2015

AN EXAMINATION OF THE PSYCHOMETRIC PROPERTIES OF THE WORKING MODEL OF THE CHILD INTERVIEW CODING SCHEME WITH BIOLOGICAL MOTHERS WHO HAVE MALTREATED

Brian D. Gustman
University of Kentucky, bdgustman@gmail.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

<div align="right">

Brian D. Gustman, Student

Dr. H. Thompson Prout, Major Professor

Dr. Kenneth Tyler, Director of Graduate Studies

</div>

AN EXAMINATION OF THE PSYCHOMETRIC PROPERTIES OF THE WORKING
MODEL OF THE CHILD INTERVIEW CODING SCHEME WITH BIOLOGICAL
MOTHERS WHO HAVE MALTREATED

_____

DISSERTATION
_____

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Education
at the University of Kentucky

By
Brian Daniel Gustman

Lexington, Kentucky

Director: Dr. H. Thompson Prout, Professor of School Psychology

Lexington, Kentucky

2015

ABSTRACT OF DISSERTATION

AN EXAMINATION OF THE PSYCHOMETRIC PROPERTIES OF THE WORKING
MODEL OF THE CHILD INTERVIEW CODING SCHEME WITH BIOLOGICAL
MOTHERS WHO HAVE MALTREATED

There are hundreds of thousands of children living in foster care in the United States on any given day. Mental health professionals may be called upon to assist with evaluating the parental capacity of these children's parents in order to inform reunification decisions. One of the key parental capacity domains to be evaluated is the relationship between parent and child (Schmidt et al., 2007). The Working Model of the Child Interview coding scheme (WMCI; Zeanah et al., 1996) is one tool for evaluating this relationship. There is a significant practice-to-research gap with this measure. To date, no peer-reviewed studies have established the WMCI coding scheme's psychometric properties; the need for such research with this measure has been cited in the literature (Sprang, Clark, & Bass, 2005).

In this dissertation, the literature was reviewed on attachment theory and internal working models of caregiving, the WMCI and other measures of similar constructs, the importance of establishing a measure's construct and criterion-validity, and forensic standards for measures. Then, it outlined a series of research questions aimed at exploring the psychometric properties of the WMCI Coding Scheme with mothers who have maltreated. Next, the methodology was described. Based on the results of the 403 biological mothers who completed the WMCI as part of court-appointed evaluations following child maltreatment, the items of the WMCI Coding Scheme were best conceptualized using two-factors: Quality and Content. This finding was consistent with the manual. Significant differences in item-level scores existed for all three WMCI descriptive classifications and for overall factor scores. Due to the instability of the two-item Content factor, this dissertation explored the use of a WMCI Total Score (combining the two factors) and using only the WMCI Quality factor and items. No significant relationships existed between any demographic characteristics and WMCI factor scores. Weak, negative correlations with other measures provided some evidence of convergent validity. In conclusion, some potential clinical/research implications for the WMCI Coding Scheme were made and limitations and future directions were described.

Brian D. Gustman

April 28, 2015

AN EXAMINATION OF THE PSYCHOMETRIC PROPERTIES OF THE WORKING
MODEL OF THE CHILD INTERVIEW CODING SCHEME WITH BIOLOGICAL
MOTHERS WHO HAVE MALTREATED

By

Brian Daniel Gustman

H. Thompson Prout, Ph.D.
Director of Dissertation

Kenneth Tyler, Ph.D.
Director of Graduate Studies

April 28, 2015

ACKNOWLEDGEMENTS

I owe many more long walks and tennis ball tosses to my dog, Wechsler. We started and finished this journey together. Your ever lolling tongue, snuggles and neuroses have kept me smiling.

Finally, thank you to my family. To my parents, Sheryl and Dan, I cannot thank you enough for all you have given and sacrificed for me over the years. You instilled in me the importance of education and pursuing a career that I would love. Dad, I have embraced your willingness to express my thoughts even when they may be unpopular. Mom, I have embraced your passion for helping others. To my sister, Kari, I cannot adequately express how central you have been and continue to be in my life and professional journeys. Thank you for your unfailing support of me. Your ability to help me hear and appreciate the harmonies around the melody of my life brings such richness. This one is also for you my little niece, Marian. I hope you dream great dreams for yourself and know I will support you my whole life. Thank you also to my grandparents, Aunt Janet, and cousins, Judy, Bill and Teresa. This accomplishment is for our family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One: Introduction and Literature Review

Childhood maltreatment is at epidemic levels across the United States, and beyond (Scott, Wolfe, & Wekerle, 2003; Wong et al., 2009). The U.S. Department of Health and Human Services (U.S. DHHS, 2010) reported that there were approximately 3.6 million reports of child maltreatment made during 2009. These reports represented approximately three million different children. Of those children, maltreatment was substantiated on 702,000 different children during the same year (U.S. DHHS, 2010). At its most severe, it was reported that some 1770 children were confirmed to have died as a direct result of their maltreatment during 2009, a rate that has steadily increased over the past 5 years. A recent study of the economic burden of child maltreatment from the number of substantiated maltreatment cases during 2008 revealed that each nonfatal child maltreatment victim averages a lifetime cost of $210,012 in government funds. For cases from 2008 alone, the total lifetime costs were approximately $124 billion (Fang, Brown, Florence, & Mercy, 2012).

General population studies reveal that actual rates of maltreatment are much higher than those substantiated by child protective service agencies. Briere and Elliott (2003) reported that approximately 37 percent of young adults in a nationally-stratified, random sample reported experiencing either child physical abuse, child sexual abuse, or both. May-Chahal and Cawson (2005) randomly sampled approximately 3,000 young adults in the United Kingdom and reported that approximately 17 percent of participants reported "concerning" to "severe" absences of care and 17 percent reported "intermediate" or "serious" lack of supervision.

Child maltreatment is a particularly devastating form of trauma because it includes a betrayal of the most basic trust: that children's primary needs for food, shelter, and safety/protection will be met. This interpersonal betrayal is particularly problematic for children because it coincides with the substantial and rapid developmental tasks of childhood (e.g., Browne & Winkelman, 2007; Cohen, Brown, & Smailes, 2001; De Bellis et al., 1999; Hildyard & Wolfe, 2002; Manly, Kim, Rogosch, & Cicchetti, 2001; van der Kolk, 2003).

