now, well over a hundred years later, we are still examining the dimensions of temperament, still looking for ways the socio-behavioral aspects of temperament (mind) and the physical brain structures (body) are linked.

Definition of Emotion Regulation

"The heart has its reasons of which Reason knows nothing" (Pascal, 1669). Within the broader construct of temperament, the more focused construct of emotion regulation connotes the exertion of control over ones' emotions, both in down-regulation (inhibition) and in up-regulation (expression) of emotions. Philosophers, among others, have long been fascinated by this competition between purposeful reason and unleashed emotion. According to Seneca (40-50 BCE), "Wisdom is only possible when the emotions are silenced and when reason does all the talking." Philosopher Thomas Paine, during the Age of Reason, reiterated this ancient argument in his work *The Crisis*, writing "To argue with a man who has renounced the use and authority of reason… is like giving medicine to the dead" (Paine, 1776). Others have taken the opposite view, as did philosopher Jean-Jacques Rousseau (1712 – 1778) who insisted on a return to reliance on emotion, intuition, and instinct (a natural state) as opposed to the rigidity of reason and rationalism of the Enlightenment. On this societal scale, both views acknowledged the existence of this combination of "brake" and "accelerator" that is emotion regulation. For measurement of emotion regulation the argument must be brought to a micro-level.

Emotion regulation has been defined as set of emotional, cognitive, behavioral and interpersonal skills which regulate and moderate the experience and expression of human emotions (Posner & Rothbart, 2002; Rothbart & Derryberry, 1981). Similarly, Forbes and Dahl (2005) describe emotion regulation as ". . . the internal and external processes involved in the initiation, maintenance, or modification of the quality, intensity, or chronometry of emotional responses" (Forbes & Dahl, 2005, p. 5).

Perhaps a more colloquial description of the construct is one reported by Gross (1998) who states that emotion regulation refers to "…the processes by which we influence which emotions we have, when we have them, and how we experience and express these emotions."

Combining the previous definitions a construct emerges that involves both positive and negatively-valenced emotions, and the processes that may serve to enhance, suppress, and sustain them, or even to replace them with other emotions (Butler & Gross, 2004). Emotion regulation involves the coordination of several interactive cognitive, behavioral and language processes, including control of attention, inhibition of motor responses, planning for goal-directed behavior, and the ability to switch positions (Rothbart & Derryberry, 1982; Rothbart, Derryberry, & Posner, 1994).

Some empirical findings support the assertion that development of optimal emotion regulation is critical in "inoculating" children against future psychopathology. There is evidence that offspring of ever-depressed parents show more difficulty in development of emotion regulation than do offspring of never-depressed parents and there is significant variability in their emotion regulation that is detectable early in life (Silk, Shaw, Forbes, Lane, & Kovacs, 2006). Further, developmental psychologists assert that healthy emotion regulation development during infancy/toddlerhood is a precursor/predictor of later mental, emotional, and behavior health (Porges, 1992; Rothbart, 1981; Santucci, Silk, Shaw, Gentzler, Fox, & Kovacs, 2008) and serves as a necessary and critical developmental task during this period.

The putative link between parents' emotional state and their offspring's emotion regulation skills is intuitively clear. In early infancy, caregivers are the primary "emotion regulator" for child, and in this way parents model emotion regulation, providing children a framework for expressing, constraining, sustaining, or replacing various emotional states. One can imagine what happens to the process when parents do not (or can not) model successful emotional regulation. Do parents with difficulty in regulation of their *own* emotions model dysregulation to their offspring? Does ineffective parenting and/or attachment disrupt successful development of emotional regulation in infants? This pathway presupposes a behavioral transmission of risk, through parents' lack of appropriate modeling of effective regulation. Finally, there is also evidence of the linkage between temperament and later affectivity difficulties (Longan & Vasey, 2008). Much of the research in unipolar depression has focused on the link between temperament dimensions (specifically Negative Affectivity, Positive Affectivity and Constraint), and mood disorders in adolescence and adulthood (Clark & Beck, 1994; Lonigan & Vasey, 2008; Rettew & McKee, 2005; Watson, Clark, & Tellegen, 1988; Whittle, Allen, Lubman, & Yucel, 2006).

Operational definition of emotion regulation. To effectively measure emotion regulation, the construct must first be operationally defined. As mentioned, it is often defined by developmental psychologists in the context of temperament models (Goldsmith & Campos, 1986; Strelau; 1983, 1998; Zuckerman, Buss & Plomin, 1975, 1984). Emotion regulation consists of internal and external processes involved in

initiating, maintaining, and modulating the occurrence, intensity, and expression of emotions (Thompson, 1994). Similar definitions are offered by Eisenberg, Guthrie, Fabes, Reiser, Murphy, Holmgren, et al. (1997); Eisenberg and Morris (2002); and Eisenberg and Spinrad (2004). Most definitions stipulate that emotion regulation is a complex and fluid set of processes, which are integrated across physiological, cognitive, psychological and behavioral levels. These many systems must act in coordination in order for emotion regulation to be functional and adaptive for the individual (Silk, Vanderbilt-Adriance, Shaw, Forbes, Whalen, Ryan, & Dahl, *in press,* 2008).

One of the inherent difficulties in examining emotion regulation is the confusion caused by the myriad disparate conceptualizations of the construct. Cole, Martin, & Dennis (2004) suggest that it is imperative that when framing a new study of emotion regulation, one must first operationally define the construct to be measured. Apparently this has not always been the rule in emotion research. It is the theoretical conceptualization that drives the types of strategies for measuring emotion regulation – a subjective change in subjective state occurring dynamically in infants who can't talk or describe how they are feeling.

Components of Emotion Regulation

To fully operationalize the construct of emotion regulation, it must be broken down into its putative component parts based upon the conceptual model (Kochanska, 1997, 2000; Rothbart, 1981; Rothbart & Posner, 2003) used for this study. Just as Aristotle described temperament in terms of humours, so do those examining emotion

regulation often define it in terms of some common dimensions (Buss & Plomin, 1984; Goldsmith & Campus 1986; Kagan, 1989; Kagan & Moss, 1962; Kochanska, 1997, 2000; Strelau, 1998). The main theoretical model tested in this study consists of three critical and distinct dimensions: (a) Negative Affectivity, (b) Effortful Control, and (c) Surgency or Surgency/Extraversion. These will first be operationally defined.

Negative affectivity. First, negative emotionality or *negative affectivity* (NA) is a component feature of emotion dysregulation that is sometimes described as fussiness, difficult temperament, distressed behavior, low frustration tolerance, or moodiness. It has been often cited as a critical feature of suboptimal emotion regulation, and appears as a common theme when parents describe their infants as "difficult" (Zuckerman, Buss, & Plomin, 1984). Completing this theoretical linkage, there is evidence that Negative Affectivity is broad predictor of psychopathology, particularly associated with affective disorders such as depression and anxiety (Lonigan & Vassey, 2008; Posner & Rothbart, 2002; Whittle, Allen, Lubman, & Yucel, 2006). Whittle and colleagues (2006) found NA to be more strongly linked with disorders of global distress like Major Depressive Disorder and Generalized Anxiety Disorder than it is with disorders of more limited distress such as Social Phobia and Panic Disorder, or those characterized mainly by avoidance behaviors such as Specific Phobia and Agoraphobia. Clearly, there is much evidence implicating NA in emotion dysregulation and in the roots of psychopathology.

Effortful control. Descartes wrote eloquently on the subject of effortful control, when he advised that "The principal use of prudence or self-control is that it teaches us to be masters of our passions." (1649/1955, p. 427). The modern concept of *effortful control* is similar, involving focus, attentional control, and purposeful activity toward self-regulating emotional states. Effortful control (EC) includes attention directed at modifying a response to stimuli, often defined as "inhibiting a dominant response in order to perform a non-dominant response" (Posner & Rothbart, 2003). Thus, high levels of EC would be necessary for optimal emotion regulation. Whittle, Allen, Lubman, and Yucel (2006) included a construct very similar to EC in their model of emotion regulation, which they called "Constraint".

While EC (or Constraint) most closely resembles the broad construct of emotion regulation, it is but one feature of emotion regulation. EC may potentially moderate the risk of internalizing or externalizing problems that are often induced by high negative emotionality (Rothbart, 2003). Children's higher levels of effortful control may serve a protective function, leading to lower levels of child conduct difficulties despite parental/familial risk factors (Garstein & Fagot, 2003). Difficulties with EC have been implicated in other types of disorders, including Attention-Deficit Hyperactivity Disorder (ADHD; Wiersema & Roeyers, 2008). They found that children with low measured EC performed significantly differently from those with higher measured EC. Children with low measured EC demonstrated a higher proportion of ADHD symptoms, made more impulsive errors during testing (on Go/No-Go tasks), and showed smaller No-Go P3 amplitudes of event-related potentials (ERPs) related to the executive attention network, a

phenotypic neurological marker of poor EC. Behavioral evidence of EC appears in typically developing children at around age 18 – 24 months (Backen Jones, personal communication, 2009). Further, higher levels of depressive symptoms, coercion, and cognitive guidance from parents, along with *lower* levels of child EC (less purposeful regulation), were associated with higher levels of child externalizing behaviors (Garstein & Fagot, 2003). Parental/family factors and child effortful control should be considered in understanding the development of behavior problems in early childhood (Garstein & Fagot, 2003).

Surgency. Finally, *surgency* or *surgency/extraversion* (S) has been defined as a measure of approach behavior or positive affectivity. It involves the behavioral goal of seeking resources, by organizing responses to obtain potential rewards (Rothbart & Bates, 1997; Rothbart & Sheese, 2007). Optimal self-regulation would entail a strong presence of surgency. Some models of emotion regulation do not include the dimension of surgency, instead focusing only on the inhibition of emotion (withdrawal, avoidance, or escape behaviors) as a measure of psychological health.

However, the theoretical model in this study defines healthy functioning as more than just the absence of negative emotion and/or inhibition of emotional expressions; it necessarily involves the essence of moving toward others, of approach behaviors, and positive expressions of emotions. Surgency could be described as a positive engagement in ones' life and activity - a moving forward within ones' life or the healthy involvement and activity that is functional for the individual, particularly in seeking rewards. Whittle,

Allen, Lubman, & Yucel (2006) and others use a similar construct in their theoretical framework which they label Positive Affectivity (PA). Buss & Plomin (1975, 1984) also defined a positive affectivity category of "sociability" by describing behaviors such as a desire to be with others, makes friends easily, prefers to be with others rather than be alone, and is not shy. Goldsmith and Campos' (1986) dimension of temperament called "Pleasure" operationally fits the construct of Positive Affectivity as well. Surgency is a behaviorally-defined construct that appears earlier than EC. Rothbart and Derryberry (1981) reported that Positive Affectivity (or surgency) was generally observable in infants at around two months of age, particularly in terms of their approach to cues of reward or novelty. This is congruent with the broad category of surgency as defined here.

The construct of surgency also figures prominently in the pioneering cognitive-behavior work in the 1970s by Lewinsohn and colleagues (Lewinsohn, 1974). Lewinsohn asserted that the etiology of depression involves a lack of response contingent positive reinforcement. This conceptual scenario, then, highlights the importance of high levels of surgency in seeking out situations and avenues for receiving contingent positive reinforcement. Emotional experience and expression are important features of our humanity and their absence can create significant mental health and interpersonal problems. Internalizing disorders such as anxiety, phobias, and even depression, typically involve over-controlled emotion and behavior to the point that the individual is not a fully active agent in his/her own life. Individuals with internalizing disorders tend to withdraw or escape from experiences which they believe may be fear or anxiety-inducing, to avoid feelings or experiences that could trigger additional pain. Of course, withdrawal only

reduces the opportunities for positive reinforcement from the world, invoking more isolation, gloom, and painful emotional experiences, which reinforces further withdrawal. And so the spiral goes. A very low level of surgency in this conceptual model would indicate suboptimal emotion regulation, in which the infant was not positively engaging with his environment, thereby limiting access to positive reinforcement. For this study, then, it is predicted that low levels of surgency would be related to later higher levels of internalizing behaviors and risk for depression.

Emotion Regulation Patterns

Optimal emotion regulation. Ben Jonson in his comic satire *Cynthia's Revels* (1600) ably describes the pattern of optimal emotion regulation:

> A creature of a most perfect and divine temper; One, in whom the Humours and Elements are peaceably met, without an emulation of Precedencie: he is neither too fantastickly Melancholy; too slowly Phlegmatick, too lightly Sanguine, or too rashly Cholerick, but all in all, so compos'd and order'd; as it is cleare, Nature was about some full worke, she did more than make a man when she made him.

If optimal temperaments are well balanced, then emotion regulation must play a part in achieving that balance. Optimal self-regulation of emotional state occurs when the one's response is functional in the given context. Thus, optimal regulation would involve a predictable combination of these three constructs (Posner & Rothbart, 2000) which involves (a) low levels of negative affectivity, (b) high levels of effortful control, and (c) high levels of surgency. The optimally self-regulated individual is not too negative in

mood, has a strong sense of impulse control and capacity for delayed gratification, and is skilled in engaging in social and meaningful ways with his or her environment thereby gaining access to contingent positive reinforcement. Thus, in terms of functional relations among constructs Negative Affectivity would be negatively correlated with both Effortful Control and Surgency, while Effortful Control and Surgency would be positively correlated with one another.

Suboptimal emotion regulation. A suboptimal regulation pattern would show high levels of Negative Affectivity (NA), low levels of Effortful Control (EC) and Surgency (S). This individual would demonstrate frequent negative affect, coupled with poor impulse regulation, control of attention and patience, and would demonstrate limited approach behaviors, thus self-limiting opportunities for contingent positive reinforcement from the environment. In this main theory, three latent constructs (factors) of NA, EC, and S, are presumed related, and presumably may be measured indirectly through the identified indicators. These three constructs putatively comprising emotion regulation will result in certain specific patterns, which can be characterized as either optimal emotion regulation or suboptimal emotion regulation as described.

The pattern of suboptimal emotion regulation is important to measure, because the later implications of this pattern are problematic. The work of Verstraeten, Vasey, Raes, & Bijttebier (2008) with adolescents provided empirical support for the existence of significant associations between temperament and later depression in adolescence. Verstraeten et al. (2008) found that individuals with higher levels of NA, lower levels of

Positive Affectivity (similar to surgency) and lower levels of EC were found associated with higher levels of depressive symptoms. Their results showed that this association between NA and PA was significantly moderated by the participant's level of Effortful Control. That is, this pattern of high negative affect and low positive affect was significantly associated with depressive symptoms *only* when the effortful control dimension was low. They also found higher levels of NA were associated with higher levels of ruminative response style, which was related to more depressive symptoms, but, again, this held true only in individuals with low EC. Earlier valid measurement of emotion regulation may allow for intervention and perhaps prevention of these undesired outcomes in adolescents.

Complementing prior work in affective neuropsychology that attempts to link neural circuitry with temperamental phenotype (Davidson & Irwin, 1999; Davidson & Henriques, 2000), the work of Whittle, Allen, Lubman, & Yucel (2006), draws parallels between specific temperamental constructs (such as NA and PA), and specific neural circuitry between functional areas of the brain. Substantial empirical evidence (Davidson, Pizzagalli, Nitschke, & Putnam, 2002; Goldapple, Segal, Garson, Lau, Bieling, Kennedy, & Mayberg, 2004; Henriques & Davidson, 1990) supports the relation between NA and activation of subcortical structures such as the amygdala and hippocampus, and some prefrontal structures such as the anterior cingulate cortex (ACC) and the dorsolateral prefrontal cortex (DLPFC). The linking of theoretical temperamental constructs and functional neural circuits and processes underscores the need (and heuristic value) in focusing on temperament (specifically ER) in exploration of the mechanism for the

transmission of mood disorders (Whittle et al., 2006). While the humours may have their roots in mythology, constructs of NA, S, and EC have their roots firmly in brain structures. Study of temperament is one way of conceptually mapping and linking modes of observed behavior to physiological phenomena.

Linking parental depression and offspring risk. An important end goal in examining emotion regulation as a depression transmission pathway is to aid prevention through early intervention. Many feel that mental disorders have at their core the dysregulation of affect, even though the mechanism relating emotion regulation to disorders is unclear (Davidson, 2000; Lonigan & Vassey, 2008). The diagram in Figure 1 depicts the presence of parental depression (both genetically and behaviorally) as a predictor of suboptimal emotion regulation in offspring. The diagram also depicts the putative relation between offspring emotion regulation to internalizing behavior problems. It is hypothesized that the first-order latent constructs Negative Affectivity (NA), Effortful Control (EC), and Surgency (S) each contribute in separate and significant ways to offspring emotion regulation outcomes. The double arrows from parent depression to offspring ER represent the two interactive paths of transmission from parent to child: genetic predisposition and parenting behaviors, that is, (a) family history of depression directly affects offspring internalizing behavior problems, and (b) depressive symptoms concurrent with parenting predicts offspring's suboptimal emotion regulation.

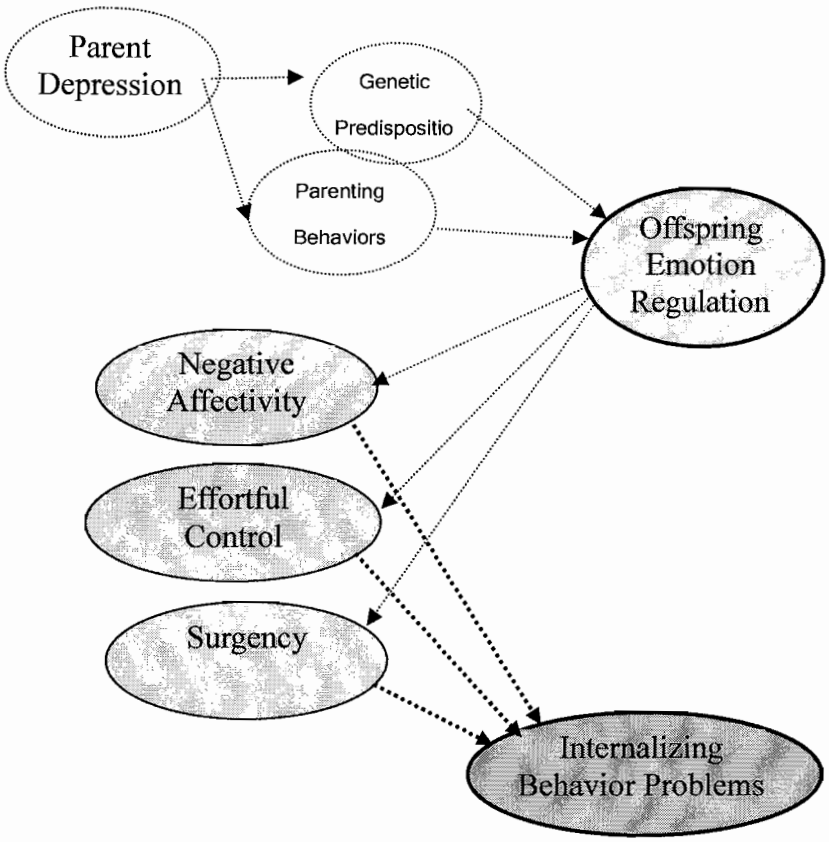

Figure 1. *Diagram of the Theoretical Model of Transmission of Risk from Infant Suboptimal Development of Emotion Regulation*

Development of emotion regulation. Gartstein and Bateman (2008) sought to explore the complex relation between infant temperament and maternal depression symptoms and the outcome of toddler depressive symptoms. They found that toddler depression symptoms were attenuated in infants with lower measured NA and mothers' reporting fewer depression symptoms.

However, when infants had higher measured NA, they showed increased toddler depression symptoms later on, regardless of maternal depression symptom status.

Many have asserted that difficult temperament automatically implies increased risk, through diverse and complex mechanisms, not the least of which is disrupted parent-child interactions (Bates, Freeland, & Lounsbury, 1979; Thomas, Chess, & Birch, 1968). Pettit and Bates (1984) on the other hand, indicated that the "difficult infant" dimension (perhaps a synonym for suboptimal emotion regulation) was relatively independent of mothers' and infant behavior. Clearly, there is no simple linear path from difficult infant to depression risk.

Figure 1 illustrates the hypothesis that depression risk may be passed down in families through development of suboptimal emotion regulation. Others have outlined the evidence that suboptimal emotion regulatory processes are implicated in different types of psychopathology. For example, in examining the diagnostic criteria from the diagnostic and statistical manual of mental disorders (DSM –IV-R) one finds that one of the criteria for diagnosing Borderline Personality Disorder is "difficulty controlling anger"; for Posttraumatic Stress Disorder (PTSD) "efforts to avoid feelings" is a criterion, and Generalized Anxiety Disorder is often characterized by clients' "difficulty controlling worry" (Kring & Werner, 2004).