Following ongoing or particularly severe child maltreatment, a relatively common practice in the United States is to remove the child from the care of the offending parent(s) and place the child in kinship care with relatives, or, when an appropriate kinship placement cannot be identified, foster care or similar placement (Budd, Clark, & Connell, 2011). Most often, parents are provided with a case plan that includes several actions the parents must take to address their caregiving risk before a family court judge considers the appropriateness of returning the children to their parent(s). Typically, the family court judge relies on a child welfare case manager, in addition to testimony by others, to decide if and when the child maltreatment risk of the parent(s) has been adequately addressed/reduced to the point that reunification of the family is in the best interest of the child.

In some instances, reunification may be successful (loosely defined as no new maltreatment referrals), but for many children reunification decisions result in additional maltreatment. Researchers have found that children returned to their caregivers following maltreatment are approximately three times more likely than the general population to experience additional maltreatment. Within two years following reunification

approximately one-fourth of children have been re-victimized by their caregivers and this figure increases to approximately one-third at three years post reunification (Connell et al., 2009). This recidivism rate appears to be relatively consistent across available studies, with another researcher reporting that approximately 37 percent of children who were reunified with their parents were the subject of additional maltreatment reports following their reunification (Johnson-Reid, 2003). Authors of both studies found that children were most likely to experience substantiated neglect following the reunification, which is consistent with the general child maltreatment literature wherein neglect is the most frequently substantiated finding (as opposed to physical, sexual or emotional abuse; Skowron & Woehrle, 2012). A history of multiple removals from caregiver(s) greatly increased recidivism risk. From the results of these studies, it is clear that current systems for making reunification decisions following removal for child maltreatment are inadequately protecting children from further maltreatment. These findings demand a better system to help child welfare agencies and family courts to inform decision-making regarding family reunification in order to better protect children.

Mental health professionals are ideally suited for consultation regarding reunification, particularly when there are questions regarding parental risks related to relationship skills, psychological functioning and substance abuse (Budd, Poindexter, Felix, & Naik-Polan, 2001). Unfortunately, mental health professionals frequently lack the experience and/or specialized assessment skills and tools necessary to make an informed expert opinion on these matters; this is further complicated by threats to professional objectivity if the mental health professional has been providing treatment services to the child(ren), parent(s), or family. Many mental health professionals have a

limited understanding of forensic psychology and standards of evidence (Budd et al., 2001). Schmidt, Cuttress, Lang, Lewandowski, and Rawana (2007), in their article on assessing the parent-child relationship for parenting capacity evaluations, discuss the problematic tendency of such evaluations to neglect the assessment of the relationship/attachment between the parent and child. Further, they emphasize the need for parenting capacity evaluations to include data related to the caregiver's own relationship history, mental health and treatment, current and prior child welfare involvement, levels of social support and motivation to change. They cite the need to include assessments that are both theoretically- and empirically-based due to the high-stakes of such assessments. Similarly, Sprang, Clark, Kaak, and Brenzel (2004) outlined the need for a multi-trait, multi-method approach to conducting comprehensive assessments for use by child welfare. They outlined five key domains to assess when determining parental capacity: "child factors, adult factors, relational factors, socioenvironmental factors, and maltreatment factors" (Sprang et al., 2004, p. 328). Both of these articles point to the need to consider the parent-child relationship as part of the evaluation process. One theoretical framework for conceptualizing this relationship is attachment theory. This dissertation will provide an overview of attachment theory, with particular attention to internal working models of parent-child relationships, translation of internal working models to other social science theories, an overview of the empirical literature on these internal working models, and a comprehensive review of the literature on the Working Model of the Child Interview (WMCI: Zeanah, Benoit, Barton, & Hirshberg, 1996). This background will then frame the proposed research questions and

hypotheses, methodology, results, discussion and directions for future research with the WMCI.

Attachment Theory

Bowlby (1958, 1973, 1977a, 1977b, 1988) was largely responsible for the development of attachment theory. Bowlby (1958) asserted that attachment is an evolutionary necessity for the physical and psychological survival and development of the child. Others have eloquently described the critical nature of attachment "as the psychological version of the immune system, designed to combat and reduce stress and fearful arousal just as the biological system combats physical disease" (Lyons-Ruth et al., 2004, as cited in Schmidt et al., 2007, p. 250). Bowlby (1958) identified the various instinctual responses of the infant that are designed to elicit a reciprocal parental response and thus meet the attachment need; these instinctual responses are: sucking, clinging, following, crying and smiling (Bowlby, 1958, p. 362). He asserted that these behaviors are more than actions or reflexes to gain access to physical, primary reinforcers, as was speculated by Freud and learning theorists; instead, he theorized these behaviors are used to meet not only these primary needs, but also social interaction and contact needs.

The assertion that social interaction and connection are also essential needs has been supported by many ethological studies. Results from some of these studies have demonstrated that nonhuman animal species engage in, and repeat, similar behaviors without contingent primary physical reinforcement; thus the proximity and behavioral responses of the caregivers must themselves be reinforcing. Bowlby's attachment theory also accounts for the need of all infants and children to have primary attachment figures, a concept inadequately accounted for by other theories devoid of relationship constructs.

Specifically, Bowlby asserts that during illness or distress a child shows a predictable preference for one caregiver over all others, regardless of whether they offer the same support. Bowlby (1958) concludes his work by outlining that it is when a primary attachment figure, most often the mother in Western cultures, consistently fails to reciprocate or haphazardly reciprocates the social response being sought by the child's attachment seeking behaviors that child psychological disturbances develop.

In later work, Bowlby (1977a, 1977b) further developed this clinical conclusion and asserted that psychotherapists of his day were approaching psychopathology inappropriately by focusing exclusively on the individual and essentially ignoring the person's early caregiving relationship experiences. Bowlby (1984) and Ainsworth (1969, 1979) asserted that it is through the consistent responsiveness of the parent to the child's physical and psychological needs that a child is able to form a secure attachment. This attachment then serves as the organizational framework for that child's social-emotional development (Bowlby, 1984). In its most simplistic form, Bowlby's construct distinguished attachment behavior from parenting behavior by defining the first as seeking and "obtaining protection" and the latter as "giving protection" (Bowlby, 1984, p. 14). Further, Bowlby asserted that both attachment and parenting behaviors of the dyad are prompted on a general level by biological programming, but the refinement of parenting and attachment behaviors is acquired through direct experience and observation of others. Although it is the early life attachments that most strongly influence individual psychological and relational development, attachments nonetheless occur throughout the lifespan and continue to shape the thoughts, feelings and behaviors of the individual.

From Bowlby's attachment theory perspective, psychopathology throughout the lifespan stems from, and results in, attachment and relationship difficulties.