De Pauw, Mervielde, & Van Leeuwen (2009) recently examined the relations between toddlers' problem behaviors and three different measures of toddler temperament. They compared the relations between children's behavioral maladaption and traits measured by three temperament models (Buss & Plomin, 1984; Rothbart, 2003;

Thomas, Chess, & Birch, 1968), and the Five-Factor personality model (Cicchetti, Ackerman, & Izard, 1995; Goldberg, 1990; Srivastava, 2009). Through a joint principal components analysis, in which they combined items from both personality and temperament scales, they found a six-factor model that included sociability, activity, conscientiousness, disagreeableness, emotionality, and sensitivity. This model did a better job of differentiating among the Child Behavior Checklist problem behavior scales (41% - 49% of problem behavior variance explained by the model) than did a single temperament or personality scale alone (23% - 37% of problem behavior variance explained).

In the same manner, then, this study seeks to first define the putative components of emotion regulation and then examine their relation to outcomes of interest. This study, however, will base the conceptual model on three theoretical dimensions of emotion regulation commonly supported in the field (Posner & Rothbar, 2003). Further analysis could include evaluating whether or not parental history or family density of depression predicts emotion regulation in offspring, and ultimately, which factors of parental depression history are best predictors of offspring emotion regulation (e.g., family density, current symptoms, depressed parent gender, etc.).

CHAPTER III

METHOD

Procedures

This study made use of an extant data set generated from the Infant Development Study (IDS), which began in 1996 and evaluated offspring of previously depressed youth. Long term goals of IDS were to inform the design of preventive interventions for children of parents with affective disorders in order to prevent potential deleterious effects of parental depression on offspring well being. Another goal was to examine the potential mechanisms for familial transmission of affective disorders and/or symptoms, which the study described as potentially genetic, parenting-related, or context-related.

In the IDS study, multiple sources of information were used, including mothers, fathers, trained observer/raters, and clinical psychologists or their highly-trained assistants. About one third of the probands were diagnosed at some point with depression, another third were diagnosed with other mental disorders, and another third had no mental disorders. The extant longitudinal data on the probands was used along with the newer data collected for the offspring.

In addition, multiple methods of data collection were used which included: detailed mailer questionnaires, behavior rating scales, developmental scales administered by trained raters, and structured parent interviews from clinical psychologists.

Participants

Participants were those in the Infant Development Study which included a sample drawn from the 770 females and 630 males from wave 3 of the Oregon Adolescent Depression Project (OADS), a longitudinal prospective study of adolescent depression. Originally, the OADS project consisted of 816 individuals (59% women; 41% men; 89% White) who started participating in the research project during their mid-adolescence and continued through their early 30s. For additional detail about this sample and study see Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993; Lewinsohn, Rohde, Seeley, Klein, & Gotlib, 2003; and Rohde, Lewinsohn, Seeley, Klein, Andrews, & Small, 2007. Teenaged participants in the sample were evaluated at approximately age 16 years, 17 years, as they approached their 24th birthday (child-bearing years), and again as they approached their 30th birthday. The Infant Development Study focused on those participants who became parents during these last assessment periods in early adulthood.

Of these, approximately 167 probands from the OADP were used in the IDS study. Because IDS examined participants throughout adolescence and again during their peak childbearing years (24 years and 30 years), extensive information on both probands and their offspring was obtained. Offspring in the Infant Development Study were evaluated from ages 3 months to 48 months and beyond. The goals in selection of measurement occasions for this study were to use the waves which provided the most relevant responses from mothers from among those available for examination. One of the issues inherent in examining extant data is that one works with the data one has.

The initial intent was to evaluate toddler problem behavior at 36 months, however the Child Behavior Check List was not administered to mothers at the 36 month measurement occasion, but rather at 48 months.

Also of note is that the probands were not exclusively those diagnosed with Major Depressive Disorder. In fact, about one third of those were diagnosed with MDD, another third were diagnosed with other psychiatric disorders, and one third had no such psychiatric diagnoses. Thus, while not a representative community sample, some participants did demonstrate a variety of conditions other than depression.

Measurement of Emotion Regulation: Operationalization

"Measurement is the Achilles' heel of socio-behavioral research" assert Pedhazur & Schmelkin, (1991, p.2). This certainly seems to be the case when trying to measure a phenomenon as complex and dynamic as emotion regulation in a toddler. The measured variables of interest in this study included: infant development of emotion regulation ratings from mothers and mothers' ratings of toddler behavioral problems (internalizing, externalizing, and global problem behaviors). Mothers' ratings of offspring temperament and behaviors were obtained through mailer questionnaires.

Instruments

From the IDS data set, items and subscale scores were taken from the Toddler Behavior Assessment Questionnaire (TBAQ; Goldsmith, 1996), the Infant Characteristics Questionnaire (ICQ; Bates, Freeland, & Lounsbury, 1979), and the Child Behavior

Checklist (CBCL; Achenbach, 1991). These instruments were delivered as part of the IDS mailer questionnaire, completed by mother and father, and are described more fully as follows.

Infant Characteristics Questionnaire (Bates, J., Freeland, C., & Lounsbury, M. 1979). The ICQ is a 28-item behaviorally anchored measure of infant temperamental difficulty completed by the parent. Bates et al., used squared multiple correlations as initial communality estimates and varimax, orthogonal rotation in their factor analysis, and reported that the ICQ yielded four coherent subscales: fussiness/difficulty, unadaptability, inactivity, and unpredictability. The ICQ is essentially a brief, screening device for infant difficultness, e.g., a way to measure and describe the fussy, hard-to-sooth, emotionally labile infant that mothers call "difficult". The instrument provides assessments of fussy and difficult infant behaviors that were moderately stable over a 10 month period (r's > .48; Moran & Pederson, 1998).

Definitions of the four factors are as follow: (a) fussiness/difficulty (fussy, hard to soothe), (b) unadaptability (unadaptable regarding infants' initial and eventual reactions to new events, people, and things), (c) inactivity (the opposite of active, sociable, and fun), and (d) unpredictability (difficult to predict infant's needs, e.g., hunger, wet). The fussy/difficult subscale shows moderate (Kirk, 1995; Tabachnick & Fidell, 2001) internal consistency ($r > .60$), adequate test-retest stability over a one-month period, good inter-parent agreement ($r = .61$) and concurrent validity with observed fussiness, demandingness, and the degree to which the child is unable to entertain him/herself

(Bates, Freeland, & Lounsbury, 1979). The ICQ was found to correlate with other measures of temperamental difficulty and with in-home ratings made by independent observers (Bates et al., 1979; Whiffen & Gotlib, 1989). It is appropriate for use with children up to 24 months of age. Relevant to this study, Teti and Gelfand (1991) also reported significant correlations between the fussy/difficult scale and concurrent maternal depression.

Toddler Behavior Assessment Questionnaire (TBAQ): Goldsmith, 1996. The TBAQ is a 110 item, parent report instrument containing 11 dimensions, which is designed to examine temperament-related behavior in 16–36 month old children. The factors measured are described as follows: (a) Activity Level (including limb, trunk, or loco-motor movement during a variety of situations, free play, confinement, or quiet activities, (b) Pleasure (smiling, laughter, and other hedonically positive vocalizations or playful activity in a variety of non-threatening or familiar situations), (c) Social Fearfulness (inhibition, distress, withdrawal or signs of shyness in novel or uncertainty-provoking social situations), (d) Anger Proneness (crying, protesting, hitting, pouting, or other signs of anger in conflict situations with a child or the caregiver), and (e) Interest/Persistence (duration of task engagement in ongoing solitary play or other activities). Many of the subscales are similar to those in the Infant Behavior Questionnaire (Rothbart, 1981). According to the manual, internal consistency reliability estimates for the subscales typically exceeded .80 (Goldsmith, 1996).

Child Behavior Check List (CBCL; Achenbach, 1991). Mothers completed this 99 item scale by rating their child's behavior in the last 2 months. Reliability and validity of this instrument are well established, with test-retest reliability in the range of .71 to .93, inter-parent reliability between .63 at age 2 and .60 at age three (Achenbach, 1991). Further, for the internalizing subscale inter-parent agreement ranged from .57 to .71, .70 to .86 for the externalizing behavior subscale and from .69 to .82 for Total Problem Behaviors subscale across ages two and three years. The CBCL has been widely used in developmental and psychopathological research, with strong evidence of high convergent and discriminant validity (Achenbach, 1991; Achenbach & Rescorla, 2001). Scores can be used to describe two empirically derived broadband syndromes of internalizing and externalizing problems, as well as total behavior problems (a combination of the two scales). The scales include the following subscales: (a) Internalizing Scale**:** Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn and (b) Externalizing Scale: Attention Problems, Aggressive Behavior, Destructive Behavior. A detailed list of the items from the Child Behavior Checklist corresponding to the subscales is presented in Appendix C.

Expected Outcomes

In fitting the measurement model, it was expected that optimal emotion regulation and suboptimal emotion regulation would show specific and divergent patterns of association with the problem behavior outcomes. Table 1 depicts how optimal emotion regulation would be characterized by negative associations with NA and positive

associations with both S and EC, that is, optimal emotion regulation in toddlers is defined by lower behavior ratings of negative emotionality and higher ratings of positive emotionality and effortful (regulatory) control.

Table 1.

Expected Patterns of Latent Constructs in Emotion Regulation

Pattern type	Negative affectivity	Surgency	Effortful control
Suboptimal emotion regulation externalizing	Positive	Negative	Negative
Suboptimal emotion regulation internalizing	Positive	Negative	Positive
Optimal emotion regulation	Negative	Positive	Positive

The expected patterns of optimal and suboptimal emotion regulation based upon the recommendation of the expert panel, and the current review of literature on emotion regulation are depicted in Table 1. It was expected that patterns for internalizing and externalizing problem behavior would differ specifically in the association with EC, in

that for externalizing behavior, EC would be negatively correlated (characterized by under-control), while for internalizing behavior, EC would be positively correlated (characterized by over-control). To examine these hypotheses, and as a method of concurrent and predictive validation, mothers' ratings of children's problem behaviors on the Child Behavior Checklist (CBCL; Achenbach, 1991) at 24 months (Wave 4) and 48 months (Wave 6) were regressed upon the best fitting measurement model, thus creating structural models of emotion regulation.

The initial research question involved examination of infant emotion regulation at ages 12 months and 36 months. However, expert panel member Backen Jones recommended that the construct of effortful control would not likely be observable at 12 months of age, and that analysis at 24 months would be more likely to provide a better estimation of emotion regulation from mother report. Further, it was initially planned to evaluate toddler problem behavior at 36 months, but upon closer examination it was discovered that mothers did not complete the CBCL portion of the survey at toddler age 36 months, but rather, did so later, at 48 months. In all, the measurement occasions initially chosen for analysis were both shifted to one year later, based upon expert recommendation. This did not substantially alter the research focus, but rather, may have made it more likely to yield interpretable results.

Indicant Selection

Subscales from the Infant Characteristics Questionnaire (ICQ; Bates, Freeland, & Lounsbury, 1979), and the Toddler Behavior Assessment Questionnaire (TBAQ;

Goldsmith, 1986) were used for selection by an expert panel of senior research scientists in the field of emotion regulation and affective disorders, which included Lisa Sheeber, PhD; Laura Backen Jones, PhD; and John Seeley, PhD. Researchers on the expert panel were asked to select those subscales from the ICQ and TBAQ, which best represented the construct of emotion regulation overall, and the specific domains of NA, EC, and S. Their selections were done independently and those chosen by all three experts were selected as indicants for this measurement model. A detailed list of recommendations from the expert panel members is presented in Appendix A. Table 2 presents the selected subscale indicants believed to be associated with the three latent constructs (NA, S, and EC) and lists the infant behavior instrument that the subscale is derived from.

Table 2.

Latent Constructs and Their Recommended Indicants

Construct	Indicant	Measure
Negative affect	Fussiness/Difficulty	ICQ
	Anger Proneness	TBAQ
	Social fearfulness	TBAQ
Effortful Control	Unadaptability (reverse)	ICQ
	Unpredictability (reverse)	ICQ
	Interest/persistence	TBAQ
Surgency	Activity level	TBAQ
	Pleasure	TBAQ
	Unsociable (reverse)	ICQ

Note. ICQ = Infant Characteristics Questionnaire; TBAQ = Toddler Behavior Assessment Questionnaire.

Measurement Models

A series of confirmatory factor analysis (CFA) measurement models were conducted based on theory and previous research about infant emotion regulation. Because the models tested were based upon theory *a priori* tests of comparative model fit were conducted with a CFA method rather than exploratory factor analysis (EFA). There

are many substantive advantages to using the CFA method when confirming mature theory (Bagozzi, 1980; Hu & Bentler, 1982; Marsh, 1987), including the ability to construct, define, and test an *a priori* model; specify or estimate model parameters; and test and compare the relative goodness of fit of various competing structural models. The general CFA model is specified in Structural Equation Modeling notation as follows:

$$\mathbf{Xi} = \Lambda\mathbf{X}\ \Xi\mathbf{k} + \Theta\delta$$

where:

Xi = a column vector of observed variables;
Ξk = ksi, a column vector of latent variables;
ΛX = lambda, an i X k matrix of structural coefficients defining the relations between the manifest (X) and latent (Ξ) variables;
Θδ = theta-delta, an i X i variance/covariance matrix of relations among the residual or error terms of X.

The equation indicates that each of the manifest variables (anger proneness, distress to limitations, soothability, etc.) may be expressed by a structural equation that interconnects lambda, ksi and theta-delta. Thus, the model for the relations between factors can be represented by a set of structural equations which include these above vectors and matrices. Additionally, these described equations will be presented in matrix form, along with another necessary matrix, the phi (Φ) matrix which is also necessary for model estimation (See Table 6). This is a variance-covariance matrix that specifies relations among all of the latent variables (Ξ) in the model as described previously. Three alternate conceptualizations of the structure of the emotion regulation instrument were compared using maximum likelihood CFA procedures with M-Plus (Muthen & Muthen, 1998-2007). In order to establish a metric for the latent variables, one path

coefficient for each latent variable was set to 1.0. Initially, the latent variables were allowed to correlate freely, variances of factor uniquenesses were estimated, and covariances of uniquenesses were fixed to zero (uncorrelated uniquenesses).

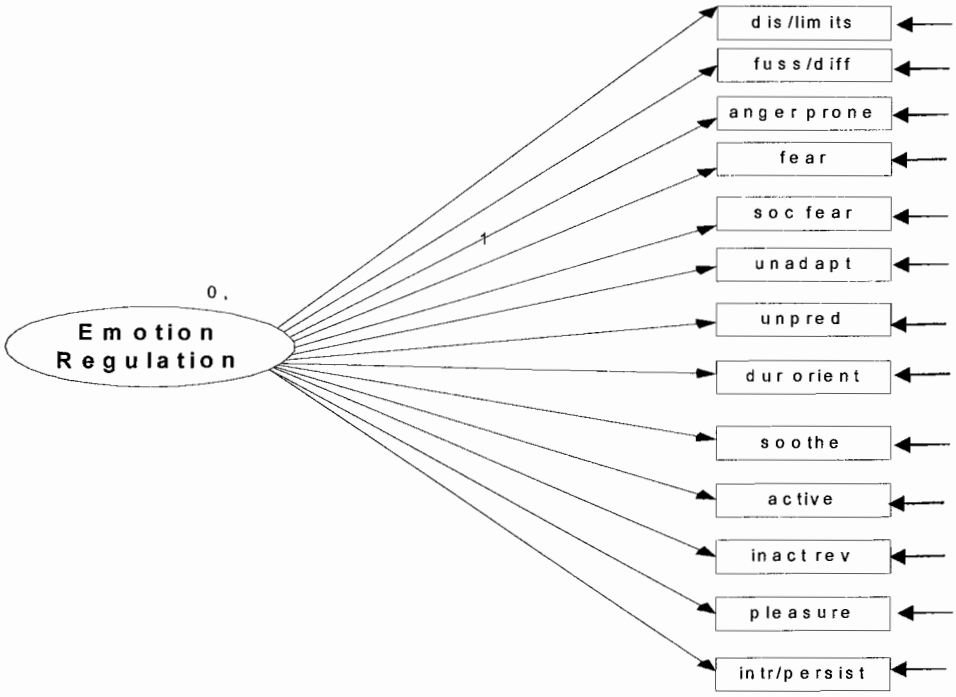

Figure 2. One Factor Model of Emotion Regulation

Model 1 (ER1) is a single emotion regulation factor which accounts for the variability in all the eleven indicants. Model residuals are designated by the arrows directed toward each indicant in the model. This model depicts a global construct of emotion regulation measured by the indicants. It represents essentially one factor of interest made up of the domains of NA, EC, and S which are defined as non-interpretable

as distinct constructs. There is empirical and theoretical support for this unidimensional model. The initial factorial analysis of the IBQ, for example, by Rothbart (1986) resulted in demonstration of a unidimentional construct of emotion regulation defined along a continuum between negative and positive affectivity. Although two factors emerged (Negative Affectivity and Positive Affectivity), results from the scree test were such that authors interpreted results to indicate that there was only one factor not present due to random variability. Further, the IBQ (Bates, Freeland, & Lounsbury, 1979) taps into a single dimension of infant temperament, namely, difficult temperament, which includes constructs of fussiness, unadaptability, unpredictability, and inactivity. The first three constructs could well overlap significantly, and fall under one general construct of negative affectivity. An infant who is unadaptable and unpredictable may also be perceived by parents as fussy.

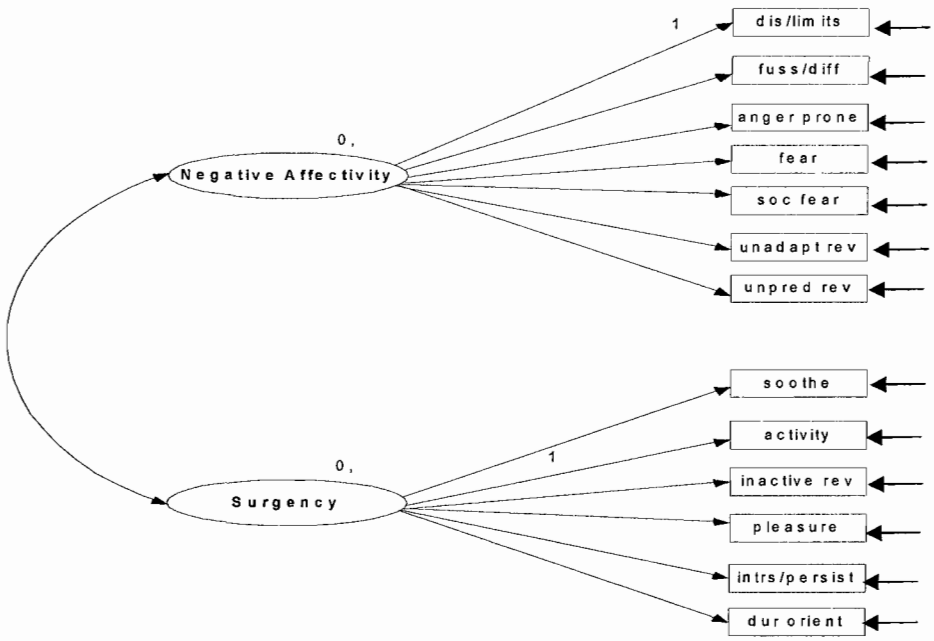

Figure 3. Two Factor Model of Emotion Regulation

Model 2 (ER2) is a two-factor model, similar to the initial factor analysis model of the IBQ (Rothbart, 1981) a temperament scale used in the IDS study with the probands up until age 24 months. Factor analytic work with the IBQ has generally yielded dimensions related to Positive and Negative Affectivity (Goldsmith & Campos, 1990; Kochanska, 1997; Rothbart, 1981). The latent construct surgency (S) is theoretically and operationally very similar to the construct of positive affectivity, as both include approach behaviors, activity, and positive affect. Many theories of development, mood disorder, and problem behavior suggest the presence of a two-factor model.

Achenbach (1988), through empirical research, found problem behaviors generally fell into two subtypes: internalizing (or over-controlled) and externalizing (or under-controlled) behaviors. Similarly, other conceptual models parallel this two-factor solution, such as the Behavioral Inhibition Scales and Behavioral Activation Scales framework suggested by Carver and White (1994) and Gray (1994), and the over-controlled and under-controlled constructs described by Wolfson, Fields, and Rose (1987). Congruent with this framework our two factor model (ER2) was created, based upon the latent factors of NA and S, and including the same eleven indicants from the one factor model. These eleven indicants were bifurcated into the latent constructs of NA and S based upon face validity, and approved for evaluation by our expert Dr. Seeley.