Ainsworth (1979) further developed Bowlby's attachment theory in her empirical studies of infants and their mothers during the child's first year of life. She identified three general classifications of infants: securely attached, avoidant and ambivalent. Ainsworth noted that securely attached infants use their mothers as a "secure base" from which they explore their environment when the mother is present in the room (p. 932). Following a brief separation from their mothers these infants seek proximity and/or contact. Infants with ambivalent (insecure) attachments appeared anxious throughout the observation, but especially when the mother separates, and when the mother returned they simultaneously seek proximity but resist contact. The infants with avoidant (insecure) attachments tend not to appear distressed, unlike the other two groups of infants, when separated from their mother, and when the mother returns they tend to avoid and/or ignore her. Ainsworth stated that "there is a strong case to be made for differences in attachment quality being attributable to maternal behavior" (1979, p. 933). She also noted, however, that certain infants that may be born "difficult" and these infants are particularly at-risk for developing one of the insecure attachments.

Attachment theory applied to child maltreatment. When attachment theory is applied to child physical and emotional abuse, Bowlby asserted that these acts are "the distorted and exaggerated versions of behavior that is potentially functional" (Bowlby, 1984, p. 12). He condemned the parent's maltreatment behavior, and he asserted the empirically-supported adage that violence begets violence. Thus, he noted that many parents who were maltreated as children grow up and in turn revisit maltreatment on their

children. Of particular emphasis in his conceptualization of parents who maltreat, he noted they tend to be anxiously attached. When a person feels a significant relationship is being threatened they tend to first feel anxious and then respond with anger in an attempt to protect the relationship. In the most extreme form this may result in the physical assault of the child, but long preceding these acts are multiple verbal assaults and rejections of the child by the parent. Following these acts of maltreatment, children learn that they cannot trust their parent(s) to meet their needs, or that their needs will only be met haphazardly by their parents. Within this theoretical framework, such a response style results in an insecure attachment (Bowlby, 1984). This manifests differently for different children but their presentation often includes the same anxiety, anger and aggression that they have experienced from their caregivers when the child interacts with others. From this developmental and social learning perspective, throughout their lifespan children then recreate interpersonal relationship styles from their early lives. They may do this by engaging in romantic and caregiving/parenting relationships that include these same elements of anxiety and anger (Bowlby, 1984).

There is solid evidence that attachment theory is applicable to populations of children who have been maltreated and their parents. Cyr, Euser, Bakermans-Kranenburg, and van Ijzendoorn (2010) in their meta-analysis of attachment security among maltreated and high-risk families found that children who had been maltreated were significantly less likely to have secure attachments than were their non-maltreated comparisons. The effect size of $d = 2.10$ was large. The authors also examined studies of secure attachments among children who had a number of high-risk environmental factors (but who had not been maltreated), and found that these children were significantly more

likely to have insecure attachments than were children with low-risk factors; however, children who were maltreated were significantly more likely to have insecure attachments than were those from high-risk backgrounds (Cyr et al., 2010). Consistent with attachment theory, these results make sense: children who have been maltreated differ from other children with other risk factors when both groups experience stress; children who have been maltreated have experienced their attachment figure to be not only inconsistent in meeting their needs, but also actively harming or neglecting their needs. In another, smaller meta-analysis the researchers examined insecure attachment, as measured by the Strange Situation, among children who have been maltreated and comparison children; the authors found 36 percent of children who had not been maltreated had insecure attachment classifications, 80 percent of children who had been maltreated were classified as insecure in their attachments (Baer & Martinez, 2006).

Overall, attachment theory was a departure from other theories of its day in that it considered not only individuals, but also relationships between individuals. It clearly articulated that the foundational relationship for most individuals is the relationship that they develop with their primary caregiver. Bowlby (1988) posited that although attachment was partly biologically driven, it was also more complex, determined by children's experiences with the parents over time and the parents' own childhood history of receiving caregiving. On some level, the quality of the attachment relationship can be directly observed in parent-child interactional procedures (Ainsworth, 1979); however, Bowlby contended that the observed behavior of the parent and child was only the surface of the matter. To better understand why children display different types of

attachments to their parents, the internal processes that direct these behaviors must be understood: what he termed internal working models (Bowlby, 1984).

Internal working models. The internal working model of attachment is perhaps most simply defined this as the "representation of the self in relation to attachment" (Main, Kaplan, & Cassidy, 1985, p. 67). Essentially, the internal working model is the blueprint or organizing framework for relationships. Generally, the first working model that is developed by an individual is their representation of their primary attachment figure. When primary attachment figures are sensitive and responsive to infants, the infants come to trust that these individuals will meet their needs. With this secure base, infants and children have the confidence to tackle other developmental tasks knowing that they can return to their attachment figure for support. On the other hand, infants and children with inconsistent or unresponsive primary attachment figures develop insecurities because they cannot predict the behavior of their attachment figure (Ainsworth, 1979). It is this internal working model or cognitive conceptualization of the relationship on which children build their relationship with their caregivers and others. Cases where children have been able to establish secure internal working models explain why these children are generally resilient to isolated lapses in caregiving; these children have come to understand and trust that their parent(s) will meet their needs from earlier experiences. Understandably, it is the latter, unpredictable internal working model of attachment that is problematic to children as they go on to develop other relationships and attempt different developmental tasks.

Researchers have found that both members of the parent-child dyad develop internal working models of the relationship. As the child develops an internal working

model of her caregiver, the parent develops an internal working model of her child and of herself as a caregiver to her child (Zeanah, Keener, Anders, & Vieira-Baker, 1987). Throughout the rest of this paper this parental internal working model of the child and the parent-child relationship will be referred to as the IWM of caregiving. Much like the child's internal working model of attachment, the parent's IWM of caregiving is thought to be the result of that parent's own experiences being cared for during her/his childhood. The IWM of caregiving may be somewhat unique from other relationship working models, because the immediate focus of the relationship is not in what the child provides to the caregiver, but what the caregiver does to meet the needs of the child (George & Solomon, 1996). This IWM of caregiving serves as a cognitive framework for the parent's attachment-related thoughts, feelings and behaviors with her/his own child (Main et al., 1985; Mayseless, 2006). Some have suggested that when an individual is a child, the working model of the relationship is only amenable to alteration through direct experience (Main et al., 1985); however, the parent's IWM of caregiving is amenable to change as that parent has new caregiving experiences and engages in metacognitive processes (Collins & Read, 1994). It is to elucidate the IWM of caregiving Zeanah and colleagues went on to develop the Working Model of the Child Interview (WMCI: Zeanah & Benoit, 1995; Zeanah et al., 1996).