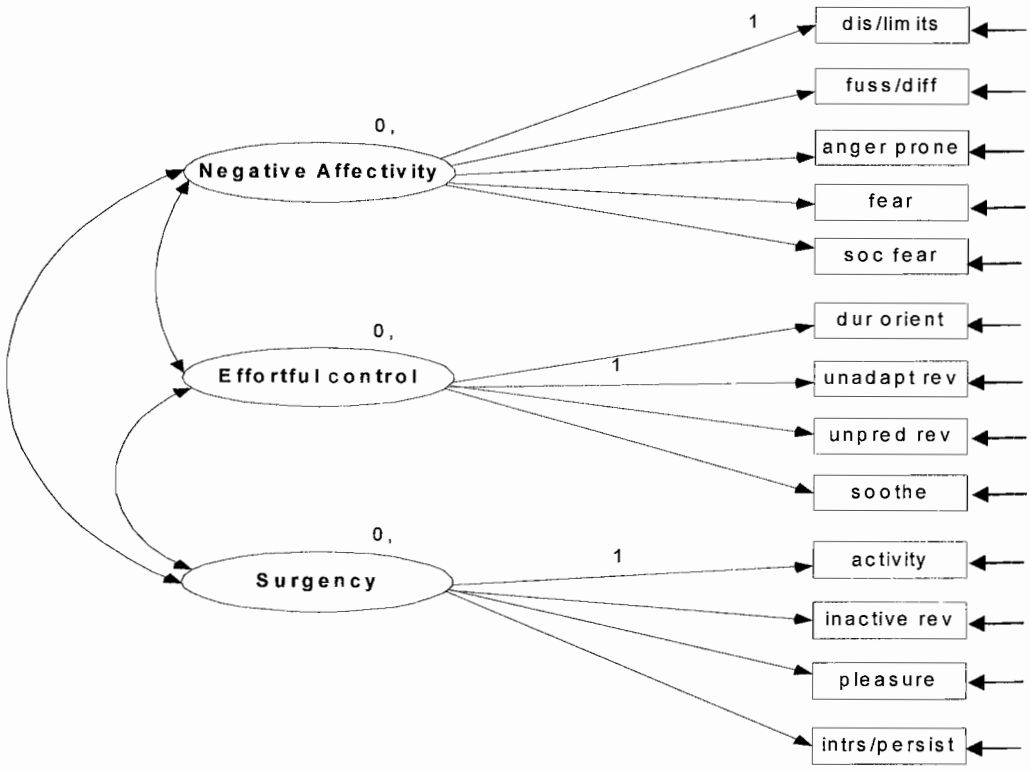

Figure 4. Three Factor Model of Emotion Regulation with Subscales

Model 3 (ER3) is a three factor solution with the three related but distinct constructs of Negative Affectivity, Surgency, and Effortful Control. This added dimension of Effortful Control brings the construct of intentionality and attention into the mix. Goldsmith (1996) also included this attentional dimension in his TBAQ as a regulatory feature. The expectation of this model is that the indicants would correlate with their designated latent factors with little overlap, and that the correlations between measured factors may be but not high (Kline, 2005; Kirk, 1995). This theoretical model is based partly upon the work of Ahadi, Rothbart, and Ye, 1993; Rothbart and Derryberry, 1981; Kochanska, DeVet, Goldman, Murray, and Putnam, 1994; Posner and Rothbart,

2003; and Whittle, Allen, Lubman, & Yucel, 2006. There is substantial theoretical and empirical support for a three factor model of emotion regulation. For example, Clark and Watson's tripartite model of emotion (Clark & Watson, 1991) which was prominent in the depression and anxiety literature of the 1990s, includes this general NA factor which theoretically influences both anxiety and depression. The other two components of the tripartite model were a physiological hyperarousal factor (PH), and a positive affect dimension (PA). These factors appear congruent with the factors of EC and S, respectively. There are distinct conceptual differences between the constructs of PH and EC, however, in that, while they both involve a type of activation, PH does not share the purposeful and attentional elements found in EC. Chorpita and colleagues also found empirical support for this three part model (Chorpita, 2002).

The theoretical model by Whittle et al. (2006), is congruent with and based upon the adult personality models of Clark and Watson (1999), Cloninger, (1986), Eysenck (1990), and Watson, Clark, and Tellegen (1984). Whittle et al. assert that considerable evidence demonstrates significant and stable relations between the latent constructs of Negative Affectivity, Positive Affectivity (here, as S or surgency) and Constraint (here, as Effortful Control or EC) and several mental disorders. Studies providing evidence of the relations between stable patterns of NA, S, and EC include reviews by Clark et al. (1994), Rettew and McKee (2005), and Watson et al. (2005).

To summarize, a series of nested models (utilizing the same eleven indicants) including a one factor, two factor, and three factor model of emotion regulation were tested for model fit with the data from the Infant Development Study in an attempt to create a well fit measurement model of emotion regulation.

Model requirements. In order to specify a CFA model, two requirements must be met: (1) the number of parameters must be equal to or less than the number of observations, and (2) every latent variable (including measurement errors and factors) has to have a scale (Kline, 2005). Further, the CFA model must meet requirements for identification. In this case, the construct of emotion regulation is believed to be made up of at least three dimensions, including Negative Affectivity, Surgency, and Effortful Control. This CFA model had three first-order factors (NA, EC, and S) and at least three indicators per factor and the model was believed to be over-identified. Through an iterative process (a series of model fittings), the number of factors were determined that best fit the data based upon models tests.

Model description: structural models

Following generation of an acceptable model, further analyses including structural models were run in order to secure concurrent and predictive validity evidence for the measurement model. Further, the models were intentionally designed to examine the unique effects of each of the latent constructs in predicting behavioral outcomes.

Since the focus of the study was on determination of the dimensions which comprise the broader construct of emotion regulation demonstration of each first order constructs unique contribution was highlighted by these structural models.

In order to provide evidence of concurrent validity, it was predicted that the model at 24 months would significantly explain outcomes of toddler internalizing/externalizing behavior at the same measurement occasion. Further, by regressing the outcome internalizing/externalizing behavior at 48 months on the latent constructs, it was hypothesized that the model would also show predictive validity, that is, the emotion regulation at 24 months would demonstrate significant power in explaining variability in internalizing/externalizing behavior at 48 months. The unique contribution of each first order construct was examined.

Finally, a more formal analysis of the relation between the model and the outcomes of internalizing/externalizing behaviors at 48 months was performed by regressing the outcome behaviors at 48 months on the latent ER constructs and on the behavioral outcome results at 24 months. This procedure demonstrated the strength of the model in predicting residual change in the problem behaviors, while holding the level of problem behaviors measured at 24 months constant.

The value of establishing evidence of predictive validity should not be underestimated in term of utility in informing intervention. Well fit models that show good predictive ability may inform focused interventions addressing those behaviors critical to optimal development. For example, parents with a history of MDD, but who successfully teach/coach infant development of emotion regulation can perhaps reduce

the effects of family history of MDD on their offspring. Here is the potential for intervention – teaching parents more contingent responding, better attunement and synchrony, and other specific methods of supporting their offspring's development of healthy self-regulatory skills. It is hoped that development of strong ER skills may buffer the infant against the risk of internalizing/externalizing behaviors in later childhood related to the effects of parental depression. Not only could a suitable measurement model identify offspring who are at greatest risk for later behavioral problems, but it could also measure behavioral improvement as an intervention outcome.

Research Questions

1. *How do we best measure emotion regulation in infants/toddlers ages 12 months to 36 months?*

2. What combination of indicants and latent constructs best explains individual differences in ER during this early developmental period (24 months to 48 months)?

These first two questions were examined by testing the definition of the construct of emotion regulation as a product of the interaction of three commonly-mentioned dimensions of emotion regulation – Negative Affectivity, Effortful Control, and Surgency (Kochanska, 1997, 2000; and Rothbart, 1981). First, different measurements of the same purported construct should correlate significantly with one another. Second, the predicted patterns within these dimensions should become apparent, such that those infants described by mother report as demonstrating suboptimal emotion regulation should show the predicted pattern of high and low scores on the three putative dimensions.

Finally, other measures of known validity and reliability can be used to provide evidence of construct validity for our instrument, such as the measures' correlation with the Child Behavior Check List (Achenbach, 1988) as reported by mothers as well.

3. *Is emotion regulation in offspring best explained by a one factor, two factor, or three factor model, i.e., NA, EC, and S or ER?*

4. *Does suboptimal emotion regulation at 24 months predict internalizing behaviors at 48 months?*

Analysis Strategy and Model Selection

Preliminary analysis. Descriptive statistics for all variables were examined, including mean, standard deviation, and frequency distributions, to examine the tenability of assumptions required for the proposed statistical analyses. The examination included calculating correlations within a correlation matrix format using the indicator scores and means and standard deviations among indicants. A summary table of the correlation matrix of subscale correlations is presented in Appendix B. The following discussion outlines the steps recommended by Kline (2005) for checking the tenability of assumptions for analysis. The two basic conditions for a CFA are that: (a) the number of free parameters is less than or equal to the number of observations, and (b) every latent variable, including measurement errors and factors both, must have some type of scale. This scaling was achieved by fixing the unstandardized residual coefficient (that is, the direct effect of the measurement error on its indicator variable) to 1.0. Factors were

scaled in a similar manner, fixing the direct effect of one of the indicants to 1.0 as well (as the reference variable). After the first two conditions had been met, it was necessary to ensure that the model included at least two indicators per factor (the "two-indicator rule" by Bollen, 1989). In the basic model, there were at minimum three measured indicators per latent factor.

One of the key distributional requirements for CFA is the assumption of normality of the distribution; including univariate, bivariate and multivariate normality. CFA using maximum likelihood estimation with non-normal distributions tends to provide standard errors which are too low, thus artificially inflating the Type I error rate. The assumptions of univariate and bivariate normality were tested and found tenable in the distributions of observed variables. This included examination of extreme skew and kurtosis, and influential case outliers. Evaluation of a correlation matrix of indicator variable scores was used to assess bivariate normality. Values of skew and kurtosis were found to be within recommended normal limits (+/-1.0 or even +/- 1.5 or 2.0 according to some; Schumacker & Lomax, 2004). Kline (2005) recommends a more liberal guideline, suggesting that standardizied skew index values between -3.0 and +3.0 may be considered within normal limits, while standardized kurtosis index of -10.0 to +10.0 may be considered roughly normal.

Kline also acknowledges that since multivariate normality requires that all univariate distributions are normal, the joint distributions of any pair are normal, and that all bivariate plots are homoscedasctistic and linear, assessing all forms of normality for several variables may present a great challenge. Kline goes on to suggest that in most

cases a careful evaluation of univariate normality may provide enough information to assume that multivariate normality is tenable, especially when distributions are scanned for extreme outliers. In this case, a popular statistical measure used to test multivariate normality, Mardia's statistic, was not available for use in SPSS, and so multivariate normality assumptions were based upon the demonstration of univariate and bivariate normality within the distributions.

Dealing with attrition and missing cases. The most parsimonious way to deal with missing data is not to have them at all. Fortunately, within the IDS study, great efforts were made to minimize attrition and missing cases, so that of 167 participants, 151 had usable data for most variables (over 90% of the cases). The Statistical Package for the Social Sciences (SPSS, 14.0, 2005) was used for evaluation of missingness with the indicants selected, including Little's test for Missing Completely At Random (MCAR). Additionally, M-Plus provided a summary of missing data patterns for the indicants.

Evaluation of model fit. The models were evaluated using both measures of absolute fit and of comparative or relative fit. For absolute fit values, the chi-square test statistic, the standardized root mean square residual (SRMR) and the Root Mean Square Error of Approximation (RMSEA; Steiger, 1989) were used. The chi-square (X^2) statistic is not known for its high utility in examining model fit, but since it is so widely used and reported, it was included in the analysis for ease of comparison with other

similar studies. The chi-square tests the null hypothesis that the target (or default) model fits the data as well as does the full or saturated model. One of the inherent problems with use of the chi-square is that it is influenced by sample size, so even well-fit models can appear misspecified given a large enough sample size.

The SRMR is an index of the amount of difference between the observed and predicted correlations, which should be a small value. For this index, SRMR = 0 indicates perfect model fit, so higher values indicate increased "badness-of-fit", the actual value the mean absolute correlation residual (no residual – perfect model fit). The criterion for acceptable model fit in terms of SRMR will be SRMR $\leq$ .05. Another measure of absolute model fit for use in this study is the RMSEA (Steiger, 1989), which also evaluates fit by examining of the size of residuals. This is a parsimony-adjusted index, as the formula has a built-in function that corrects for model complexity. The RMSEA index is somewhat different in that it does not assume the model's perfect fit with the population, and thus does not require a true null hypothesis (Kline, 2005). For evaluation of model fit, RMSEA standards recommended by Browne & Cudeck (1993) were used, which suggest that (a) an RMSEA value of .05 or less represents a "close fit", (b) an RMSEA between .05 and .08 represents "reasonably close fit", and (c) an RMSEA above .10 indicates "an unacceptable model."

For evaluating comparative model fit, differences in the three factor model were examined by comparing their chi-square values and the Tucker-Lewis Index values. Cut-off values for the chi-square test of model fit were established *a priori* at $p > .05$. The Tucker-Lewis Index (TLI; Tucker & Lewis, 1973) was used as a measure of comparative

model fit. For this fit index larger values indicate better model fit. While according to Hu & Bentler (1999), values of .90 or higher indicate acceptable model fit, the newer and more stringent cut points recommended by Yu (2002) of TLI $\geq$ .95 was used as criteria for good model fit. This comparative fit measure uses its own distinct formula for comparing the chi-square ratio to the fit of the "null" model in which the items are assumed to share no common variance (opposite of the saturated model). A model that is well-fitted across the board using absolute fit and relative fit provides stronger evidence of the validity of the measurement model in estimating the data observed.

Akaiki's information criterion (AIC) serves as a measure of relative fit – a ranking criterion - to compare the model' fit in a different way, by incorporating both measures of model fit and parsimony. It is a predictive fit index, typically used to compare competing nonhierarchical models using the same data. For AIC, the underlying assumption of the criterion is that for the optimal model, the parameter estimates do not represent a "true" value, but rather an approximation (Burnham & Anderson, 2002). For the AIC, greater parsimony and fit are indicated by lower values which are compared among models tested (according to guidelines from Burnham & Anderson, 2002). To compare, difference scores (ΔAIC) were calculated by subtracting the AIC value of best performing model from the AIC values of the other models tested. To determine relative fit, the following guidelines were established *a* priori and were used for comparison: a) ΔAIC values $\leq$ 2.0 were considered similarly-fitted models, b) ΔAIC values $\geq$ 4.0 but $\leq$ 7.0 demonstrate less model fit, and those with ΔAIC values > 10.0 were considered to show very poor fit relative to the best model.

Validity analyses. For evidence of both concurrent and predictive validity, mothers' ratings of toddlers' problem behaviors (on the Child Behavior Checklist) were regressed on the resultant latent constructs comprising emotion regulation in the best fitting measurement model. As described, one the hypotheses was that the latent factor NA would be positively correlated with externalizing problem behavior, while both EC (representing an attentional control function) and S (representing a dimension of positive affectivity) would be negatively correlated with externalizing problem behavior. In contrast, it was hypothesized that both NA and EC should be positively correlated with internalizing behavior, which may be characterized by the dual presence of negative emotionality coupled with over-controlled emotionality. It is hypothesized that S would be negatively correlated with internalizing behavior and low surgency would be indicative of withdrawal or avoidance behaviors.

These results support the construct validity of the measurement model, and the importance of infant emotion regulation as a transmission pathway for or early indicator of affective disorders. This knowledge could indicate a critical entry point for early intervention and prevention of a potentially deleterious development of behavioral problems.

CHAPTER IV

RESULTS

Indicant Selection

Upon initial review of item content it was apparent that several of the subscales analyzed within the models spanned more than one of the three hypothesized ER dimensions. For example, Backen Jones suggested that the TBAQ subscale "expresses pleasure" associated with strongly with S, but also moderately with EC (Kirk, 1995). Similarly, the ICQ subscale "unpredictability" was strongly associated with NA, and also moderately associated with *both* S and EC. To improve the possibility of getting a clean simple structure for the indicants it was recommended by expert panel members Seeley and Backen Jones to use individual items on the survey instruments as indicants rather than subscale scores.

For the selection process it was recommended that an item pool be generated for each dimension (NA, S, and EC) selected from the previously recommended subscales. The items from the recommended subscales were initially chosen based upon face validity, that is, those items from the subscales were chosen that best represented the component constructs of Negative Affectivity, Surgency, and Effortful Control. The item pools were subjected to Principal Components Analysis (PCA) for each dimension in order to identify the most salient items to be included as indicants in the measurement model.

Normally, PCA is used when seeking to provide an empirical summary of the data, whereas principal axes analysis is suggested for testing a theoretical model (Tabachnick & Fidell, 2001). However, principal axes analysis can create situations in which communalities may not always be estimable or may be invalid (e.g. generating values greater than 1 or less than 0), which results in items being dropped from the analytical model. Instead, the purpose of the analysis in this instance was data reduction, for which PCA is aptly suited. The focus was on reducing the numbers of indicants from a large pool of items to a smaller, more fruitful few items.

The *a priori* PCA item selection guidelines used included identifying a minimum of three indicants per construct demonstrating at least a .60 path coefficient on the unrotated first general principal component. This resulted in the eleven items presented in Table 3. The result was a more specific model, using far fewer questions, as the subscale scores consisted of the combination of several items per subscale.

Missing Data

An analysis of missingness was performed using SPSS and Little's test for Missing Completely At Random (MCAR) suggested that the data was indeed MCAR ($p >$ .05). In addition to the minimal and random nature of the missing data, the M-Plus procedure uses all available data using full information maximum likelihood to estimate the model parameters, thus, each parameter is estimated directly, without first filling in the missing data values for each case (Muthen, 1998-2007).

Measure Reliability (Cronbach's alpha)

Examination of the reliability of the items within each latent construct was done by calculating Cronbach's alpha for each dimension. Since the scales were similar but used slightly different wording, the Cronbach's standardized alpha was used and reported. The items associated with the NA factor, $\alpha = .826$; S, $\alpha = .763$; and for EC, $\alpha = .727$. All reliabilities were strong as expected, and associated with a unidimensional factor.

Descriptive Statistics and Distributional Assumptions.

Distributional assumptions of multivariate normality were tested by examining univariate and bivariate normality with SPSS 14.0 (2005) and were found tenable and were presented in Table 4. A few of the indicants ("stay upset" and demonstrated skew values slightly higher than the acceptable range (absolute values greater than 1.0), however, these differences were not extreme. The estimated kurtosis value from the descriptive statistics in statistical software SPSS 14.0 was higher than desired, 2.495, for this same variable. In terms of adequate sample size, the ratio of cases to number of variables used exceeded the criterion of 5:1, in fact, the ratio of cases to variables for this sample was 14:1 (N = 154; number of variables = 11).

Table 3

Individual Indicants, Sources, and Predicted Associated Constructs

Source	Construct	Item
ICQ	NA	How much does your child fuss and cry in general?
ICQ	NA	How easily does your child get upset?
ICQ	NA	Please rate the overall degree of difficulty your child would present for the average mother.
ICQ	NA	How easy or difficult is it to take your child places?
TBAQ	EC	When you removed something our child should not have been playing with, how often did s/he stay upset for 10 min or longer?
TBAQ	S	When playing quietly with one of her/his favorite toys, how often did your child smile?
TBAQ	S	When playing quietly with one of her/his favorite toys, how often did your child sound happy?
TBAQ	S	When being gently rocked or hugged, how often did your child smile?
TBAQ	S	When being gently rocked or hugged, how often did you child giggle?
TBAQ	EC	When upset, how often did your child change to feeling better within a few minutes?
TBAQ	EC	When you are comforting your upset child, how often does s/he calm down quickly?

Descriptive statistics for the selected individual indicants, in terms of mean, standard deviation and sample size used in the analysis are presented in Table 4.

Table 4.

Descriptive Statistics for Individual Indicants

Items	Mean	Standard Deviation	N
Fuss and cry	3.01	1.120	150
Easy upset	3.53	.895	150
How difficult	3.14	1.176	148
Take places	3.12	1.404	149
Stay upset	6.19	.956	149
Sound happy	6.05	.717	150
Smile at play	5.56	1.059	150
Feel better	5.40	.988	151
Calm down	5.63	1.071	150
Often smile	5.56	1.247	151
Often giggle	4.57	1.619	151

Note. Mean values only calculated to two decimal places by software, SPSS. SPSS is a registered trademark of SPSS Inc.

The additional assumptions for using confirmatory factor analysis were found tenable as well, including: (a) a minimum of three indicators per latent construct, (b) use of continuous observed variables, (c) the number of free parameters was less than or equal to the number of observations, and (d) every latent variable (factors and measurement error) had a scale (Kline, 2005). As a result of this change in indicants, the path diagrams structure remained essentially in the same configuration, while the subscales were replaced by individual items as indicants. The revised path diagrams are as follow in Figures 5 through Figure 7.

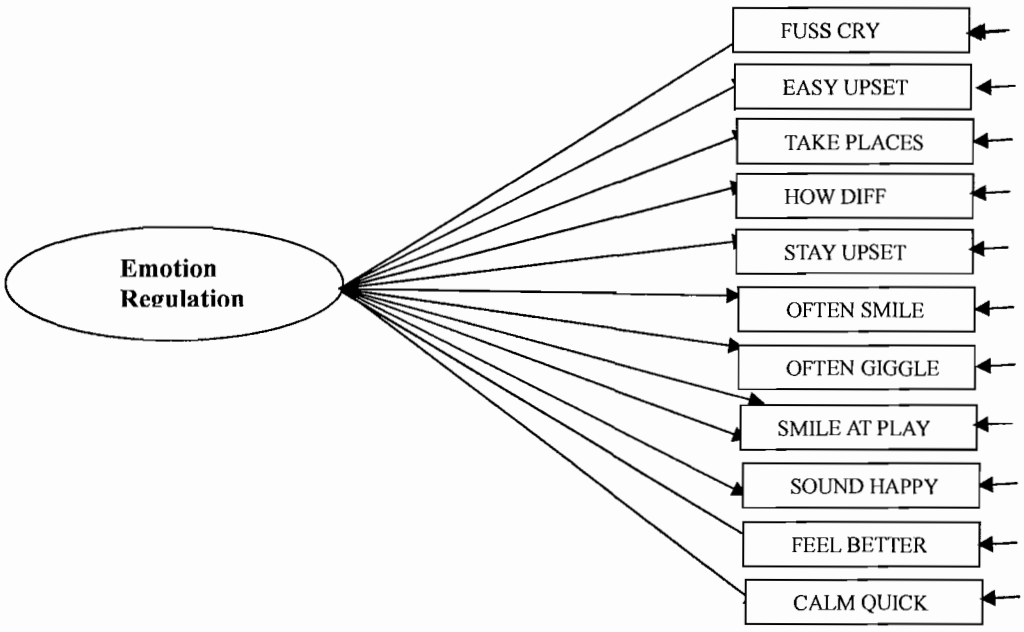

Figure 5. *One Factor Model of Emotion Regulation*

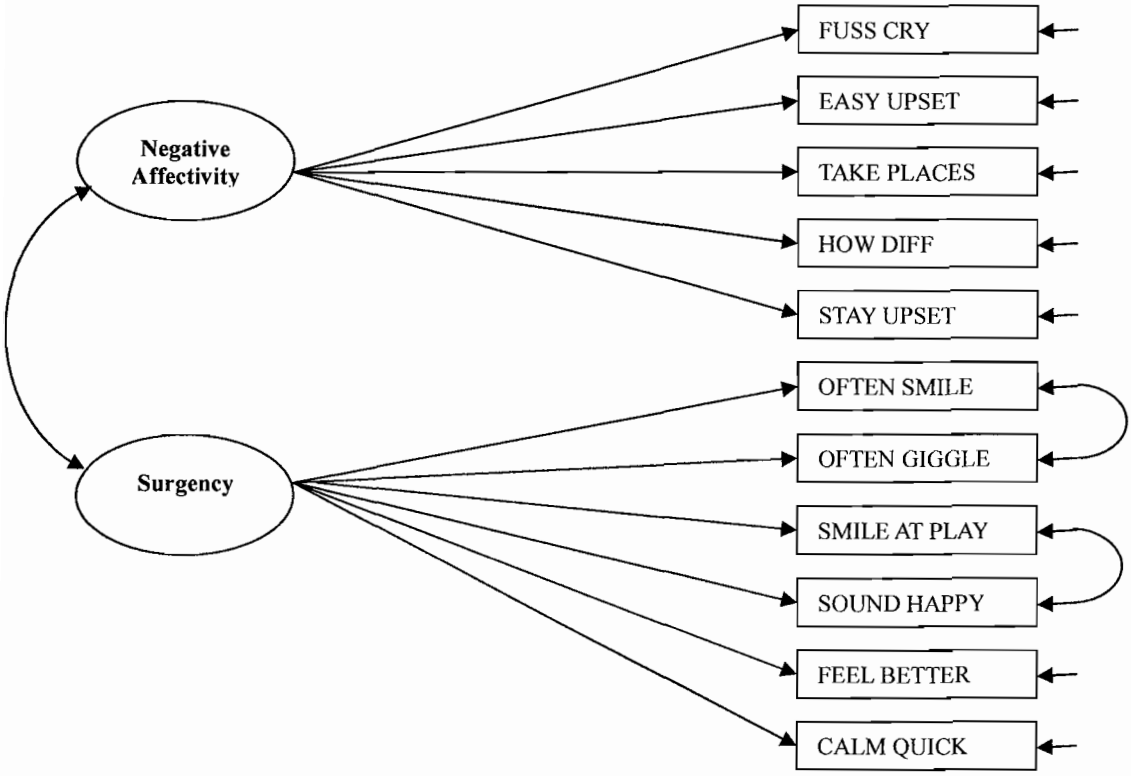

Figure 6. *Two Factor Model of Emotion Regulation*

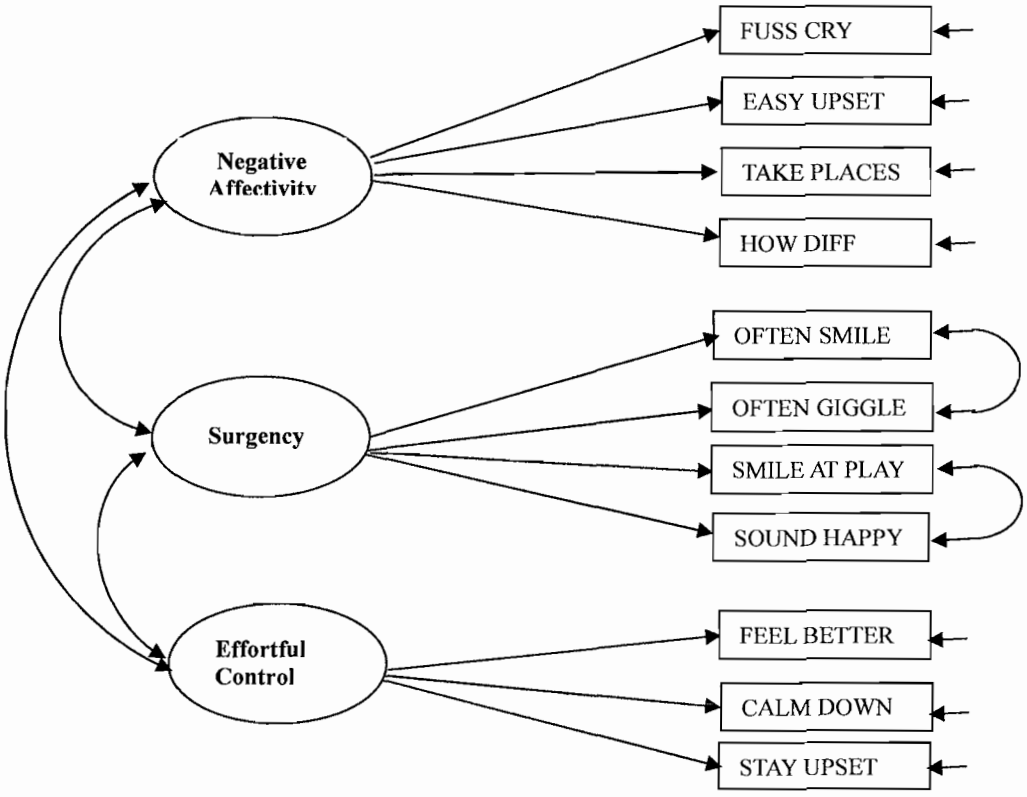

Figure 7. *Three Factor Model of Emotion Regulation*

The three separate models were analyzed with confirmatory factor analysis using Analysis of Moment Structures software (AMOS 6.0; Arbuckle, 2005). Of these models tested, only the two factor and three factor models converged. Results of the model fit indices are presented in Table 5.

Table 5.

Emotion Regulation: Model Fit Indices and Selection Criteria

	Indicators of model fit						Model selection criteria	
Model	X^2	X^2 df	X^2 *p*-value	TLI	RMSEA	SRMR	AIC	ΔAIC
ER2	191.889	43	.0000	.884	.148	.120	4771.180	31.602
ER2.1	104.967	41	.0000	.845	.099	.091	4688.256	48.678
ER3	106.285	41	.0001	.842	.100	.066	4689.576	49.998
ER3.1	52.287	39	**.0757**	**.966**	**.046**	**.049**	4639.578	0.000

Note. TLI = Tucker-Lewis fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; AIC = Akaike information criterion. Values meeting or exceeding model fit indices used are bolded.

As recommended by Kline (2005) a two-step modeling method was performed in which the "pure" measurement model must be fit before moving on to create a structural model. First, the proposed models were tested for goodness of fit using the individual survey items as indicants. The one factor, two factor, and three factor models all contained the same indicants, and the intention was to compare the model fit of the various nested models.

The one factor model, presented in Figure 2 and consisting of the eleven selected items, failed to converge, thus no model fit criteria were available and were assumed to be unmet. After the initial analysis was conducted, with a maximum of 1,000 iterations, the iteration maximum was increased to 10,000 and then to 20,000 with no successful model convergence.

Both the two-factor model and the three factor model converged successfully. Although the two factor model did converge, the model did not meet acceptable fit criteria, as shown in Table 5. Even though modification indices for the two-factor model indicated that fit could be improved by allowing two sets of residuals to correlate, this modification still did not result in an acceptable model fit for the two-factor model.

The three-factor model, predicting relations between indicants and putative emotion regulation constructs of NA, S, and EC, did successfully converge, and proved better fitted than both the saturated and independence models (as expected). However, the model met some acceptable model fit standards and failed to meet others previously outlined.

Modification indices indicated that allowing two sets of residuals (consistent with the two-factor model described previously) to correlate would likely improve model fit substantially. Clearly, however, the model as specified did not meet the *a priori* fit criteria. The specified residual pairs were significantly correlated, as shown in the correlation matrix of item-level indicants found in Appendix B ($r = .480$, $p = .0001$; $r = .448$, $p = .004$). The items were also functionally related to one another because they were responses to the same stimulus scenario or "question stem". For example, one of the items pairs were as follows:

"When gently rocked or hugged, how often does your child *smile*?" and "When gently rocked or hugged, how often does your child *giggle*?" These items are part of a set of questions that comes from the same question stem. They represent related but distinctly different observed child behaviors. It can be argued that smiling is a representation of passive positive affectivity response, while giggling represents a more active behavioral form of positive affectivity, or perhaps a more exuberant or intense form of expression. With the high correlations between items, if the results were due to method effects rather than the actual response, rather than testing the differing responses, the model may have been testing the "stem" part of the question: when gently rocked or hugged, how often does your child X?"

Considering that the original item pool was substantial, selection of these "pairs" was an unfortunate choice. Regardless of the possible reasons, the fit indices clearly indicate that the one. two, and three factor models as specified did not meet fit criteria.

While this model was used as the baseline measurement model for testing of subsequent structural models, the extreme limitations of the model in terms of generalizability and theory are here noted. A full discussion of the lack of model fit and an examination of the issues surrounding correlated residuals is presented in the discussion section.

The items, associated factors, and standardized parameter estimates for the three-factor model with correlated residuals are presented in Table 6, while the unstandardized parameter estimates, covariances, and associated factors are presented in Table 7.

Table 6.

Standardized Parameter Estimates for the Three-Factor Model of Emotion Regulation with Correlated Residuals

	Factor		
Indicants	Negative affectivity	Surgency	Effortful control
How difficult child for average mom	.683		
How easy/difficult take child places	.596		
How easy upset	.734		
How much fuss cry	.912		
How much giggle when held/rocked		.497	
How much smile when held/rocked		.659	
How much smile when play w/toys		.672	
How much sound happy play w/toys		.521	
How often stay upset 10 min			.633
How often cheer up in 5 min			.692
How often feel better quickly after upset			.717
Factor correlations			
Negative affectivity		-.210	-.506
Surgency			.354
Effortful control			

Table 7.

Unstandardized Parameter Estimates for the Three-Factor Model of Emotion Regulation

Model results	Estimate	*SE*	Estimate/*SE*	*p*-value
Negative affectivity by				
How much fuss cry	1.000	0.000	999.000	999.000
How difficult child average mom	0.797	0.094	10.318	0.000
How easy/difficult take child places	0.837	0.114	7.372	0.000
How easy upset	0.644	0.062	8.515	0.000
Surgency by				
How much giggle when held/rocked	1.000	0.000	999.000	999.000
How much smile when held/rocked	1.013	0.190	5.333	0.000
How much smile when play w/toys	0.891	0.392	2.273	0.023
How much sound happy play w/toys	0.464	0.240	1.936	0.053
Effortful control by				
How often feel better quickly after upset	1.000	0.000	999.000	999.000
How often stay upset 10 min	0.847	0.141	6.018	0.000
How often cheer up in 5 min	1.043	0.173	6.048	0.000
Factor covariances				
Negative affectivity with surgency	-0.170	0.111	-1.533	0.125
Effortful control with negative affectivity	-0.360	0.082	-4.370	0.000
Effortful control with surgency	0.199	0.083	2.400	0.016
Residuals				
Giggle when held/rocked with				
Smile when held/rocked	0.620	0.303	2.045	0.041
Smile when play w/toys with				
Sound happy play w/toys	0.214	0.130	1.642	0.101

Structural Models

Even though the three factor measurement model was determined *not* to meet the *a priori* fit indices criteria, the attempt was made to validate a structural model of emotion regulation that predicted criterion outcome measures. To validate the putative construct of emotion regulation as defined by the three factor measurement model, specific outcomes from the Child Behavior Check List (mothers' ratings) were regressed upon the three latent construct factors (Model ER3) in separate analyses. The choice of analyzing the separate structural models (as opposed to one all-inclusive structural model) was made in order to address more specifically the main purpose of the study, e.g., to identify those dimensions that make up the construct of emotion regulation. In separate structural analyses, the various unique contributions of the three related (but distinct) first-order constructs could be estimated.

One potential drawback in this choice of model development is that since the model only demonstrates the unique variance attributable to each first order latent construct, the common or shared variance contributions are not depicted within the graphic representation of the model. This may be misleading, in that the shared variance does not appear in the diagrams and may be inadvertently overlooked. While in some of the models the unique effects of NA or EC may be small, the amount of common explanatory variance shared between the two may be quite substantial. In this way, then, the total combined effects of the three dimensions may be obscured. The total combined model effect or proportion of explained variance is presented in the diagram as the R^2 value in models with only one manifest variable.

In the models measuring residual change in behavior ratings from 24 months to 48 months, the R^2 is replaced with a semi-partial squared correlation, or sr^2, which is the variance explained by the model holding the 24 month behavior rating score constant.

Future analyses could include the creation of a single structural model that would evaluate the relations between the latent constructs *and* the various outcomes of interest simultaneously. This type of model would limit the problem of capitalizing on chance due to multiple tests and would require no alpha correction to control for family-wise or experiment wise Type I error.

The outcomes of interest in the structural models included Internalizing behaviors, Externalizing behaviors, and an *ex post facto* analysis of Aggressive behavior (a dimension on the Externalizing behavior scale of the CBCL). The following path diagrams represent the structural models estimating unique contributions of variables within the model to the outcomes. The abovementioned CBCL subscales were regressed on the three factor model in these analyses. A Bonferroni correction to limit capitalization on chance due to the multiple-comparison estimates was applied to each family of models. A series of three model tests were performed on each type of outcome, e.g., internalizing, externalizing and aggressive behaviors. First, the behavior ratings at 24 months were regressed on the three latent constructs. Then, the behavior ratings at 48 months were regressed on the same constructs. Finally, the behavior ratings at 48 months were regressed upon the latent constructs and the behavior ratings of the same scale measured at 24 months.

These constituted the three separate model tests within the same "family" of tests. Accordingly, to control for Family-Wise Type I Error, the alpha level for the families of models was set at $\alpha = .05$, so that the alpha level for the individual model tests within each family was set at $\alpha = 0.5/3 = .0167$. Test statistics demonstrating *p*-values above .0167 were determined non-significant.

The presentation of the path diagrams for Internalizing behaviors begins with Figure 8. As in all path diagrams presented, bolded values indicate estimated path parameters meeting the criterion of $p < .0167$.

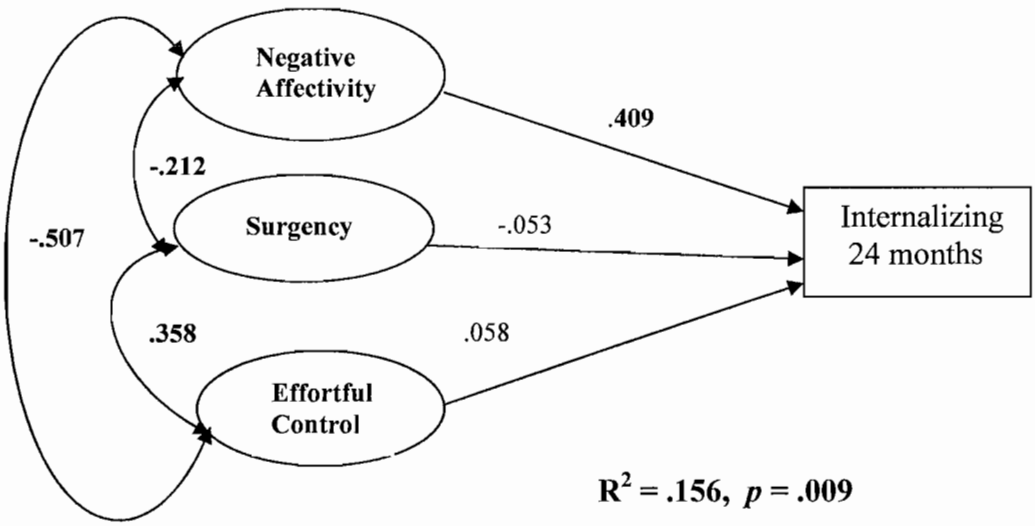

Figure 8. *Structural Model of Emotion Regulation Predicting Internalizing Behavior at 24 Months. Path coefficients in bold face are significant at p < .0167*

Internalizing behaviors. First, the hypothesis that emotion regulation should be associated with internalizing behavior, as measured by the CBCL Internalizing Scale, was

tested using mothers' ratings of child behavior in Wave 4, when the toddlers were aged 24 months. This analysis served as a measure of the model's concurrent validity and is presented in Figure 8.. Results showed that the model accounted for a small but statistically significant portion of the variance in internalizing behaviors (R^2 = .156, p = .009). Further, of the three latent constructs in the model only NA was statistically significantly associated concurrent internalizing behaviors (b = .409, p = .0001), while S and EC were not. As hypothesized, NA was positively associated with internalizing behavior.

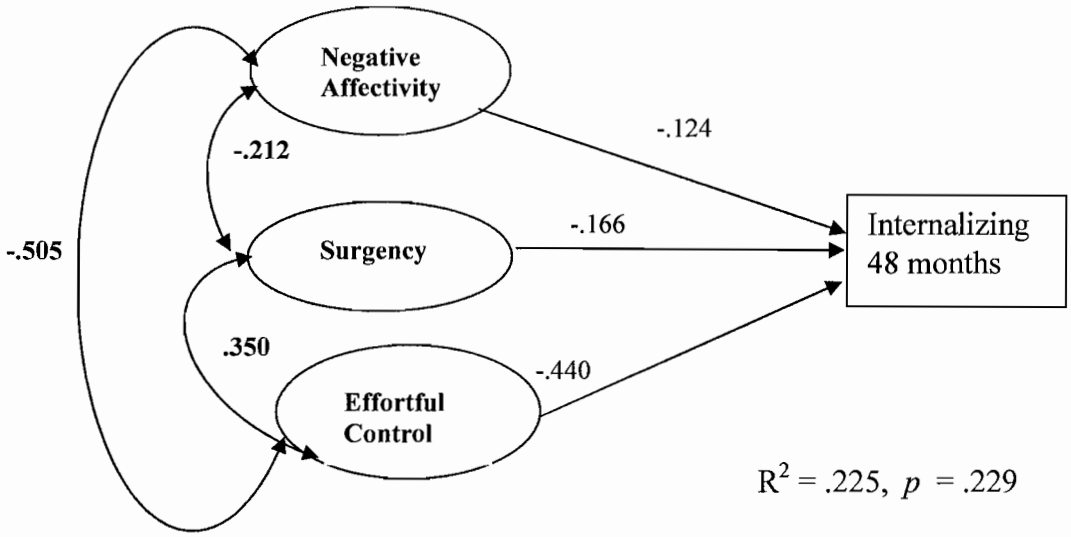

Figure 9. *Structural Model of Emotion Regulation Predicting Internalizing Behavior at 48 Months. Path coefficients in bold face are significant at p < .0167.*

Predicting internalizing behavioral outcomes. One form of predictive validity of the model was evaluated by regressing CBCL outcomes at Wave 6, when toddlers were aged 48 months, on the latent constructs.

Results presented in Figure 9 showed that, when measured at 24 months, neither the model nor the unique contribution of any of the latent constructs significantly predicted internalizing behavior at 48 months.