From my own cognitive-behavioral theoretical (CBT) orientation, the IWM of caregiving is readily assimilated. From a CBT framework, the IWM of caregiving is analogous to a parent's cognitive schemas about her child and her relationship with her child (McBride & Atkinson, 2009). Cognitive theory emphasizes the importance of early life experiences and interactions on the development of core beliefs and the automatic

11

thoughts that are subconsciously activated and largely predict our resulting affective

states and behaviors. From a related and integrative standpoint, internal working models

of attachment are well understood from a social information processing theory

perspective, too. Social information processing theory would support that individuals

with secure attachments and internal working models, also would be able to process all

forms of later attachment and relationship information (Dykas & Cassidy, 2011). Their

early positive life experiences would give them a positively-biased relationship schema.

Similarly, Dykas and Cassidy (2011) assert, from their review of studies, those

individuals with insecure attachments either resist processing attachment related

information, or interpret it through a negative relationship schema.

　　　These theories suggest that the IWM of caregiving, or cognitive schemas, become

relatively stable and resistant to change, and can result in self-fulfilling prophesies

leading the person to interact with others in ways that reinforce the original thoughts

about the child and the parent-child relationship (McBride & Atkinson, 2009; Zeanah &

Anders, 1987). In cognitive theory, schemas are dormant until triggered by specific life

events. In the case of the IWM of caregiving, these caregiving schemas have been

dormant until the parent learns of the pregnancy and active even more once the child is

born. A parent with a balanced IWM of caregiving, which developed from her own

experiences receiving care as a child, are theoretically connected to more positive affect

and responsive caregiving behavior. A parent with a nonbalanced IWM of caregiving is

more likely to recall pathological early life experiences of their own that interfere with

their ability to respond to their own child (Prather, 2007). These early life experiences

related to receiving care, serve as a cognitive framework for the parent's IWM of the

relationship with her child as conveyed through attachment-related thoughts, feelings and behaviors about the child (Main et al., 1985; Mayseless, 2006). This combination of attachment theory and cognitive theory principles has led to the development of cognitive-interpersonal therapeutic approaches (McBride & Atkinson, 2009).

Although it has largely fallen out of favor in research, transactional theory may also be useful in understanding IWMs of caregiving. Unlike attachment theory, which focuses on the unidirectional influences of the caregiver(s) on the child, transactional theory considers that both individuals in the dyad influence each other (Ciciolla, Gerstein, & Crnic, 2014). For instance, there may be certain characteristics of the parent and child that influence how the other member of the dyad engages with each other. These parent and child influences will be more specifically discussed and tested later; however, it is worth brief discussion here. Attachment theory readily explains how parent characteristics influence the IWM of caregiving for that parent (e.g. based on their own experiences with receiving care as children, traumatic life events, etc.). Transactional theory extends this to include the possibility that children do not all start life as a blank slate, but instead bring their own characteristics to the relationship. Children are born with different temperaments and some are also born with significant developmental delays that result in more social, emotional, and behavioral needs than for children who are typically developing. There has been considerable research on the relationship between early childhood developmental delays and caregiving behaviors. Generally, parents of children with developmental delays display less warmth and sensitivity in their parenting styles, possibly because the child does not respond in expected ways to warmth and sensitivity (e.g., Fenning, Baker, Baker, & Crnic, 2007). In the context of

transactional theory, it would be expected that parents IWMs of caregiving are not only shaped by their own attachment history, but also by the interactions with, and responses from, their children.

Aligned with a transactional theoretical approach to considering parent-child relationships, there has been some research conducted to examine the effects of clinical problems in both mothers and children and how the child attaches to her/his mother. Specifically, van IJzendoorn, Goldberg, Kroonenberg, and Frenkel (1992) conducted a meta-analysis of 34 clinical studies on attachment, utilizing the Strange Situation Procedure (Ainsworth, Blehar, Waters, & Wall, 1978, as cited in van IJzendoorn et al., 1992). The authors selected studies that examined maternal problems (e.g. child maltreatment perpetrators, mental illness, teen mothers, etc.) and/or child problems (e.g. prematurity, deafness, physical problems, autism, etc.). They then selected 21 nonclinical comparison samples to use as a basis for classification analyses. The results of this study were that child problems did not result in significant differences in attachment classifications as compared to nonclinical classifications. This result would suggest that child problems present from birth do not significantly alter a child's ability to form a secure attachment with his mother. The authors found that maternal problems did significantly alter the attachment classifications of their children, with significantly fewer children being classified as securely attached (van IJzendoorn et al., 1992). These results would at least partially support the emphasis of attachment theory on caregiving behaviors influencing attachment-related behaviors, and simultaneously call into question the bidirectional child-to-parent influence on the attachment classification.

Attachment, internal working models, and psychopathology. The relationship

between attachment theory and psychopathology, follows a model similar to cognitive-

behavioral conceptualizations. In the attachment theory framework, parents' internal

working models of their relationship with their child (parent-child relationship

cognitions) are substantially related to their parenting behaviors. These parenting

behaviors then influence the attachment styles of their children (children's thoughts about

their caregiver and predictability of caregiving), which then influence the behaviors of

the child (Bowlby, 1977a). In the extreme, the behaviors of the child become symptoms

of psychopathology.

A review of the literature supports this link between childhood attachment,

internal working models of caregiving, and later child and adolescent psychopathology.

Cohn, Cowan, Cowan, and Pearson (1992) examined these pathways and found that

among parents of preschool-aged children, those parents who had insecure internal

working models of attachment also were rated as less warm and less structured in their

interactions with their children. Additionally, these authors found that reciprocally the

children of parents with insecure internal working models of attachment were less warm

toward their parents during interactions. Another study went further to examine the

predictive validity of attachment to externalizing and internalizing problems in children

two years after the attachment classification procedure (Moss et al., 2006). The results of

this study revealed significant relationships between attachment classification and

externalizing and internalizing problems. Specifically, children with attachment

classifications that were insecure (ambivalent or controlling) had significantly higher

externalizing symptoms on a widely used measure of general child externalizing

15

symptoms. Similarly insecurely attached (controlling group only) children were rated as having more internalizing symptoms than securely attached children. The findings of this study are particularly useful because both internalizing and externalizing problems were rated not only by the child's mother, but also by the child's teacher (Moss et al., 2006). By including the teacher rating, the researchers were able to reduce the likelihood that the relationship between the attachment classification and the symptoms were solely attributable to the mother's perceptions of her child.

One form of attachment, the disorganized attachment, has been the focus of many research studies examining the relationship between attachment and developmental psychopathology. In a review of the literature on disorganized attachment, children were most often classified as disorganized when their mother's internal working model of attachment included unresolved losses or traumatic experiences (Green & Goldwyn, 2002). The literature further supports the association between parent-child interactive behaviors and disorganized attachment in the child. When mothers' interactions with their children have disrupted emotional communication, hostility, and intrusiveness, their children are more likely to have disorganized attachment (Lyons-Ruth, Bronfman, & Parsons, 1999). This disorganized attachment related to a number of different types of childhood and adolescent problems. Children who had an early childhood disorganized attachment have been found to be significantly more at risk for experiencing peer rejection, poorer self-regulation skills during adolescence, atypical classroom behaviors, and have higher levels of psychopathology than securely attached children (Green & Goldwyn, 2002).

The link between child attachment relationships and child and adolescent psychopathology has also been studied by others. One group of researchers found that when children and adolescents experience some negative parenting behaviors, those individuals with secure attachments were able to manage this without developing problem behaviors (Bosmans, Braet, Van Leeuwen, & Beyers, 2006). These same researchers found that children and adolescents with insecure attachments had significantly more problem behaviors across age groups. These findings support the importance of examining internal working models of caregiving, parent-child interactive behaviors, and infant/child attachment classifications. They support the theorized link between attachment and later difficulties during childhood and adolescence.

Criticisms and limitations of attachment theory. Attachment theory has been widely used to conceptualize the attachment difficulties of children who have experienced abuse and neglect. Researchers have employed it to help develop the framework for childhood onset psychiatric diagnoses such as Reactive Attachment Disorder, and parent-child relationship problems. Reactive Attachment Disorder has been widely criticized by many social scientist practitioners and researchers, which by extension calls into question the utility of attachment theory for diagnosis (Zeanah, 1996; Ziberstein, 2006).

There are also criticisms that attachment theory has not been particularly useful when directly translated from theory to intervention. Most attachment interventions have focused on altering caregiver behaviors to make them more consistent and responsive to the child; however, this does not address the problem that the child has already started to develop a relatively stable IWM of the caregiver. In cases where early life has involved

considerable instability, abuse, neglect, and lack of positive affect, simply changing the behaviors of the adult may not be enough to alter the child's IWM (Slater, 2007).

Perhaps the strongest criticism of attachment theory has been the emphasis on the mother-child relationship, often to the minimization or exclusion of the father-child relationship (Slater, 2007). Some early interpretations of attachment theory blamed mothers for their children's psychiatric disorders. In particular, mothers who worked outside the home were strongly criticized. This criticism went so far as to suggest that families who placed their children into daycare settings were damaging the parent-child attachment. Importantly, these generalizations were refuted by a longitudinal study of daycare conducted by the National Institute of Child Health and Human Development (Slater, 2007). This and other research on attachments in early childhood suggest that children are capable of developing multiple caregiver attachment relationships. A more current interpretation of attachment theory emphasizes the importance of these multiple caregiving relationships, rather than solely focusing on a primary attachment relationship (Etelson, 2007).

Empirical Studies of Internal Working Models

General studies of internal working models. Carlson, Sroufe, and Egeland (2004) had a number of significant findings in their longitudinal study of the predictive utility of early experience and relationship representations on later adolescent social-emotional behavior. They found that the direct relationship between early experiences, defined as attachment quality and life experiences through toddlerhood, and adolescent social functioning were not significantly related. They did find that measures of relationship representation correlated ($r = .36$) as strongly with adolescent social

18

functioning as did teacher social behavior ratings ($r = .33$). A significant limitation of this

research study was although the method of social behavior ratings from early childhood

through adolescence remained constant, relationship representation at each age of

assessment was assessed with different measures with different modes of responding.

Based on their results, it would appear that particularly problematic was a projective

family drawing and rating system designed by the authors which did not significantly

correlate the other of the other assessment measures (which did correlate with the other

internal working model measures). Thus, it would appear that this measure lacked

adequate validity with this sample.

In an exploration of the relationship of fathers' internal working models of

relationships to parenting behaviors, the authors found several significant findings

(Newland, Coyl, & Freeman, 2008). Specifically, they found that fathers' working

models of their relationship with their own mother and working models of their

relationship with their romantic partner predicted child attachment security. After testing

for the ways that fathers play with, discipline and co-parent as mediators of the

relationship of fathers' working models to child attachment security, working models

remained a significant predictor (Newland et al., 2008). Despite the lack of a fully

mediated model, the authors found that fathers' positive working models of their own

mother were significantly correlated with less severe discipline techniques (e.g.,

spanking) and their working model of their relationship with their partner significantly

correlated with parenting consistency and co-parenting behaviors (Newland et al., 2008).

Another study of working models of relationships that assessed both mothers' and

fathers' working models of their relationship with their own mothers attempted to

demonstrate the relationship between parents working models and their behavior with their own children (Macfie, McElwain, Houts, & Cox, 2005). Despite their somewhat small sample size, they found that there was in inter-generational transmission of role-reversal. Inconsistent with their initial hypotheses, they found that the role-reversal was specific to the gender of the child; mothers who were role-reversed with their mothers were role-reversed with their daughters, but this finding did not hold for mothers and their sons. Similarly, fathers who reported role reversal with their mothers predicted their partner's role-reversal with their sons, but not with their daughters (Macfie et al., 2005). This study was unique in its analysis of child gender specificity of aspects of parents working models and in that this and the aforementioned study by Newland et al. (2008) were the only studies examining fathers working models in relationship to their parenting or attachments with their children.

      Measures of internal working models. The measurement of the construct of internal working models of caregiving is the primary focus of this dissertation. One of the criticisms of attachment theory and especially internal working models has been on the measurement of these constructs (Zeanah & Anders, 1987). The difficulty of measuring something that by definition is internal has been part of the struggle to take the constructs from theory to research and practice. The importance of quantifying and classifying the internal working model of caregiving was the impetus behind the development of the WMCI and the WMCI Coding Scheme (Zeanah & Anders, 1987; Zeanah et al., 1996). This dissertation explored the psychometric properties of the WMCI, to provide empirical evidence on how well the WMCI Coding Scheme functioned as a measure of this construct. Several other measures of internal working models of caregiving, or

conceptually similar constructs were also in use at the time of this dissertation. Therefore, I reviewed some of these other measures before providing an overview of the WMCI and my reasoning for selecting it for this study.