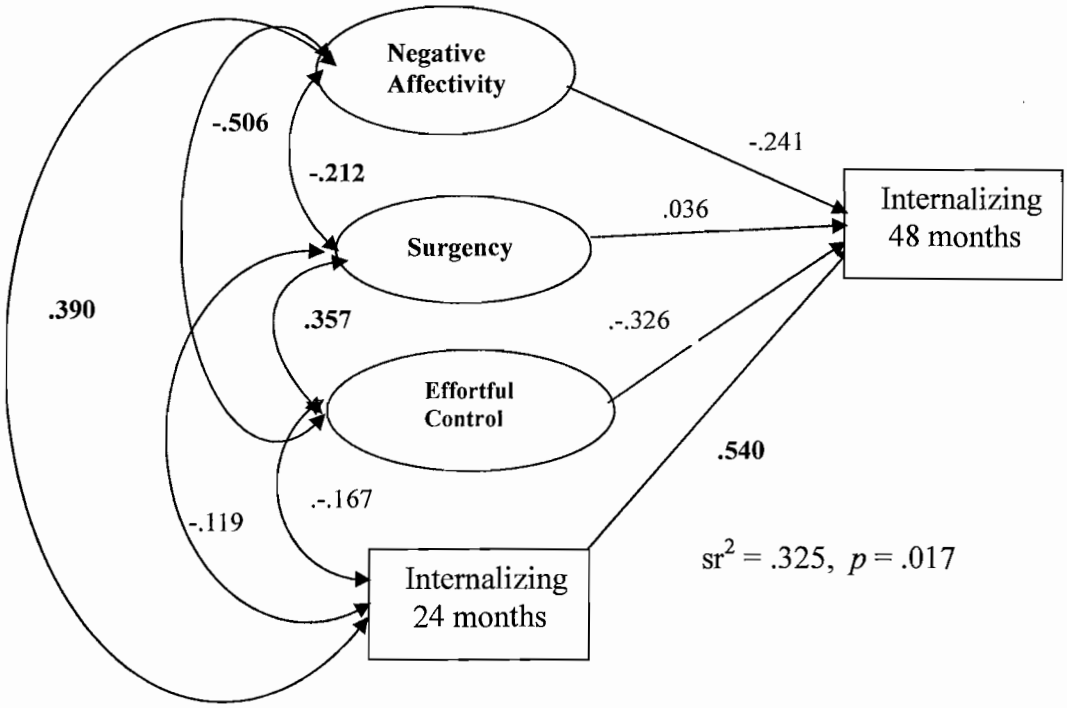

Figure 10. *Structural Model of Emotion Regulation Predicting Internalizing Behavior at 48 Months, Controlling for Internalizing Behavior at 24 Months. Path coefficients in bold face are significant at p < .0167.*

Predicting residual change in internalizing behavioral outcome. Finally, an additional type of structural model was examined that provided a more formal analysis of the model's ability to predict residual change in internalizing problem behaviors. The

model was evaluated by regressing CBCL scores at 48 months on latent constructs *and* CBCL scores at 24 months. This analysis, presented in Figure 10 evaluated the effects of ER while holding constant the 24 month internalizing CBCL scores. In this case, results showed that the model failed to statistically significantly account for variance in internalizing behavior at age 48 months ($sr^2 = .325$, $p = .017$). Further, none of the latent constructs, NA ($b = -.241$, $p = .377$), S($b = .036$, $p = .903$) or EC($b = -.326$, $p = .210$), uniquely significantly predicted residual change in internalizing behaviors at 48 months.

However, even though unique variance did not reach criterion of statistical significance, this estimation did not take into account the amount of shared or common variance among the latent constructs. For example, NA and EC have a strong negative correlation to one another, while there is also a moderate to strong correlation between EC and S. By only evaluating the unique contribution of each construct in predicting the outcome of internalizing behavior, the shared or common variance among variables is masked. Note, however, that the model -explained variance is denoted in the figure as the squared semi-partial correlation, or sr^2. This is the variance explained by the model separate from the variance explained by the 24 month behavior rating (which is held constant). A more complete structural model that regresses the outcome variable on the global construct of ER would examine the total amount of variance (shared and unique) that predict the outcome of problem behaviors.

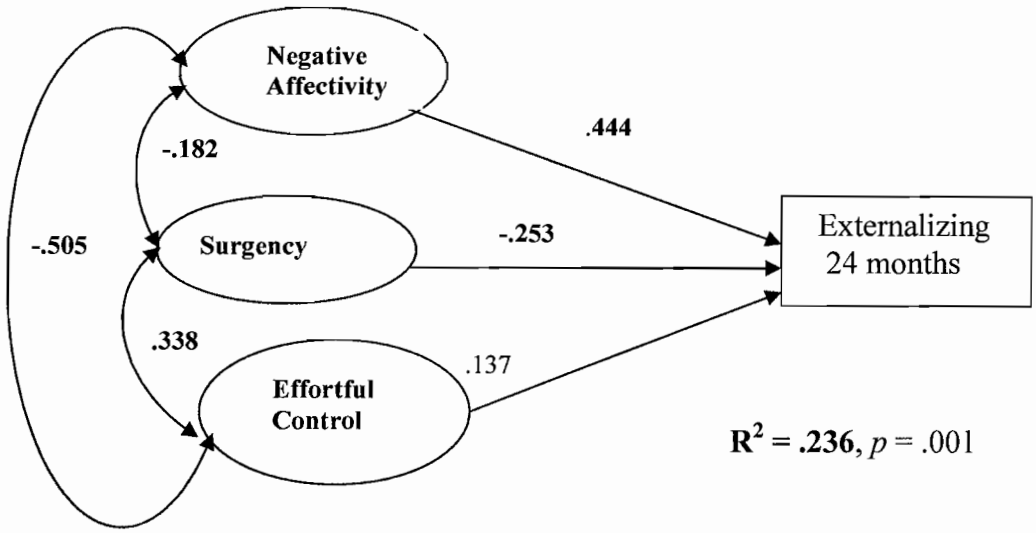

Figure 11. *Structural Model of Emotion Regulation Predicting Externalizing Behavior at 24 Months. Path coefficients in bold face are significant at p < .0167.*

Externalizing Behaviors

The same analyses were conducted with the outcome of Externalizing Behaviors using the CBCL as shown in Figures 5.1 through 5.3. Coefficients significant at $p < .0167$ are bolded. Results of the analysis of the structural model for externalizing behaviors (Figure 11) indicated that the model significantly accounted for a small to moderate (Kirk, 1995; Tabachnick & Fidell, 2001) amount of the variance in externalizing behavior at 24 months ($R^2 = .236, p = .001$). Similarly, both NA ($b = .444, p = .000$) and S ($b = -.253, p = .009$) were significantly associated with externalizing behavior at age 24 months, with NA showing a positive correlation with externalizing behavior and S showing a negative correlation, as hypothesized.

Again, it should be clearly noted that the path coefficients from the latent constructs to the outcome of problem behaviors represent only the amount of unique variance contributed by each latent construct, and does not include the common variance shared among the constructs.

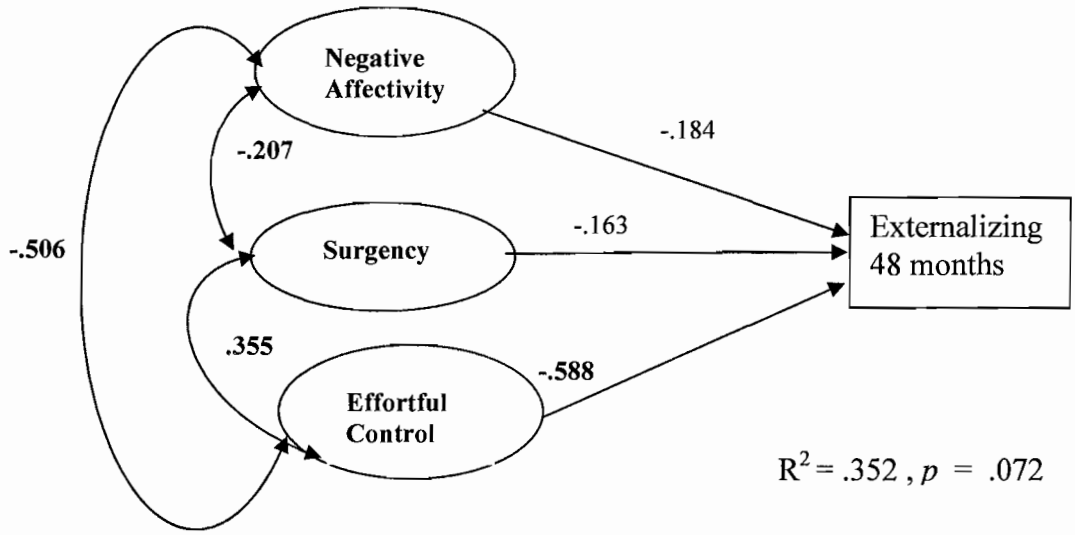

Figure 12. *Structural Model of Emotion Regulation Predicting Externalizing Behavior at 48 Months. Path coefficients in bold face are significant at $p < .0167$.*

Predicting externalizing behavioral outcomes. Results in Figure 12 showed that when measuring ER at age 24 months, the model did not account for statistically significant variability in externalizing behavior at age 48 months. NA (b = -.184, p = .492) and S (b = -.163, p = .571) measured at 24 months did not significantly predict externalizing behaviors at 48 months, however EC did (b = -.588, p = .014).

This supports the hypothesis that EC may be an important part of the ER model, as it is significantly negatively associated with externalizing behavior as much as two years beyond measurement at 24 months.

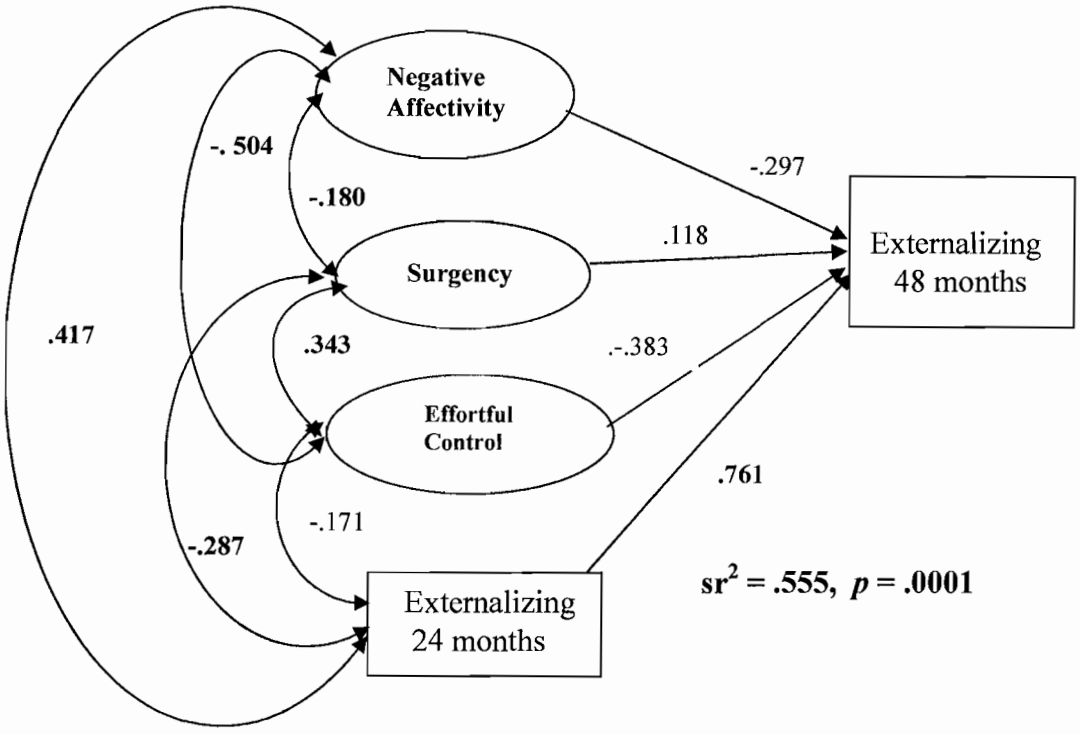

Figure 13. *Structural Model of Emotion Regulation Predicting Externalizing Behavior including Externalizing Behavior at 24 Months. Path coefficients in bold face are significant at p < .0167.*

Predicting residual change in externalizing behaviors. Externalizing problem behavior scores on the CBCL at 48 months were regressed on latent constructs and externalizing scores at 24 months. The effects of the model were evaluated in predicting residual change of externalizing behavior while holding constant the prior CBCL scores

obtained at 24 months. Results presented in Figure 13 showed that the ER model measured at 24 months accounted for about half of the variance in externalizing behaviors at 48 months when externalizing scores at 24 months were held constant (sr^2 = .555, p = .0001). None of the individual constructs NA (b = -.297, p =.189), S (b = .118, p = .608) and EC (b = -.383, p = .086) were uniquely significantly predictive of residual change in externalizing behaviors at 48 months, but taken together, the three part model significantly predicted the externalizing outcome.

Post Hoc Analysis of Aggressive Behavior

Because of the strong association between EC and externalizing behaviors, a *post hoc* decision was made to analyze the structural model's parameter estimates in predicting outcomes of aggressive behavior, to explore whether or not aggression was the primary dimension of externalizing behavior that was associated with effortful control. Results are presented in Figures 14 through 16.

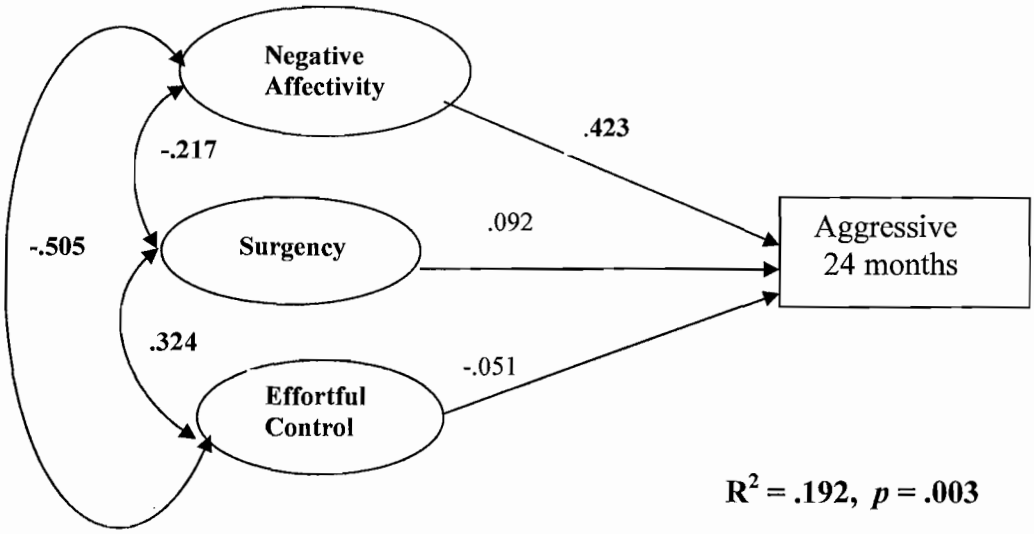

Figure 14. *Structural Model of Emotion Regulation Predicting Aggressive Behavior at 24 Months. Path coefficients in bold face are significant at p < .0167.*

Aggressive behaviors. Based upon the results from analysis of the Internalizing and Externalizing outcomes, the same analyses were used *ex post facto* to specifically predict aggressive behaviors (a component dimension of the externalizing scale) from the model. CBCL scores for aggressive behavior were regressed upon the same model at age 24 months and age 48 months. Results presented in Figure 14 show that the model explained a small but statistically significant amount of variance in aggressive behaviors at 24 months ($R^2 = .192,\ p = .003$). Of the latent constructs, only NA was significantly associated with aggressive behavior at 24 months ($b = .423,\ p < .001$) in terms of unique variance.

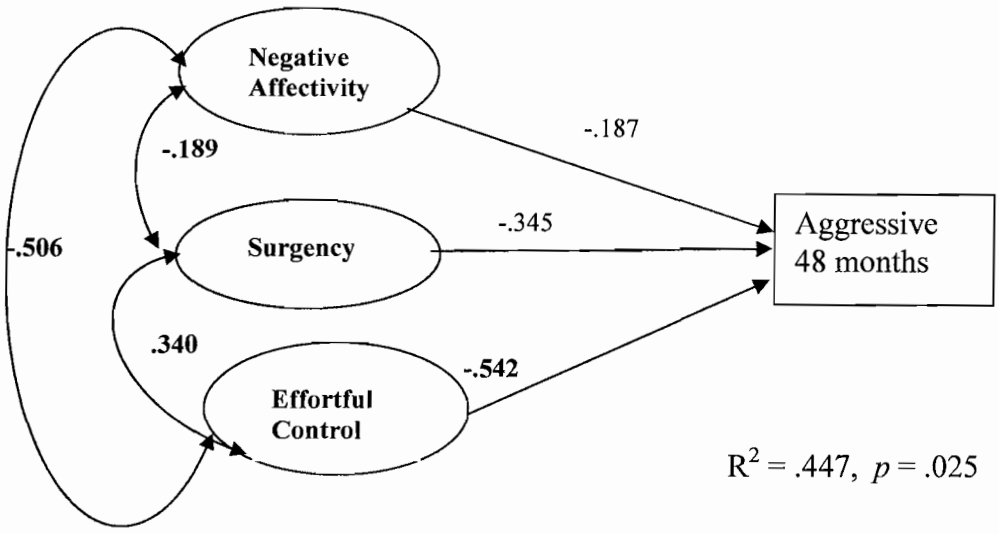

Figure 15. *Structural Model of Emotion Regulation Predicting Aggressive Behavior at 48 Months. Path coefficients in bold face are significant at p < .0167.*

Predicting aggressive behaviors. Figure 15 showed that the model (measuring ER at age 24 months), did not account for statistically significant variability in aggressive behavior at age 48 months using the Family-wise Bonferroni adjusted alpha level (α= .05/3) or $p < .0167$, ($R^2 = .447, p = .025$). None of the factors, NA ($b = -.187, p = .453$), S ($b = -.345, p = .157$) and EC ($b = -.542, p = .019$) measured at 24 months *uniquely* predicted aggressive behaviors at 48 months.

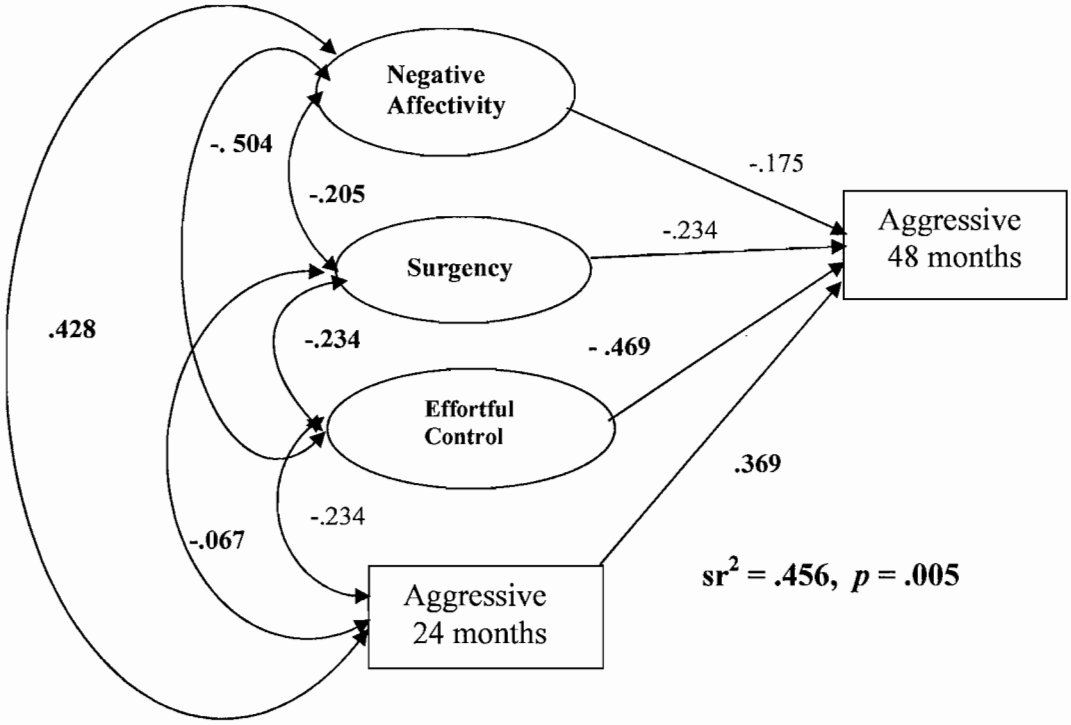

Figure 16. *Structural Model of Emotion Regulation Predicting Aggressive Behavior at 48 Months, Controlling for Aggressive Behavior at 24 Months. Path coefficients in bold face are significant at p < .0167.*

Predicting residual change in aggressive behaviors. Aggressive behavior scores on the CBCL at 48 months were regressed on latent constructs and aggression scores at 24 months. As described previously with internalizing and externalizing behaviors, the effects of the model in predicting residual change of aggressive behavior were analyzed by holding constant the prior CBCL aggressive behavior scores obtained at 24 months. Results presented in Figure 16 showed that ER measured at 24 months accounted for about 45% of the variance in aggressive behaviors at 48 months ($sr^2 = .456, p = .005$).