Adult Attachment Interview. In the literature, one of the most frequently used measures of internal working models is the Adult Attachment Interview (AAI; George, Kaplan & Main, 1985). The AAI is a semi-structured interview that was designed to elicit the internal working models of adults. Specifically, it asks about the adult's memories of being parented, why the interviewee believes the parent(s) acted the way they did during the interviewee's childhood, how these have affected the interviewee throughout life, how the interviewee has handled losses and traumatic life experiences, and how the interviewee sees her/his own parenting or future parenting. After completion of the interview the narrative is coded according to the following elements: coherence (quality, quantity, relation, and manner). The coding system ultimately leads the interviewer to classify the adult attachment as autonomous/secure, dismissing, or preoccupied. The autonomous classification is one where the narrative is clear and consistent, and incorporates both positive and negative life experiences. Dismissing classifications are typically highly positive despite a lack of supporting positive experiences or ignoring/minimizing negative experiences. The AAI preoccupied classification is characterized by narratives that are confused or angry and the narrative frequently lacks focus or direction. Both dismissing and preoccupied classifications are considered insecure classifications (van IJzendoorn, 1995).

The psychometric properties of the AAI have been widely researched (Bakermans-Kranenburg & van IJzendoorn, 1993). Generally, these studies found that

AAI classifications were unrelated to intelligence, social desirability, and autobiographical memory of the parent. The AAI has demonstrated concurrent and predictive validity with regard to observed parent-child interaction and infant attachment (e.g., McFarland-Piazza, Hazen, Jacobvitz, & Boyd-Soisson, 2012; van IJzendoorn, 1995). In a meta-analytic study of the AAI and its predictive validity combined effect sizes of around 1.00 were found using a variety of criterion measures, which is a large effect size. The demographic composition of the samples included in the meta-analysis included mothers and fathers, adolescent mothers, mothers living in low SES, and mothers with psychiatric diagnoses. There have been relatively fewer studies examining the use of the AAI with fathers. One study that examined these father-child relationships, did demonstrate convergent validity. This study failed to report whether fathers' AAI scores from its coding scheme were similar to those of mothers, though the authors of the study collected AAI interviews from the spouses of the fathers in the study (McFarland-Piazza et al., 2012). Despite some of its limitations, due to the frequency of use of the AAI and the substantial literature supporting the reliability and validity of the measure, it has been described by some as the 'gold standard' of adult attachment measures (Rivas, Handler, & Sims, 2010).

Parent Development Interview. The Parent Development Interview (PDI: Aber, Slade, Berger, Bresgi, & Kaplan, 1985) is a 45-question interview that is designed to assess parents' mental representations of their relationship with their child (i.e., IWMs of caregiving). The interview takes between 1.5 – 2 hours to administer, due to secondary follow-up probes, when deemed necessary, in addition to the standard questions. The interview asks parents to describe their child, their relationship with their child (including

both positive and negative examples), and then describe themselves as parents (strengths and weaknesses). The interview is then coded along three dimensions: *parental representation of affective experience, child affective experience,* and *state-of-mind.* Multiple items comprise each dimension and are rated on scales with varying response options 9-point, 3-point, and 5-point Likert-type scales (Slade, Belsky, Aber, & Phelps, 1999).

In their exploratory factor analysis of the PDI, Slade et al. (1999) selected a 3-factor solution for the 16 PDI items. The factors were: *joy-pleasure/coherence, anger,* and *guilt-separation distress*. This three factor solution accounted for 51% of the variance in the original items, which is only minimally acceptable. Only items with factor loadings at .60 or higher were retained for use in this study, which resulted in six items being removed. The final EFA solution was a 3-factor solution with three items on each of the first two factors previously mentioned, and four items on the third factor (Slade et al., 1999). These researchers then went on to examine whether factor scores from the PDI were significantly related to AAI classifications; they found that two of the three factors did relate in the expected direction with AAI *Autonomous* classifications. Furthermore, they found significant relationships between PDI *joy-pleasure/coherence* and *anger* factor scores and mothering behaviors. Mother's with higher *joy-pleasure/coherence* scores had significantly more positive mothering behaviors during an observation and significantly fewer negative mothering behaviors than those with lower scores. Additionally, mothers with higher PDI *anger* factor scores had significantly fewer positive mothering behaviors (Slade et al., 1999). It is worth noting that this sample was relatively small for some of the analyses performed, with a total sample size of 125

mothers. Also of import, the study only included married, working- and middle-class Caucasian families.

Another more recent study examined the relationships between the two PDI factor scores, which were significant in the aforementioned study, and mother-child interactive behaviors (Dollberg, Feldman, & Keren, 2010). They found the strongest support for the PDI *joy-pleasure/coherence* factor, which significantly correlated positively with maternal sensitivity, maternal limit setting, and maternal assistance/support; this factor negatively correlated with maternal intrusiveness and child negative affect. The PDI *anger* factor only significantly correlated positively with the other PDI factor score, and maternal intrusiveness; no significant negative correlations were obtained for any of the other interactive behaviors (Dollberg et al., 2010).

Despite the rather limited published literature on the Parent Development Interview, there has been a published factor analysis on it and published guidance on how to use the PDI factor scores. It also provides an advantage over some other measures of internal working models in that it is specific to each parent-child dyad, and that it relies on the use of factor scores, rather than on broad classifications. Further research to abbreviate the interview process and identify additional items to replace those that have not loaded significantly in factor analytic studies may advance the utility of this measure from both a research and clinical standpoint.

Mental Representation of Caregiving Scale. Reizer and Mikulincer (2007) developed the Mental Representation of Caregiving Scale (MRC) as a self-report measure of internal working models thought to contribute to parental responsiveness. Their factor analysis of this 27-item measure produced a five factor scale: two factors

related to working model of self as caregiver – perceived ability to recognize others' need, perceived ability to provide effective help; one factor relating to working models of others – appraisal of others as worthy of help; and two factors related to motives for caregiving – egotistic motives for helping, and altruistic motives for helping (Reizer & Mikulincer, 2007). Follow-up one-way MANOVAs revealed that there were significant differences between fathers' and mothers' scores on the measure, with women reporting significantly more positive representations than men, and men scoring significantly higher on the egotistic motives for helping factor. Reizer and Mikulincer (2007) continued to psychometrically establish their scale by conducting studies of construct, convergent and divergent validity which generally supported the MRC measure. A significant limitation of this psychometric validation of the MRC scale was although the authors tested for differences between fathers' and mothers' working model of caregiving scores and found differences to exist, they failed to explore these differences in their studies related to convergent and discriminant validity. It could be hypothesized that perhaps the MRC measure is more valid of either mothers' or fathers' representations. Additional work is needed in this area to determine the meaning of maternal/paternal differences on the MRC.

Insightfulness Assessment. The Insightfulness Assessment (IA: Oppenheim & Koren-Karie, 2002) represents a relatively unique approach to assessing IWMs of caregiving, through a construct the authors identify as parental insightfulness. This measure involves first recording the parent-child dyad engaged in several short interactions with each other (e.g., free play, semi-structured cooperative play, etc.) based on the developmental age of the child. Then the parent is interviewed after watching the

video vignettes of the parent with her child. The clinician then codes the interview in a way modeled after the Working Model of the Child Interview. The interview transcript is then coded on 10 items with a 1 – 7 Likert-type scale. Items include: *Insight into child's motives, Openness, Complexity in description of child, Maintenance of focus on child, Richness of description of child, Coherence of thought, Acceptance, Anger, Worry,* and *Separateness from child.* These scores are then used to classify the parent's narrative according to four categories: *Positively Insightful, One-Sided, Disengaged,* or *Mixed.* The authors note that the categorical classification is more than a simple summation of scores from the 10 items, but rather categories are assigned based on specific constellations of item scores.

To date, the published literature on the Insightfulness Assessment is relatively limited. In their original article, introducing the IA, the authors cite three previous studies that examined the utility of this measure. They reported that these studies demonstrated significant relationships between the IA categories and child attachment classifications, with the exception of the *Disengaged* classification (Koren-Karie, Oppenheim, Dolev, Sher, & Etzion-Carasso, 2002; Oppenheim, Koren-Karie, & Sagi, 2001, as cited in Oppenheim & Koren-Karie, 2002). The third study was a paper presentation where the authors of the measure examined maternal insightfulness pre- and post-treatment for preschool children with clinical problems (Oppenheim, Goldsmith & Koren-Karie, 2001, as cited in Oppenheim & Koren-Karie, 2002). Results of this study demonstrated improvements in maternal insightfulness which correlated with improved child behavior.

In one of the few other published studies utilizing the Insightfulness Assessment, the relationships between IA categories and item scores were used to explore their

relationship to child attachment security and maternal sensitivity among clinically depressed and non-depressed mothers (Ramsauer et al., 2014). Their results demonstrated that clinically depressed mothers differed significantly from non-depressed mothers in their IA categories and on all IA items with the exception of the *Richness of description of child* item. The findings regarding IA categories/item scores and attachment security were mixed; however, IA categories were significantly related to maternal sensitivity, accounting for 70% of the variance with mothers with clinical depression, and 51% of the variance with mothers who were non-depressed (Ramsauer et al., 2014). The available empirical literature on the Insightfulness Assessment is very limited, and as such should be used with considerable qualification of these limitations or used only descriptively rather than diagnostically.

   Parenting Stress Index. A very different measure of parent-child relationship is the Parenting Stress Index (PSI; Abidin, 1990). The PSI is a 120-item self-report questionnaire designed to assess parenting stress in the context of the parent's relationship with one of the parent's children. The measure provides three different stress scores. The first two are stress scores related to perceived characteristics of the child and characteristics of the parent, respectively. The third is a measure of situational stressors, in an attempt to discriminate between parenting and life stress (Doll, 1989). The PSI has been manualized and includes hand-scoring procedures or a computer-scored program can be utilized. In either case, the raw scores on the PSI are converted to percentiles. The manual suggests that a Total Stress Score (comprised of the items on the Parenting Stress and Difficult Child factors) above the 90th percentile indicate at-risk parent-child relationship difficulties and warrant referral for intervention (Abidin, 1995). The manual

27

indicates that 90 percent of parents who score at or above this cutoff are true positive cases of parents needing intervention; however, the cutoff misses approximately 25 percent of abusive parents. In fact, some research suggests that abusive parents have somewhat unique profiles on the PSI, with either extremely elevated scores or very low scores on the Total Stress Index (Abidin, 1990).

The PSI was normed on a sample of 2,633 parents of typically developing children and children with behavior or medical problems. The norming sample included children ranging in age from 0 years, 3 months to 13 years, 0 months, with the majority of the sample under the age of 5 years-old (Heinze & Grisso, 1996). A limitation of this measure is that the normative sample was predominantly Caucasian (95%) and married. There are separate norms for Hispanic, Dutch, Portuguese, Italian and Chinese individuals, though it should be noted that these have been developed in countries where these are the dominant language, rather than within the United States (Abidin, 1990). Some studies have attempted to extend the PSI to use with ethnically diverse populations in the United States. Though these studies are somewhat limited, scores on the PSI with mothers who are African-American are similar to that of the norming sample, scores with mothers who are Hispanic were more elevated (Heinze & Grisso, 1996). At the time of their review of the PSI in 1996, Heinze and Grisso (1996) found approximately 200 published studies that utilized the PSI for a variety of uses. Their conclusion was that the PSI was an acceptable measure for the use of screening parental capacity for court evaluations, with the limitation that it has only been established with female caregivers and that results should be interpreted with caution without more in-depth clinical interviews (Heinze & Grisso, 1996).

Working Model of the Child Interview

The Working Model of the Child Interview (WMCI: Zeanah et al., 1996) was designed to measure internal working models of caregiving and will be the focus of this dissertation. The WMCI has been widely utilized in clinical and research applications with parents and their infants and young children since its creation in the late 1980s and early 1990s. The WMCI is an open-ended question, semi-structured interview designed to quantify and qualify/categorize the internal working models of caregiving that parents have of their young children. Unlike some other measures, the WMCI is about a specific parent-child dyad relationship, rather than globally about the parents' internal working models of caregiving. From an attachment theory perspective, these internal working models of caregiving comprise half of the information needed to fully understand the parent-child relationship, with the other half being those behaviors that are directly observable between the parent and child (Zeanah et al., 1997). The WMCI is both a semi-structured interview and a coding scheme for assessing the parent's overall internal working model of caregiving, and the elements of that internal working model of caregiving. A more complete description of this measure is provided in the Method chapter of this dissertation.