In this structural model, neither NA (b = -.175, p = .492), S (b -.234, p = .400), nor EC was significantly predictive of aggression (b = -.469, p = .044).

Also, aggressive behavior measured at 24 months did *not* statistically significantly predict aggression at 48 months (b = .369, p = .031) in terms of unique variance explained. Table 8 summarizes the results of the structural model testing which presented the unique contributions of each of the latent constructs to explaining variability in problem behaviors at 24 months and 48 months. Statistically significant positive correlations or path coefficients ($p < .0167$) are represented by a "+" in the columns while statistically significant negative correlations or path coefficients ($p < .0167$) are represented by a "-". The total ER model was superior to any of the individual latent constructs in predicting problem behavior outcomes. Next strongest in predicting problem behavior outcomes was the latent construct of Negative Affectivity.

Table 9 presents the summary of the model and latent constructs' relations to specific outcomes.

Table 8.

Summary of Model and Constructs in Predicting Outcomes at 24 Months and at 48 Months

Ratings of problem behavior	ER model	NA	EC	S
Internalizing@24 mo.	+	+		
Internalizing @48 mo.				
Internalizing @48 mo. holding Internalizing @24 constant				
Externalizing @24 mo.	+	+		-
Externalizing @48 mo.			-	
Externalizing @48 mo. holding Externalizing @24 constant	+			
Aggressive @24 mo.		+		
Aggressive @48 mo.			-	
Aggressive @48 mo holding Aggressive @24 constant	+		-	

Note. ER = Emotion regulation; NA = Negative affectivity; EC = Effortful control; S = Surgency

Overall model results. In sum, the ER model plus the covariate measured problem behavior at 24 months predicted residual change of problem behaviors at 48 months beyond the behavior rating scale given at 24 months. This ER model (with the 24 month behavior rating held constant) explained about 33% of the variance in internalizing behavior at 48 months ($sr^2 = .325, p = .017$) , while internalizing behavior at 24 months alone only explained about 24% of the variance in internalizing behaviors ($sr^2 = .237, p = .0001$). Similarly, for predicting externalizing behavior at 48 months, ER (and externalizing behavior at 24 months held constant) explained about 55% of the variance ($sr^2 = .555, p = .0001$) while externalizing behavior at 24 months alone explained only about 37% of the variance ($sr^2 = .367, p = .0001$). Finally, for aggressive behavior the ER model (with aggression at 24 months held constant) explained 46% ($sr^2 = .456, p = .005$) of the variance and aggressive behavior at 24 months explained only 28% of the variance in aggressive behavior at 48 months ($sr^2 = .284, p = .031$).

The best fitting three factor measurement model (the one fitted *post hoc* with two pairs of correlated residuals) was found to be significantly associated with some of the concurrent measures of internalizing and externalizing problem behavior, and to significantly predict both outcomes upon measurement two years later.

CHAPTER V

DISCUSSION

This section provides a discussion of the procedures, analyses, and results of the study. First, the process of fitting the measurement model is summarized, followed by a discussion of the observed pattern of correlations between latent constructs. Observations about the importance of the construct effortful control in understanding emotion regulation are presented next, followed by a discussion of model characteristics in predicting problem behaviors. Finally, the study's limitations are presented along with recommendations for future studies in emotion regulation.

Statistical Analyses and Model Selection

The primary purpose of this study was to develop a valid measurement model of emotion regulation for use with toddlers that could potentially provide a way to examine emotion regulation as a putative link between maternal depression and child risk. The main focus of the model development was to test the critical latent constructs that make up the broader construct of emotion regulation. Due to this focus, separate structural models were developed and tested to demonstrate the unique variance contributions of each of the three theoretical constructs.

Because depression is a disorder of increasing global magnitude and the disorder is often chronic and progressive, there is a sense of urgency in preventing its onset if possible – and as early as possible. The familial aggregation of depression provides a target population of at-risk offspring – those offspring with parents who have experienced depression. This study focused on the development and evaluation of a measurement model as a foundational step towards understanding a putative pathway by which familial depression may be transmitted.

Procedural summary. Three basic measurement models were developed *a priori*, and focused on three latent constructs, Negative Affectivity, Surgency, and Effortful Control. The selection of these constructs was based upon previous empirical research and theory. Mothers' report of infant/toddler behavior at 24 months and 48 months was selected as the source of the data for this study from the Infant Development Study. Initially, subscale values, recommended by an expert panel and taken from two commonly used infant/toddler temperament scales were used as observed indicants of the latent variables. These proved unsatisfactory in the role of indicants, due to the inability through confirmatory factor analysis to obtain a clean simple structure. Subsequently, at the suggestion of the expert panel, individual items from the recommended subscales were used instead. The eleven resultant items were presented in Table 3.

Model fit. According to model fit indices and selection criteria presented in Table 5, all model fit criteria were met or exceeded for the three factor Model ER3.1 *only* after allowing two pairs of indicant residuals to correlate. As discussed, the procedure of allowing correlated residuals is not appropriate for a confirmatory factor analytic model, in which theory, not data, should drive the model development. Thus, the *post hoc* fitting using modification indices criteria indicates only that the model was *not* fit as originally specified. Allowing residuals to correlate changed the nature of the analysis from one of model confirmation (theory-driven) to one of model exploration (data-driven).

The one factor model failed to converge, even after substantially increasing maximum iteration limits. Information-theoretic criteria for the two factor and three factor models presented in Table 5 suggest that of the measurement models evaluated, the three factor model (with two sets of correlated residuals) fit the data best (AIC = 4639.578). Figure 5. shows the path diagram of this model with standardized parameters included.

Use of modification indices. MacCallum, Roznowski, and Necowitz (1992) argued that the "common practice to modify [that] model to improve its fit." (p. 490) may result in a model that fits – but only the sample with which it was tested. Of several critical issues that MacCallum et al (1992) cover, one of the most persuasive arguments against the use of fit modifications is that of the capitalization on chance. By making *ex post facto* adjustments in association, the likelihood increases that the model is being fit to the random idiosyncrasies of a particular sample. An even more important consideration is that these post hoc modifications actually change the *nature* of the

analysis from a theory-driven confirmatory factor analysis to an empirically-driven exploratory analysis, in which data, not theory, dictate the model design. The end result is a model that is extremely well fitted to the sample-specific idiosyncrasies of one sample, warts and all, which fails to replicate with other samples. MacCallum's studies showed that even with moderate to large sample sizes (N = 300-400), some models with post hoc correlated residuals, for example, failed to replicate with other independent samples. This, among many studies, led MacCallum to assert that only rarely do such "over-fitted" models reveal a correct model.

Muthen (2008), recognizing the problem of using modifications to fit confirmatory factor analysis models, described a relatively new procedure to address some of these issues with a type of EFA-SEM modeling analysis. The ESEM provides an EFA-type of measurement model with rotations that can be used in structural equation modeling in M-Plus. This follows Brown's 2002 suggestion that rather than over-modifying CFA models that do not initially fit well, creating a measurement model through EFA may be preferable. Lance and Vandenberg (2008) go so far as to call the "appropriateness" of allowing correlated residuals in a CFA model an "urban legend" which should be avoided in most cases. Still, Cole, Ciesla, and Steiger (2007) noted in a current review of SEM studies in five top-tier journals published by the American Psychological Association, that between 26.6% and 31% of those studies modification indices suggestions were used to improve model fit without so much as a warning about capitalization upon chance, instability of parameter estimates with small samples, and the likelihood of limited generalizability.

Rather than admonishing researchers to "just say no" to correlated residuals, MacCallum does offer some specific recommendations for model improvement, including instructions to: (a) test alternative *a priori* models to create a model that does not require post hoc modifications, (b) cross-validate the model with independent samples or randomly split the extant sample if large enough, (c) clearly state limitations in the decision to correlate residuals, (d) acknowledge that the more modifications, especially at later stages of a fit sequence increases the capitalization on chance and decreases generalizability. While Kenny (2008) provided some rules about when residuals could arguably be allowed to correlate (theoretical meaningfulness and consistent rule application) the inherent problems with correlated residuals are simply too powerful to consider the procedure appropriate.

What is appropriate, however, is examining the resultant poor-fitting model (independent residuals) for reasons why the model did not fit. In this case, it may be that correlated residuals improved fit due to method effects in the questions themselves. Along these lines, Saris and Aalberts (2003) made special case for allowing correlated residuals when used in a *specific* type of Multitrait-Multimethod measurement models, not in confirmatory factor analysis. According to Saris and Aalberts, after examining several potential explanations for observed correlated residuals, e.g., method effects, relative answers, acquiescence bias, and variation in response functions, they analyzed seven separate data sets and determined that "method effect" was the most likely explanation for observed correlated residuals. While Saris and Aalberts' argument does *not* apply to this confirmatory model, the logic behind their assertion may shed light on

the underlying reasons that correlated residuals resulted in improved "fit". Exploration of this underlying reason may essentially help inform "what went wrong" in the specification of this particular measurement model. It may well be that the pairs of items were simply too intercorrelated (thus residuals were non-independent) due to method effects.

Methods effects are generally thought to refer to different types of data collection formats (e.g., interview, rating scales, direct observation), Smolkowski (2007) noted that the concept, in a broader sense, could include responses to the same form of question or to the same question stem, as in this instance. Smolkowski asserted that responses to the same question stem may correlate more highly with each other than with responses from other question stems. Smolkowski goes so far as to suggest that in certain special cases (although they should be used conservatively) correlated residuals may be acceptable (and even necessary). However, most methodological authorities, such as MacCallum would conclude that such attempts at *post hoc* model fitting are inappropriate.

In sum, examining the items themselves, and exploring the possibility of method effects on response dependencies, may lead us to improvement of model specification, such as: (a) removing one item from a pair that correlate too highly or for which modification indices recommend correlating, (b) refraining from using item responses belonging to the same question or stimulus stem, and (c) initially using a larger item pool to allow dropping out of items that correlate too highly. Even when specified models do not meet fit criteria, there is often something to be learned from the way in which the model does not fit.

Items. Items used as indicants associated strongly and unilaterally with their predicted latent factor, and not with the alternate factors, revealing a clean simple structure. Further, the three latent dimensions, at face value and by expert recommendation, were congruent with the operational definitions of Negative Affectivity, Surgency, and Effortful Control commonly described in the field of infant temperament (Rothbart & Derryberry, 1981; Posner & Rothbart, 2003; Ahadi, Rothbart, & Ye, 1993; Kochanska, DeVet, Goldman, Murray, & Putnam, 1994; Rothbart, Ahadi, and Evans, 2000; and Whittle, Allen, Lubman, & Yucel, 2006). Thus, there is conceptual support for the three factor model using these constructs.

One of the inherent weakness of confirmatory factor analysis, along with other similar forms of modeling, however, is that there may be other alternate models which fit the data as well or better as any of the three tested here. Alternate models will be presented in the discussion of study limitations.

Linking ER to internalizing behavior. The theoretical foundation of this study was that a measurement model of emotion regulation would provide a better understanding of the potential mechanism by which depression risk is transmitted from parent to offspring. It was hypothesized that evidence of a significant link between suboptimal emotion regulation and internalizing behaviors, specifically, would point to the beginnings of depressive behaviors early on in a child's development. However, the emotion regulation model was also predictive of externalizing behaviors as well as internalizing behaviors. Several factors could account for the emergence of the association of both externalizing

and internalizing behaviors with emotion regulation. First, the inherent difficulties of assessing behavior in very young children may blur the lines between overt externalizing and internalizing behaviors. Some researchers describe the emotionality of very young children by a broad dimension of distress that may not readily be distinguishable into divisions of externalizing and internalizing (Sheeber, personal communication, 10/17/2007). That is, there may not be, at 24 months, the degree of specificity required to distinguish successfully between externalizing and internalizing problems, rather, parents observed a more generalized, but noticeable, degree of distress in their offspring. A substantial body of literature supports the relative difficulty in detection of internalizing problems as compared with externalizing problems (Cicchetti, 1984; Merrell, 1999).

Patterns of association. In order to evaluate the hypothesized association patterns between factors and problem behavior outcomes, structural models were created and tested for each outcome of interest. Suboptimal patterns of the three constructs, NA, S, and EC, were hypothesized to be associated with at least two different types of problem behavior, e.g., internalizing and externalizing. It was expected that positive correlations with NA and EC coupled with a negative correlation with S would be associated with internalizing behaviors, as internalizing behavior is often characterized by negative emotionality and over-control.

For externalizing behaviors, a pattern of positive correlation with NA and negative correlations with both S and EC was expected, characterized by reduced regulatory capacity over the expression of negative emotionality. These expected patterns

of correlation within the structural models were not found in the results, as in many cases, even when the model was statistically significantly associated with problem behavior outcome, the individual latent constructs were not. However, both NA and EC were at times found to be significant, salient dimensions in both association with and prediction of problem behaviors.

To summarize, within the ER model the three latent factors made specific contributions to explained variability in specific behavioral outcomes. Negative affectivity was significantly associated with internalizing behaviors; both NA and S were significantly associated with externalizing behavior, and only EC was negatively associated with and predictive of aggressive behaviors. This is conceptually logical, since the behaviors making up the aggression scale involve more overt, active behavior than do the rest of the target behaviors included in the internalizing and externalizing scales. Thus, the regulatory or inhibitory control dimensions of EC would likely impact the degree of overt aggressive behavior in a young child.

The importance of effortful control. The inclusion of the effortful control construct proved particularly useful in predicting externalizing behavior and aggression at 48 months. This finding was important, because its inclusion was central to this theoretical model, and constituted an addition that made this model unique compared with many traditional ER models. Many models of emotion regulation and emotionality are comprised of a two-dimensional schema, in which some form of negative and positive emotions are represented (Clark, Watson, & Carey, 1988), either through approach and

avoidance dimensions (Davidson et al.1992), through constructs of inhibition and activation (Gray, 1994; Clark & Watson, 1994), or by a single factor continuum from negative to positive affectivity (Rothbart, 1981). Rothbart and Posner's (2001, 2003) description of effortful control as inhibition of a dominant response in order to perform a non-dominant response was demonstrated in this study's finding that those toddlers who showed difficulty inhibiting the dominant response to perform the non-dominant at 24 months, also showed residual change in aggressive behaviors from 24 months to 48 months. Effortful control within the model significantly predicted residual change in both aggressive behavior and the broader subscale of externalizing behaviors at 48 months, suggesting that at age 24 months, a child's regulatory skill development (EC) was a more salient dimension than negative affectivity in predicting later externalizing problems.

It is this regulatory function, effortful control, which may have direct ties to such physiological regulatory processes such as vagal tone – the degree of variability in the heart rate for each respiratory cycle. Vagal tone, or the control over the "braking and accelerating" aspect of heart rate represents one type of physiological experience and expression of emotion in the body. Perhaps future research may focus on the simultaneous measurement or covariance of overt behavioral and physiological dimensions of regulatory control. Negative affectivity, however, was positively associated with all three problem behavior categories, either by concurrent association or some form of prediction.

This is congruent with the emotion regulation literature that describes NA as a critical feature in both depressive and anxiety problems (Downey & Coyne, 1991; Whittle, Allen, Lubman, & Yucel, 2006).

Model Characteristics in Predicting Problem Behavior

In predicting problem behavior at 48 months, ER measured at 24 months was not as useful. For both internalizing and externalizing behaviors, the predictive power of the model was not significant without using the covariate of problem behaviors measured at 24 months. The strongest evidence of the model's overall predictive validity came from the models in which problem behaviors were regressed upon the three latent constructs and the problem behavior measured at 24 months. This analysis addressed the question, given equivalent problem behavior ratings at 24 months, how much does the model of ER predict increases in problem behaviors? The ER model thus tested, statistically significantly predicted all three problem behavior dimensions (internalizing, externalizing, and aggression) at 48 months.

Given the relatively strong predictive ability of the problem behavior measured at 24 months in predicting behavior at 48 months, one might ask, why not simply predict problem behavior later by measuring problem behavior early on? Why examine emotion regulation instead? The first answer is that ER predicted much more than a single dimensional construct like externalizing behavior; it predicted all three problem behavior outcomes measured.

The second answer is that the ER model (along with problem behavior at 24 months as a covariate) was a better predictor than the same behavior measured at 24 months. For each of the problem behaviors measured, the ER model measured additional dimensions beyond the scope of the specific behavior rating scale in predicting later problem behavior. Finally, if the purpose was to develop a screening measure for depression, there is evidence that previous depressive symptoms are the best predictor of future symptoms (Lewinsohn, Roberts, Seeley, & Allan, 1997; Lewinsohn, Seeley, Solomon & Zeiss, 2000). However, the stated purpose of this study was to develop a theoretical model of a transmission mechanism, not just a screening instrument. Informing successful intervention requires more than just measuring behavior; it requires development of a comprehensive conceptual model of the underlying mechanism to understand the critical dimensions in the origins of psychopathology.

For a measure containing only eleven items, one that is easily administered to parents or caregivers by mailer questionnaire, telephone, or over the internet, the ability to predict the residual change of problem behaviors in children two years before they occur seems promising, if not invaluable. The value in terms of potential prevention of human suffering cannot be underestimated. Pediatricians routinely screen for hearing, vision, growth, and other developmental milestones at well-baby visits. In light the potential value in preventing future problems, shouldn't they screen for emotion regulation as well?

Limitations

There are several limitations in this study, beginning with the age of the extant data set. One of the most critical limitations of the model of emotion regulation in this study was the lack of specificity in predicting internalizing behaviors. Rather than finding a clear link between suboptimal emotion regulation and internalizing behaviors, the externalizing behaviors were more strongly predicted by the model. To enhance specificity, perhaps the addition of a broader range of items to the current set would more directly target internalizing features of toddler distress. Additional items may also enhance the reliability of the instrument. While items selected showed good reliability in association with their expected factors, the breadth of the item content was inadequate. Dimensions such as sadness and withdrawal were not adequately assessed by the items used.

Further, this study began in 1996, and so at this writing, the "infants" in the Infant Development Study are becoming teenagers. Infants raised in the 1990s are not the same as infants raised in the first decade of the new millennium; the world has changed. Further, although the large-scale study used the most up-to-date measures at the time, improvements and revisions have been made to the instruments used, and like any assessment instruments, they may have become dated or less relevant over time. As in many studies, the participants in this study were geographically and ethnically homogenous, which may limit the extent of generalization to other populations.

Especially in the arena of parenting and child development, cultural expectations, contexts, parenting practices and behaviors are particularly diverse, and thus, these parents and offspring may represent a rather narrow demographic.

On a cautionary note, the use of structural equation modeling carries the inherent tendency of encouraging over-interpretation or over-generalization of results and reification of constructs (Kline, 2005). A complicated variance/covariance matrix and beautifully drafted path diagrams can tempt one to conclude that the model tested represents *the* model of the construct, not simply *one* possible model out of many.

Also, while the IDS study surveyed mothers, fathers, and used trained clinicians to observe parent-child dyads during scripted laboratory tasks, the scope of this study was limited to mothers' ratings of toddler behavior. The evaluation of data from a variety of sources would provide a more comprehensive evaluation of the construct of emotion regulation. Comparison of mother and father report could be examined as well as comparing mothers' and fathers' perceptions of child behavior, to that of trained observers. The original IDS data contains far more information about infant development of emotion regulation (among other things) than the small section examined in this study. Along those lines, the nature of the data is longitudinal, and analyses that make good use of the data gathered over time should be used. Multi-level modeling, including longitudinal growth modeling techniques could be used to examine trajectories of parent and child characteristics over the course of their development and to look for variables in parents, families, infants, and environments that influence the development of emotion regulation over time.

Use of seven or eight measurement occasions would provide better understanding of the trajectory of infant development of emotion regulation over time and the factors that influence its development.

Recommendations for Future Studies

Future studies could build on this measurement model by examining fathers' and clinicians' observations, in addition to mothers' observations, to provide a more comprehensive analysis of regulatory development. Development of a multi-trait, multi-source model may help provide answers to the question, how much of the variability in toddler behavior is due to true individual differences and how much depends upon who is reporting? Do mothers and fathers generally agree in their perceptions of offspring emotionality and behavior, or are there significant differences between their perceptions and experiences with the child? Studies such as these could examine parents' unique roles in shaping emotional development within the family.

As mentioned, future studies should strive for a larger and more comprehensive item pool in order to more fully tap into the domain of toddler emotion regulation. An enlarged item pool may have allowed the removal of items from the same question stems, and substitution of other items introducing a better representation of the domain of emotion regulation. The model tested failed to demonstrate acceptable fit, in part, perhaps due to the overlapping variance due to method effects. Selection of other items, for example, may have expanded or improved the assessment of behaviors related to unhappiness, sadness, crying, and withdrawal, which were not covered in this item pool.

These additional dimensions could have more thoroughly covered the domain of emotions requiring regulation from the toddler. Further, in addition to other items, different model configurations using the existing items could have been used. For example, in this study, the primary focus was identifying the component dimensions of emotion regulation, so that the problem behavior outcomes were regressed upon the individual dimensions of ER, rather than the global composite construct of ER. One of the drawbacks of the present models, as mentioned earlier, is that by modeling the unique variance attributable to each separate latent dimension, the shared or common variance explaining the outcomes was essentially masked.

Further, many other structural models could be tested with these same data. In particular, a broader structural model could be designed, in which the observed outcomes of internalizing and externalizing behaviors at 48 months could simultaneously be regressed upon the latent constructs and early problem behavior measures. The composite structural model –such as the one depicted in Figure 17, which includes only ER as a broad construct (not individual NA, S, or EC) may serve to better depict the total variance in outcomes explained by the model (Stevens, personal communication, May 13, 2009).

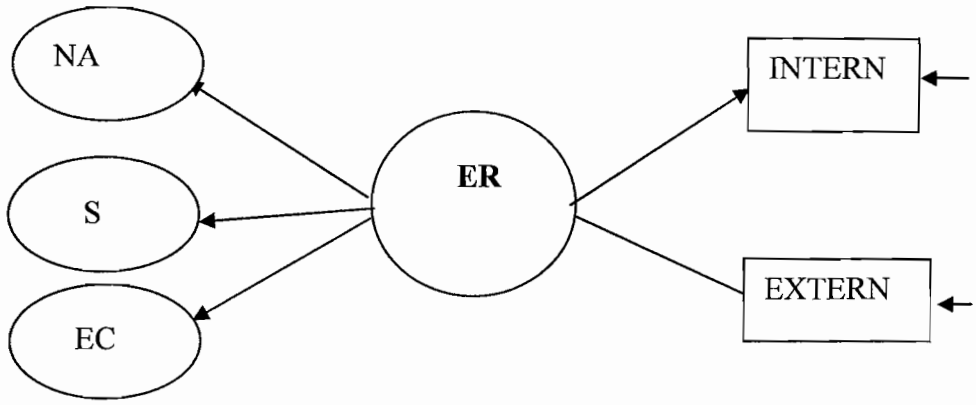

Figure 17. *Combined Structural Model for Emotion Regulation and Problem Behavior Outcomes*

In addition, the finding that early measured EC was a more critical factor in predicting later problem behaviors brings up the possibility that EC may be a mediator between problem behavior at 24 months and the same behavior at 48 months. A mediation model, in the tradition of Baron & Kenny, 1986 would certainly be a possible future direction for examining this phenomenon.

In terms of improving the statistical analysis, greater exploration of the distributional properties should be done prior to further analysis. While tests of univariate and bivariate normality, and scatterplot visual analysis were used to estimate multivariate normality, the absence of a strong individual measure for assessing multivariate normality (like Mardia's statistic, which evaluates multivariate skew and kurtosis) left some questions unanswered. In addition, as in many studies using survey response data, the numerical responses were analyzed as continuous equal-interval data, while technically they were more akin to ordinal data, which are often treated somewhat differently.

Schumacker & Lomax, 2004, assert, however, that for ordinal data with < 15 values may be assumed to be normal within the +/-1.5 or even +/- 2.0 range of skew and kurtosis. Still, it would have been preferable to have used a specific test of multivariate normality to evaluate this important distributional assumption for structural equation modeling.

Because the IDS longitudinal data is rich with parental history, a critical line of questioning left unexplored by this study is, how does parental mental health history affect infant/toddler emotion regulation and later toddler problem behaviors? About a third of the offspring in the study had a parent with depression, while a third had other mental diagnoses, and a third had no known mental health issues. These participants could be stratified into groups by mental health category in order to evaluate whether or not parental mental health status was associated with offspring emotionality and problem behavior. For participants with Major Depressive Disorder, factors to examine include parents' age at onset of depression, severity, recurrence, timing of depressive episodes relative to offspring's birth, prenatal episodes vs. depression episodes concurrent with childrearing, and family density of depression. Given a valid measurement model, it would be possible to examine the relation between parents' depression characteristics and trajectories that move toward suboptimal emotion regulation and emergence of problem behaviors.

Understanding the complex and interactive transmission pathways may lead us to earlier and better intervention strategies to support optimal development of emotion regulation for those at risk, and ultimately toward prevention of this rapidly increasing global health burden.

Perhaps someday schools will expand the annual health screening to include not only height, weight, vision, and dental health, but emotional health screening as well.

APPENDIX A

INDICANT SELECTION RECOMMENDATIONS FROM EXPERT PANEL

Emotion regulation IDS	Negative affectivity	Surgency	Effortful control
ICQ			
Fussiness/difficulty	Yes		
Unadaptability	Yes	Little	
Unpredictability (persistent)	Yes	Little	Little
Inactivity (unsociable)		Yes	
TBAQ			
Tend to express pleasure		Yes	Yes
Interest/persistence		Little	Yes
Activity level		Yes	Little
Social fearfulness	Yes	Yes	
Anger proneness	Yes		Some
Soothability	Yes	Little	Yes
CBCL			
Internal scale (total)			
Emotionally reactive	Yes		Little
Anxious/depressed	Yes		
Somatic complaints	Yes		
Withdrawn	Little	Little	
Externalizing scale (total)			
Attention problems		Yes	Some
Aggressive behaviors	Yes	Yes	

Note. ICQ = Infant Characteristics Questionnaire; TBAQ = Toddler Behavior Assessment Questionnaire; CBCL = Child Behavior Checklist 2-3

APPENDIX B

CORRELATION MATRIX AND STATISTIAL SIGNIFICANCE FOR MODEL INDICANTS

		Cry fuss	Easy upset	Take places	Diff child	Stay upset	Smile cuddle	Sound happy	Feel better	Smile play	Giggle play	Calm down
Cry fuss	Pearson r	1										
	Significance											
Easy upset	Pearson r	**.693**	1									
	Significance	.000										
Take places	Pearson r	**.524**	**.398**	1								
	Significance	.000	.000									
Diff child	Pearson r	**.596**	**.465**	**.547**	1							
	Significance	.000	.000	.000								
Stay upset	Pearson r	**-.323**	**-.244**	- .129	**-.289**	1						
	Significance	.000	.003	.119	.000							
Smile cuddle	Pearson r	-.097	-.808	-.051	-.091	.077	1					
	Significance	.239	.332	.543	.276	.354						
Sound happy	Pearson r	-.026	-.039	.015	-.039	.108	**.646**	1				
	Significance	.755	.637	.853	.644	.193	.000					
Feel better	Pearson r	-.321	**-.253**	-.112	**-.262**	**.486**	**.188**	**.257**	1			
	Significance	.000	.002	.176	.001	.000	.021	.001				
Smile play	Pearson r	-.150	-.081	**-.179**	-.152	.078	**.464**	**.337**	**.188**	1		
	Significance	.068	.324	.030	.066	.344	.000	.000	.021			
Giggle play	Pearson r	-.071	-.154	-.080	-.099	-	**.380**	**.208**	.059	**.643**	1	
	Significance	.387	.061	.331	.234	.031 .711	.000	.011	.468	.000		
Calm down	*Pearson r*	**-.355**	**-.207**	**-.191**	**-.320**	**.426**	**.211**	**.267**	**.496**	**200**	.075	1
	Significance	.000	.011	.021	.000	.000	.010	.001	.000	.014	.359	

APPENDIX C

CHILD BEHAVIOR CHECK LIST FOR INTERNALIZING AND EXTERNALIZING BEHAVIORS (ACHENBACH, 1988)

Internalizing			Externalizing		
Anxious/depressed	Somatic complaints	Withdrawn	Attention problems	Aggressive behavior	Emotionally reactive
Dependent	Aches pains	Acts young	Concentrate	Can't wait /frustrated	Disturbed change
Feelings hurt	Things out	Avoids eye	Can't sit still	Defiant/demanding	Twitching
Upset by separate	Headaches	No answer	Clumsy	Selfish/stubborn	Mood changes
Looks unhappy	Nausea	Refuses activity	Shifts quickly	Destroys others	Sulks
Nervous	Stomachache	Unresponse affect	Wanders	Disobedient /no guilt	Upset by new
Self conscious	Too neat	Little affect		No guilt /frustrated	Whining
Fearful	Vomiting	Little interest		Punishment/screams	Worries
Sad		Withdrawn		Fights/hits others	
				Hurts accidentally	
				Angry moods/temper	
				Uncooperative/wants attention	

REFERENCES

Achenbach, T. (1991). Manual for the Child Behavior Checklist/2-3 and 1991 profile. Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M., & Rescorla, L.A. (2001). *Manual for the ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth and Families.

Ahadi, S., Rothbart, M., & Ye, R. (1993). Children's temperament in the US and China: Similarities and differences. *European Journal of Personality*, *7*, 359–378.

Andrews, J., Lewinsohn, P., Hops, H., & Roberts, R. (1993). Psychometric properties of scales for the measurement of psychosocial variables associated with depression in adolescence. *Psychological Report, 73*, 1019-1046.

Angold, A., & Costello, E.J. (1993). Depressive comorbidity in children and adolescents: empirical, theoretical, and methodological issues. *American Journal of Psychiatry,* 150, 1779-1791.

Arbuckle, J. (2005). *Amos 6.0 user's guide*. Chicago, IL: SPSS Inc.

Asparouhov, T. & Muthen, B. (2008). Exploratory structural equation modeling. Accepted for publication *Structural Equation Modeling.* www.statmodel.com.

Ayissi L, Hubin-Gayte M. (2006). Newborn neonatal irritability and maternal post-partum depression/Irritabilité du nouveau-né et dépression maternelle du post-partum. *Neuropsychiatrie de l'enfance et de l'adolescence, 54(*2), 125–132.

Baron, R. & Kenny, D. (1986). The moderator-mediator variable distinction in empirical psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51(6),* 1173-1182.

Bates, J., Freeland, C., & Lounsbury, M. (1979). Measurement of infant differences. *Child Development, 50,* 794 – 803.

Bayley, N. (1969). *Bayley Scales of Infant Development.* New York: Psychological Corporation.

Beach, Henry, Stowe, & Newport (2005). Maternal depression: An adverse early environment. In: Riecher-Rossler, A., Steiner, M. (Eds.): Perinatal Stress, Mood and Anxiety Disorders. From Bench to Bedside. Basel, Karger, *Bibliotheca Psychiatrica, 173,* 70-84.

Beardslee, W., Schultz, L. & Selman, R. (1987). Level of social-cognitive development, adaptive functioning, and DSM-III diagnoses in adolescent offspring of parents with affective disorders: Implications of the development of the capacity for mutuality. *Developmental Psychology, 23*(6), 807-815.

Beck, A. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy,* 1, 5–37.

Billings, A. & Moos, R. (1985). Psychosocial processes of remission in unipolar depression: Comparing depressed patients with matched community controls. *Journal of Consulting and Clinical Psychology, 53(3),* 314-325.

Birmhauer, B., Ryan, N., Williamson, D., Brent, D., Kaufman, J., Dahl, R., Perel, J., & Nelson, B. (1996). Childhood and adolescent depression: A review of the past 10 years, Part 1. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 1427-1439.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.

Brent, D., Perper, J., Moritz, G., Liotus, L., Schweers, J., Balach, L., & Roth, C. (1994). Familial risk factors for adolescent suicide: a case-control study. *Acta Psychiatica Scandinavica, 89,* 52-58.

Brent, D. (2006). Screens and doors: The management of adolescent depression in primary care. *Archives of Pediatric Adolescent Medicine, 160*(7), 755-756.

Brody, A., Saxena, S., Silverman, D., Alborzian, S., Fairbanks, L., Phelps, M., Huang, S., Wu, H., Maidment, K., & Baxter, L. Jr. (1999) Regional brain metabolic changes in patients with major depression treated with either paroxetine or interpersonal therapy: preliminary findings. *Psychiatry Research, 91(3),* 127-39.

Brody, A., Saxena, S., Mandelkern, M., Fairbanks, L., Ho, M., & Baxter, L. (2001). Brain metabolic changes associated with symptom factor improvement in major depressive disorder. *Biological Psychiatry, 50(3),* 171-8.

Burnham, K. & Anderson, D. (2002). *Model selection and multi-model inference: A practical information-theoretic approach* (2nd ed). New York: Springer-Verlag.

Burt, K., Van Dulmen, M., Carlivati, J., Egeland, B., Sroufe, L., Forman, D., Appleyard, K., & Carlson, E. (2005). Mediating links between maternal depression and offspring psychopathology: the importance of independent data. *Journal of Child Psychology and Psychiatry, 46(5)*, 490–499.

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

Butler, E. & Gross, J. (2004). Hiding feelings in social contexts: Out of sight is not out of mind. In *The Regulation of Emotion, Eds. Philippot, P. and Feldman R.* Lawrence Erlbaum Associates, Publishers: Mahwah, New Jersey.

Calkins, S. D. (1997). Cardiac vagal tone indices of temperamental reactivity and behavioral regulation in young children. *Developmental Psychobiology, 31*, 125–135.

Calkins, S. & Hill, A. (2007). *Caregiver influences on emerging emotion regulation: Biological and environmental transactions in early development.* In: Gross, J. (Ed.) Handbook of emotion regulation. New York, NY, US: Guilford Press.

Cantwell, D., Lewinsohn, P., Rohde, P., & Seeley, J. (1997). Correspondence between adolescent report and parent report of psychiatric diagnostic data. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(5), 610-619.

Carver, C., and White, T. (1994). Behavioral inhibition, behavioral activation, and affective responses impending reward and punishment: The BIS/BAS Scales," *Journal of Personality and Social Psychology*, 67 (2), 319- 333.

Caspi, A., Sugden, K., Moffitt, T., Taylor, A., Craig, I., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A., & Poulton, R. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science, 301*(5631), 386-389.

Champagne, P. & Meeney, M. (2001). Like mother, like daughter: Evidence for non-genetic transmission of parental behavior and stress responsivity. *Progressive Brain Research, 133,* 287-302.

Chisholm, D., Sanderson, K., Ayuso-Mateos, J., & Saxena, S. (2004). Reducing the global burden of depression: Population-level analysis of intervention cost-effectiveness in 14 world regions. *British Journal of Psychiatry, 184,* 393-403.

Chorpita, B. (2002). The tripartiate model and dimensions of anxiety and depression: an examination of structure in a large school sample. *Journal of Abnormal Child Psychology, 30(2),* 177-190.

Cicchetti, D. (1984). The emergence of developmental psychopathology. *Child Development*, 55, 1-7.

Cicchetti, D., Ackerman, B., & Izard, C. (1995). Emotions and emotion regulation in developmental psychopathology. *Development and Psychopathology, 7*, 1–10.

Clark, D., & Beck, A. (1999). *Scientific foundations of cognitive theory and therapy of depression.* New York, NY: John Wiley & Sons.

Cloninger, C. R. (1986). A unified biosocial theory of personality and its role in the development of anxiety states. *Psychiatric Developments, 3,* 167-226.

Cole, P., Martin, S., & Dennis, T. (2004). Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. *Child Development, 75(2),* 317-333.

Conway, A. and McDonough, S. C. , 2006-06-19 "Longitudinal predictors of emotion regulation from infancy to preschool" *Paper presented at the annual meeting of the XVth Biennial International Conference on Infant Studies, Westin Miyako, Kyoto, Japan* Retrieved 04/09/2009 from http://www.allacademic.com/meta/p94335_index.html

Davidson, R. (1994). Asymmetric brain function, affective style, and psychopathology: The role of early experience and plasticity. *Development and Psychopathology, 6,* 741-758.

Davidson, R. & Fox, N. (1989). Frontal brain asymmetry predicts infants' response to maternal separation. *Journal of Abnormal Psychology, 98,* 127-131.

Davidson, R., Fox, N., & Kalin, N. (2007). Neural bases of emotion regulation in nonhuman primates and humans. In J.J. Gross (Ed.), Handbook of emotion regulation. New York: Guilford.

Davidson, R. & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences, 3*(1), 11-21.

Davidson, R. J., Pizzagalli, D., Nitschke, J. B. & Putnam, K. M. (2002). Depression: Perspectives from affective neuroscience. *Annual Review in Psychology, 53,* 545–574.

Davie, J. (2007). *Dialogues and Essays By Lucius Annaeus Seneca*, Seneca, John Davie, Tobias Reinhardt. (John Davie, Trans.) Contributor Tobias Reinhardt. Edition: revised. Published by Oxford University Press.

Dawson , G., Klinger, L., Panagiotides, H., Hill, D., & Spieker, S. (1992). Frontal lobe activity and affective behavior of infants of mothers with depressive symptoms. *Child Development, 63(*3), 725-737.

DeBary, W. T., Chart, W. T., & Watson, B. (1960). *Sources of Chinese tradition.* New York: Columbia University Press.

De Pauw, S., Mervielde, I., & Van Leeuwen, D. (2009). How are traits related to problem behavior in preschoolers? Similarities and contrasts between temperament and personality. *Journal of Abnormal Child Psychology, 37(3),* 309-325.

Descartes, R. (1955). *The philosophical works of Descartes* (Trans. E. S. Haldane & G. R. T. Ross). New York: Dover. (Original work published 1649)

Downey, G., & Coyne, J. (1990). Children of depressed parents: An integrative review. *Psychological Bulletin, 108,* 50-76.

Edborg, M.. Matthiesen, A., Lundh, W., & Widstrom, A. (2000). Some early indicators for depressive symptoms and bonding 2 months postpartum – a study of new mothers and fathers. *Archives of Women's Mental Health, 8*(4), 221-231.

Eisenberg, N., Guthrie, I., Fabes, R., Reiser, M., Murphy, B., Holmgren, R., et al. (1997). The relations of regulation and emotionality to resiliency and competent social functioning in elementary school children. *Child Development*, *68*, 295–311.

Eisenberg, N., & Morris, A. (2002). Children's emotion-related regulation. In H. Reese, & R. Kail (Eds.), *Advances in child development and behavior, 30*, 189–229. San Diego, CA: Academic Press.

Eisenberg, N., & Spinrad, T. L. (2004). Emotion-related regulation: Sharpening the definition. *Child Development*, *75*, 334–339.

Eysenck, H. J. (1990). Biological dimensions of personality. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 244-276). New York: Guilford.

Farmer, R., Seeley, J., Kosty, D., and Lewinsohn, P. (2008 *under review)* Refinements of the hierarchical structure of externalizing psychiatric disorders: Patterns of lifetime liability from mid-adolescence through early adulthood. *Manuscript under review (October 2008): Journal of Abnormal Psychology.*

Fernald, R. (2003). How does behavior change the brain? *Integrative and Comparative Biology, 43,* 771-779.

Forbes, E., Cohn, J., Allen, N., & Lewinsohn, P. (2004). Infant affect during parent-infant interaction at 3 and 6 months: Differences between fathers and mothers and influence of parent history of depression. *Infancy, 5*(1), 61-84.

Forbes, E. & Dahl, R. (2005). Neural systems of positive affect: Relevance to understanding child and adolescent depression? *Developmental Psychopathology, 17,* 827-850.

Garber, J., & Flynn, C. (2001). Predictors of depressive cognitions in young adolescents. *Cognitive Therapy and Research, 25,* 353-376.

Garson, G. David (n.d.). "Confirmatory Factor Analysis", from *Statnotes: Topics in Multivariate Analysis.* Retrieved 04/01/2009 from http://www2.chass.ncsu.edu/garson/pa765/statnote.htm

Gartstein, M. & Fagot, B. (2003). Parental depression, parenting and family adjustment, and child effortful control: Explaining Externalizing behaviors for preschool children. *Journal of Applied Developmental Psychology, 24*(2), 143-177.

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

Goldapple, K., Segal, Z., Garson, C., Lau, M., Bieling, P., Kennedy, S., & Mayberg, H. (2004). Modulation of cortical-limbic pathways in major depression: treatment-specific effects of cognitive behavior therapy. *Archives of General Psychiatry, 61*(1), 34-41.

Goldberg, L. R. (1990). An alternative "Description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology,* 59, 1216-1229.

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

Goldsmith, H., & Campos, J. (1982). Toward a theory of infant temperament. In R. N. Emde & R. J. Harmon (Eds.), The development of attachment and affiliative systems. New York: Plenum Press.

Goldsmith, H., & Campos, J. (1986). Fundamental issues in the study of early temperament: The Denver twin temperament study. *Advances in Developmental Psychology,* Hillsdale, NJ: Erlbaum.

Goodman, S., & Gotlib, I. (1999). Risk for psychopathology in the children of depressed mothers: a developmental model for understanding mechanisms of transmission. *Psychological Review, 106,* 458-490.

Gotlib, I, Lewinsohn, P., & Seeley, J. (1995). Symptoms versus a diagnosis of depression: Differences in psychosocial functioning. *Journal of Consulting and Clinical Psychology, 63*, 90-100.

Gross, J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology, 74*, 224-237.

Gross, J.J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*, 271-299.

Gross, J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 9*, 281-291.

Hammen, C. & Brennan, P. (2003). Severity, chronicity, and timing of maternal depression and risk for adolescent offspring diagnoses in a community sample. *Archives of General Psychiatry, 60*(3), 253-258.

Hankin, B., Fraley, R., & Abela, J. (2005). Daily depression and cognitions about stress: Evidence for a traitlike depressogenic cognitive style and the prediction of depressive symptoms in a prospective daily diary study. *Journal of Personality and Social Psychology, 88*(4), 673-685.

Henriques, J. & Davidson, R. (1990). Regional brain electrical asymmetries discriminate between previously depressed and healthy control subjects. *Journal of Abnormal Psychology*, *99*, 22–31.

Henriques, J. & Davidson R. (1991): Left frontal hypoactivation in depression. *Journal of Abnormal Psychology*, *100*, 535–545.

Hollon, S., Shelton R., Wisniewski, S. et al. (2006). Presenting characteristics of depressed outpatients as a function of recurrence: preliminary findings from the STAR*D clinical trial. *Journal of Psychiatric Research, 40,* 59–69.

Holmans, P., Weissman, M., Zubenko, G., Scheftner, W., Crowe, R., DePaulo, J. et al. (2007). Genetics of recurrent early onset depression: Final genome scan report. *The American Journal of Psychiatry, 164(2),* 248-258.

Horowitz J., & Garber, J. (2006). The prevention of depressive symptoms in children and adolescents: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 74*(3), 401-415.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Hume, D. (1896). A Treatise of Human Nature By David Hume, Lewis Amherst Selby-Bigge. Edition: 2, reprint Published by Clarendon press. Retrieved 4/09/2009 from http://books.google.com/books?id=5zGpC6mL- MUC&dq=a+treatise+on+human +nature+hume&source=gbs_summary_s&cad=0.

Jacques, D.H. (1879). *The temperaments.* New York: S.R. Wells & Co., Publishers.

Janson, H. & Mathiesen, K. (2008).Temperament profiles from infancy to middle childhood: Development and associations with behavior problems. *Developmental Psychology, 44*(5), 1314-1328.

Jonson, Ben. Ben Jonson's Plays. (1910). Everyman's Library. Edited by Ernest Rhys originally from Ben Jonson, *The workes of Benjamin Jonson* (London: Will Stansby, 616). *STC* 14751.

Kane, P., & Garber, J. (2004). The relations among depression in fathers, children's psychopathology, and father-child conflict: A meta-analysis. *Clinical Psychology Review, 24*(3), 339-360.

Keller, M., Lavori, P., Beardslee, W., Wunder, J., & Ryan, N. (1991) Depression in children and adolescents: New data on "undertreatement" and a literature review on the efficacy of available treatments. *Journal of Affective Disorders, 21,* 163-171.

Kendler , K., Thornton L., Gardner, C..(2000). Stressful life events and previous episodes in the etiology of major depression in women: an evaluation of the "kindling" hypothesis. *American Journal of Psychiatry, 157*, 1243–51.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences: 3rd Ed.* Pacific Grove, CA: Brooks/Cole.

Klein, D., Lewinsohn, P., Seeley, J., & Rohde, P. (2001). A family study of major depressive disorder in a community sample of adolescents. *Archives of General Psychiatry, 58(1),* 13-20.

Kline, R. (2005). *Principles and practice of structural equation modeling, 2nd Ed.* New York; London: The Guilford Press.

Kochanska, G. (1997). Multiple pathways to conscience for children with different temperaments: From toddlerhood to age 5. *Developmental Psychology*, 33, 228–240.

Kochanska, G., & Fowler, D. C. (2000). Temperament as a moderator of pathways to conscience in children: The contribution of electrodermal activity. *Psychophysiology*, 37, 788–795.

Kochanska, G., DeVret, K., Goldman, M., Murray, K., & Putnam, S. P. (1994). Maternal reports of conscience development and temperament in young children. *Child Development*, *65*, 852–868.

Kochanska, G. (1997). Children's temperament, mother's discipline, and security of attachment: Multiple pathways to emerging internalization. *Child Development*, 66, 597–615.

Kring and Werner, 2004 In *The Regulation of Emotion, 2004 Eds. Phillipot and Feldman* Lawrence Erlbaum Associations Publishers.

Kubic, M., Lytle, L., Birnbaum, A., Murray D., & Perry, C. (2003). Prevalence and correlates of depressive symptoms in young adolescents. *American Journal of Health Behavior, 27(*5), 546-553.

Landis, R. Edwards, B., & Cortina, J. (2008). On the practice of allowing correlated residuals among indicators in structural equation models. In *Statistical and Methodological Myths and Urban Legends: Doctrine, Verity and Fable in the Organizational and Social Sciences.* Eds. Charles E. Lance, Robert J. Vandenberg. CRC Press.

Lewinsohn, P., Clarke, G., Seeley, J., & Rohde, P. (1994). Major depression in community adolescents: Age at onset, episode duration, and time to recurrence. *Journal of the American Academy of Child & Adolescent Psychiatry, 33,* 809-818.

Lewinsohn, P., Rohde, P. Klein, D., & Seeley, J. (1999). Natural course of adolescent major depressive disorder: I. Continuity into young adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry, 38,* 56-63.

Lewinsohn, P., Rohde, P., & Seeley, J. (1995). Adolescent psychopathology III. Clinical consequences of comorbidity. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 510 – 519.

Lewinsohn, P., Rohde, P, Seeley, J, Klein, D., & Gotlib, I (2000). Natural course of adolescent major depressive disorder in a community sample: Predictors of recurrence in young adults. *American Journal of Psychiatry, 157,* 1584-1591.

Lewinsohn, P, Rohde, P, Seeley, J, Klein, D & Gotlib, I. (2003). Psychosocial functioning of young adults who have experienced and recovered from major depressive disorder during adolescence. *Journal of Abnormal Psychology, 12,* 353-363.

Lewinsohn, P., Roberts, R., Seeley, J., & Allen, N. (1997). The Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychology and Aging, 12,* 277-287.

Lewinsohn, P., Seeley, J., Solomon, A. & Zeiss, A. (2000). Clinical implications of "subthreshold" symptoms. *Journal of Abnormal Psychology. 109*(2), 345-351.

Levinson, D. (2005). The genetics of depression: A review. *Biological Psychiatry.* Retrieved online 5/12/2008 from http://www.stanford.edu/group/gbf/LevinsonGeneticsDepression.pdf

Lieb, R., Isensee, B., Hofler, M., Pfister, H, & Wittchen, H., (2006). Parental major depression and the risk of depression and other mental disorders in offspring: A prospective-longitudinal community study. *Archives of General Psychiatry, 59,* 365-374.

Lonigan, C. & Vasey, M. (2009). Negative affectivity, effortful control, and attention to threat-relevant stimuli. *Journal of Abnormal Child Psychology, 37(3),* 387-399.

Lovejoy, M., Graczyk, P., O'Hare, E., & Neuman, G. (2000). Maternal depression and parenting behavior: A meta-analytic review. *Clinical Psychology Review, 20(5),* 561-592.

MacCallum, R., Roznowksi, M.,& Necowitz, L. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Quantitative Methods in Psychology, 3*(3), 490 – 504.

Marmorstein, N, Malone, S, & Iacono, W. (2004). Psychiatric disorders among offspring of depressed mothers: Associations with paternal psychopathology. *American Journal of Psychiatry, 161*(9), 1588-1594.

Martin, A., & Cohen, D. (2000). Adolescent depression: A window of (missed) opportunity? *American Journal of Psychiatry, 157*(10), 1549-1551.

McGrath, J., Records, K., & Rice, M. (2007). Maternal depression and infant temperament characteristics. *Infant Behavioral Development, Published online 2007 August 21. doi: 10.1016/j.infbeh.2007.07.001.*

Merikangas, K., Dierker, L, & Szamari, P. (1998). Psychopathology among offspring of parents with substance abuse and/or anxiety disorders: A high risk study. *Journal of Child Psychology and Psychiatry, 95,* 711-720.

Merrell, K. (1999). Behavioral, social and emotional assessment of children and problems. Lawrence Erlbaum Associates. New Jersey.

Moore, G., & Calkins, S. (2004). Infants' vagal regulation in the still-face paradigm is related to dyadic coordination of mother-infant interaction, *Developmental Psychology, 40*(6), 1068-1080

Mondimore & Potash (2006). Familial aggregation of illness chronicity in recurrent, early-onset major depression pedigrees, *American Journal of Psychiatry, 163*(9), 1554 -1560.

Monroe, S. & Harkness , K. (2005). Life stress, the "kindling" hypothesis, and the recurrence of depression: Considerations from a life stress perspective. *Psychology Review, 112*, 417–45.

Moran, G. & Pederson, D. (1998) Proneness to distress and ambivalent relationships. *Infant Behavior and Development, 2(3),* 493-503.

Morris, A., & Silk, J., Steinberg, L., Myers, S, Robinson, L. (2007). The role of the family context in the development of emotion regulation. *Social Development, 16(2), 361-388.*

Mrazek, R., & Haggerty, R. (Eds.). (1994). *Reducing risks for mental disorders: Frontiers for preventive research.* Washington, DC: National Academy Press.

Mullins, L., Siegel, L., & Hodges, K. (1984). Cognitive problem-solving and life event correlates of depressive symptoms in children. *Journal of Abnormal Child Psychology, 13*(2), 305-314.

Muthén, L.K. and Muthén, B.O. (1998-2007). *Mplus user's guide*. Fifth Edition. Los Angeles, CA: Muthén & Muthén

Myung, I.J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology, 44,* 190-204.

Neville, H. (1995). Developmental specificity in neurocognitive development in humans. In: M. Gazzaniga (Ed.). *The cognitive neurosciences.* MIT Press, Cambridge, MA.

Nuyen, A.T. (1984). David Hume on Reason, Passions and Morals *Hume Studies* Volume X, 1, 26 - 45. Retrieved online 2/6/2009 from http://www.humesociety.org/hs/

Orvaschel, H., Walsh-Allis,G., & Weijai, Y. (1988). Psychopathology in children of parents with recurrent depression. *Journal of Abnormal Child Psychology, 16,* 17-28.

Paine, T. (1776). *The Crisis.* Retrieved on 4/07/2009 from http://www.uhistory.org/Paine/crisis/index.htm

Pascal , *B.* (2009). *Pensees* tr. Trotter, W. (Carnegie-Mellon) Retrieved online 4/9/2009 from http://books.mirror.org/gb.pascal.html.

Pauli-Pott, U., Mertesacker, B., Bade, U., Bauer, C., & Beckmann, D. (2000). Contexts of relations of infant negative emotionality to caregiver's reactivity/sensitivity. *Infant Behavior & Development, 23*(1), 23–29.

Pickens, J., & Field, T. (1995). Facial expressions and vagal tone of infants of depressed and non-depressed mothers. *Early Development and Parenting, 4(*2), 83-89.

Porges, S. (1992). Vagal tone: A physiological marker of stress vulnerability. *Pediatrics, 90,* 498-504.

Porges, S., Doussard-Roosevelt,J., Portales, L., & Suess, P. (1994). Cardiac vagal tone: Stability and relation to difficultness in infants and 3-year-olds. *Developmental Psychobiology, 27*(5), 289-300.

Porges, S., Doussard-Roosevelt, J., Portales, A., & Greenspan, S. (1996). Infant regulation of the vagal "brake" predicts child behavior problems: A psychobiological model of social behavior. *Developmental Psychobiology, 29*(8), 697-712.

Porges, S.W. (1996). Physiological regulation in high-risk infants: A model for assessment and potential intervention. *Development and Psychopathology,. 8,* 43-58.

Posner, M. & Rothbart, M. (2002). *Development and Psychopathology, 12*, 427–441. Cambridge University Press.

Post, R., Rubinow, D., Ballenger, J. (1986). Conditioning and sensitisation in the longitudinal course of affective illness. *British Journal of Psychiatry, 149*, 191–201.

Propper, C., & Moore, G. (2006). The influence of parenting on infant emotionality: A multi-level psychobiological perspective. *Developmental Review,* 427-460.

Provost, B., Crowe, T. & McClain, C. (2000). Concurrent validity of the Bayley Scales of Infant Development II Motor Scale and the Peabody Developmental Motor Scales in two-year-old children. *Physical & Occupational Therapeutic Pediatrics, 20*(1), *385-401.*

Putnam, S., Turk, N., & Stifter, C. (1998). Reactions to frustration and novelty in infancy: Physiological and behavioral consistencies and inconsistencies. *Infant Behavior and Development*, *21,* 44.

Radloff, L. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385-401.

Raudenbush, S. & Bryk, A. (2002). *Hierarchical linear models: Applications and analysis methods.* Newbury Park, CA: Sage Publications.

Roberts, R., Lewinsohn, P., & Seeley, J. (1991). Screening for adolescent depression: A comparison of depression scales. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*(1), 58-66.

Roberts R., Lewinsohn, P., & Seeley, J. (1995). Symptoms of DSM-III-R major depression in adolescence: evidence from an epidemiological survey. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*(12), 1608-17.

Rohde, P., Lewinsohn, P., & Seeley, J. (1991). Comorbidity of unipolar depression: II Comorbidity with other mental disorders in adolescents and adults. *Journal of Abnormal Psychology, 100,* 214-222.

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

Rothbart, M., Ahadi, S.,& Evans, D. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology*, *78*, 122–135.

Rothbart, M., & Bates, J.E. (2006). Temperament. In W. Damon, R. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 99– 166). New York: Wiley.

Rothbart, M., & Derryberry, D. (1981). Development of individual differences in temperament. In M.E. Lamb & A. Brown (Eds.), *Advances in developmental psychology* (Vol. 1, pp. 37–86). Hillsdale, NJ: Erlbaum.

Rothbart, M., Ellis, L., Rueda, M., & Posner, M. (2003). Developing mechanisms of temperamental effortful control. *Journal of Personality, 71(6),* 1113-1144.

Rothbart, M., & Hwang, J. (2003). Behavioral genetics studies of infant temperament: Findings vary across parent-report instruments. *Infant Behavior and Development, 26(1),* 112-114.

Rothbart, M., & Posner, M. (2006). Temperament, attention, and developmental psychopathology. In D. Cicchetti & D. Cohen (Eds.), *Developmental psychopathology: Vol. 2. Developmental Neuroscience* (2nd ed., pp. 465–501). New York: Wiley.

Rothbart, M.K., & Sheese, B.E. (2007). Temperament and emotion regulation. In J.J. Gross (Ed.), *Handbook of emotion regulation* (pp. 331–350). New York: Guilford.

Sadek, N. & Bona, J. (2000). Subsyndromal symptomatic depression: A new concept. *Depression and Anxiety, 12*(1), 30 – 39.

Santor, D. & Coyne, J. (1997). Shortening the CES-D to improve its ability to detect cases of depression. *Psychological Assessment,* American Psychological Association, Inc., *9*(3), 233-243.

Santucci, A., Silk, J., Shaw, D., Gentzler, A., Fox, N., & Kovacs, M. (2008) Vagal tone and temperament as predictors of emotion regulation strategies in young children. *Developmental Psychobiology, 50*, 205-216.

Saris, W. E., & Aalberts, C. (2003). Different explanations for correlated disturbance terms in MTMM studies. *Structural Equation Modeling*, 10(2), 193-213.

Schumacker, Randall E. and Richard G. Lomax (2004). *A beginner's guide to structural equation modeling, Second edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Seeley, J., Rohde, P., Lewinsohn, P., & Clarke, G. (2002) .*Depression in youth: Epidemiology, Identification, and Intervention.* In M.R. Shinn, H.M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches.* Bethesda, MD National Association of School Psychologists.

Sheeber, L. (2007). Personal communication October 18, 2007.

Siegal, J., Platt, J., & Peizer, S. (1976). Emotional and social real-life problem-solving thinking in adolescent and adult psychiatric patients. J*ournal of Clinical Psychology, 32*, 239-242.

Silk, J., Shaw, D., Forbes, E., Lane, T., & Kovacs, M. (2006) . Maternal depression and child internalizing: the moderating role of child emotion regulation. J*ournal of Clinical, Child, and Adolescent Psychology, 1,* 116-26.

Silk, J., Steinberg, L., & Morris, A. (2003). Adolescents' emotion regulation in daily life: Links to depressive symptoms and problem behavior. *Child Development, 74,* 1869 – 1880.

Silk, J. Vanderbilt, Adriance, Shaw, Forbes, Whalen, Ryan, Dahl (2008 in press).

Smolkowski, K. (2007). *Correlated errors in CFA and SEM models.* Unpublished manuscript.

Srivastava, S. (2009). *Measuring the Big Five Personality Factors.* Retrieved *5/26/2009* from http://www.uoregon.edu/~sanjay/bigfive.html.

Sroufe, L. (2000). Early relationships and the development of children. *Infant Mental Health Journal, 21*(1-2), 695-702.

Stevens, Joseph. (2009) Personal communication. May 13, 2009.

Steinhausen, H. & Metzke, C. (2000). Adolescent self-rated depressive symptoms in a Swiss epidemiologic study. *Journal of Youth and Adolescence, 29*(4), 427- 440.

Strelau, J. (1983). Temperament personality activity. London; New York: Academic Press.

Styron, W. (1990). *Darkness visible: Memoir of madness*. New York: Random House; pp. 56-7.

Substance Abuse and Mental Health Services Administration [SAMHSA] (2004). National Survey on Drug Use and Health. www.oas.samhsa.gov, http://www.mentalhealth.samhsa.gov/child/childhealth.asp.

Sullivan, P., Neale, M., & Kendler, K. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *American Journal of Psychiatry* 157:1552-1562.

Tabachnick, B. & Fidell, L. (2001). *Using multivariate statistics, Fourth Edition.* Needham Heights, MA: Allyn & Bacon.

Thomas, Chess & Birch (1968). *Temperament and behavior disorders in children.* New York, New York University Press

Thompson, R. A. (1994). Emotion regulation: A theme in search of a definition. In N. A. Fox (Ed.), *Monographs of the Society for Research in Child Development* (Serial No. 240 ed.,Vol.59, pp. 25–52). Chicago, IL: University of Chicago Press.

Tolstoy, L. (1875). *Anna Karenina* tr. by Constance Black Garnett (Barnes & Noble Classics Series) Edition: illustrated. Published by Spark Educational Publishing, 2003.

Trevarthen, C. (1994). The concept and foundation of infant intersubjectivity. *In: Intersubjective Communication and Emotion in Ontogeny: Between Nature, Nurture and Culture,* ed. S. Bråten, pp.3-5.

Tucker, L. & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1 – 10.

VanMeenen, K. & Wigfield, A. (2005). Doctoral dissertation. Brain electrical activity in infants of depressed and anxious mothers. *Unpublished dissertation.*

Verstraeten, K., Vasey, M., Raes, F., & Bijttebier, P. (2009 . Temperament and risk for depressive symptoms in adolescence: Mediation by rumination and moderation by effortful control. *Journal of Abnormal Child Psychology, 3*(3), 349-361.

Watson, D., Clark, L., & Tellegen, A. (1988). Development and validation of brief measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54,* 1063-1070.

Weissman, M., Warner, V. & Wickramaratne, P. (2005). Families at high and low risk for depression: A 3-generation study. *Archives of General Psychiatry, 62,* 29-36.

Weissman, M., Wickramaratne, P., Nomura, Y., Warner, V., Pilowsky, D, & Verdeli, H. (2006). Offspring of depressed parents: 20 years later. *American Journal of Psychiatry, 163,* 1001-1008,

Whiffen, V., & Gotlib, I. (1989). Infants of postpartum depressed mothers: Temperament and cognitive status. *Journal of Abnormal Psychology, 98,* 274-279.

Whittle, S., Allen, N., Lubman, D., & Yucel, M. (2006). The neurobiological basis of temperament: Towards a better understanding of psychopathology. *Neuroscience and Biobehavioral Reviews, 30,* 511 – 525.

World Health Organization (1998). Family environments of adolescents with lifetime depression: Associations with maternal depression history. *Journal of the American Academy of Child & Adolescent Psychiatry, 37*(11), 1152-1160.

World Health Organization, (2002). World Health Report 2002. *Reducing risk, promoting healthy life.* Geneva: WHO.

Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.* Unpublished doctoral dissertation, University of California, Los Angeles.

Zuckerman, M. (2005). *Psychobiology of personality*. Ed. Marvin Zuckerman, University of Delaware, Cambridge University Press.