WMCI as compared to other measures. Although there are multiple measures of internal working models of caregiving available, I selected the WMCI for this dissertation study for a number of reasons. The first reason is the considerable literature base for the WMCI. It is a measure that has been used in many research studies for approximately the past 30 years. A review of many of these studies will follow. Another reason the WMCI was selected is because it was developed from the Adult Attachment

Interview, which, as previously cited, is the most well-established measure of internal

working models. The WMCI's specificity to the internal working model of caregiving for

a specific parent-child dyad, unlike the AAI, allows for more specific examination of that

one relationship, rather than general internal working models of caregiving. Some other

measures are also specific to parents' internal working models of caregiving to a specific

child; however, these measures have been considerably less used in peer-reviewed

journal articles. When comparing the WMCI to other measure of internal working models

of caregiving, notably lacking were psychometric studies evaluating the WMCI coding

scheme. Such studies are an important step prior to using the WMCI or any other

measure to make clinical decisions. From a scientist-practitioner framework, the dearth of

empirical studies detailing the development of the WMCI coding scheme and related

narrative classifications, was concerning. This dissertation reviewed many studies that

addressed the reliability and/or validity of the WMCI, and then outlined an attempt to

begin to address some of the lingering psychometric gaps.

Reliability studies. Essential to the usefulness of any measure, is that it reliably,

or consistently, measures the construct of interest. There are many studies that have

focused on the reliability of the WMCI. Theran, Levendosky, Bogt, and Huth-Bocks

(2005) examined the test-retest reliability of the WMCI among a sample of women. They

administered the WMCI to women during their last trimester of pregnancy and then

readministered the measure approximately one year later. They oversampled for women

whom had been in domestically violent relationships. All of the WMCI coders in the

interview were trained on the coding procedure and had at least 80 percent inter-rater

reliability before coding the procedures for analysis. In this risk sample, at the

administration when the child was approximately a year old, 59 percent of women had balanced representations, 21 percent had disengaged representations and 20 percent had distorted representations. They found that the most reliable classification across administrations was the balanced representation, with 79 percent of mothers maintaining the balanced classification. The next most stable classification was the disengaged representation with 48 percent remaining classified as disengaged, 37 percent of women who initially were evaluated as distorted remained distorted. The authors also examined stability by collapsing classifications into balanced and non-balanced representations. When they did so they found that overall stability of the measure was 71 percent. Balanced classifications were significantly more stable than non-balanced classifications, 79 percent for the former and 62 percent for the latter (Theran et al., 2005).

In another study of the test-retest stability of the WMCI, the researchers found that prenatal maternal WMCI balanced classifications were 89 percent concordant with WMCI classifications of those same mothers 12 months later (Benoit, Parker, & Zeanah, 1997). A similarly high level of stability was observed for WMCI distorted classifications, with 85 percent remaining distorted. The stability for the WMCI disengaged classification was not stable with only 13 percent remaining disengaged (Benoit et al., 1997). In this sample, one year postnatal maternal WMCI classifications were 65 percent balanced, 3 percent disengaged, and 32 percent distorted.

Validity studies. One of the early validity studies of the WMCI was conducted by Benoit and colleagues (Benoit, Zeanah, Parker, Nicholson, & Coolbear, 1997). In their study they collapsed the results from three different studies into one dataset for analysis to create a sample of 54 mothers of children with clinical problems and 45 mothers of

children without clinical problems. Children were between the ages of less than one month old and 67 months old. Overall, inter-rater agreement was 76 percent. Mothers of children with clinical problems were significantly more likely to have nonbalanced representations than were mothers of children without clinical problems. They found that only 9 percent of mothers of children with clinical problems had balanced representations, and 38 percent of mothers of children without clinical problems had this type of representation. Specific to the items on the WMCI coding scheme, the authors found that there were significant differences between mothers with children with clinical problems and those without clinical problems on all of the quality and content scales, with the exception of infant/child difficulty. Had the authors applied a Bonferroni correction to their analysis, the items of fear for safety and coherence would have also been non-significant. The authors of this study reported that the WMCI appeared sensitive to detecting mother-child relationships with clinical problems, but that it lacked specificity. The cross-sectional design of this study is somewhat limiting in that it is unable to situate the WMCI representation and the child's clinical problems temporally. Thus, it is impossible to determine if, consistent with attachment theory, the representation leads to the development of child problems or if there are bidirectional interactions as would be suggested under transactional models. Another significant limitation of this study is that it collapsed participants from three different studies, conducted at different times, in two different countries, and with qualitatively different clinical problems – ranging from sleep disorders to failure-to-thrive.

In another study, the researchers examined the predictive validity of WMCI by examining the relationship between prenatal representations and maternal parenting

behaviors when the child was approximately one year old (Dayton, Levendosky, Davidson, & Bogat, 2010). In this study, the authors found that mothers with balanced representations displayed significantly more positive parenting behaviors, showing joy, warmth, and sensitivity, during their interactions with their child did mothers with distorted or disengaged representations. Mothers who were classified as distorted engaged in significantly more hostile interactions, such as sarcasm, mocking/teasing the child, than did mothers with balanced or disengaged representations. Additionally, mothers with disengaged representations significantly differed from either mothers with balanced or distorted representations in their use of controlling parenting, which was the extent to which the mother blocked instead of partnered with the child to achieve the child's goal during the playtime.

When considering parent-child attachment, attachment theory emphasizes the importance of both mental representations of the child (IWMs) and parent-child interactions. Huth-Bocks, Theran, Levendosky, and Bogat (2011) examined the relationship between IWMs, as measured with the WMCI, and infant-mother attachments, as measured by the Strange Situation procedure. The results of the study were initially non-significant when concordance rates were examined between balanced/secure, disengaged/avoidant, and distorted/ambivalent classifications, respectively. The relationships were statistically significant when WMCI classifications were dichotomized into balanced and non-balanced (combining disengaged and distorted), and Strange Situation classifications were dichotomized into secure and insecure (collapsing avoidant and ambivalent classifications).

Crawford and Benoit (2009) created an alternative coding scheme for the WMCI in response to the development and refinement of a disorganized attachment classification, which occurred after the WMCI was designed. The WMCI-Disrupted was developed by adding to the coding system items similar to those from the AMBIANCE (Lyons-Ruth, Bronfman, & Parsons, 1999, as cited in Crawford & Benoit, 2009). Crawford and Benoit (2009) found that the prenatal WMCI-Disrupted score, as scored by raters blind to all other assessments, correlated strongly with the same parent's AMBIANCE score obtained between 12 and 18 months of age, ($r = .637$). Essentially, they demonstrated that the constructs measured by the AMBIANCE observational tool could be tapped reliably at the representational level using the WMCI-Disrupted coding scheme. Of additional interest, the WMCI-Disrupted coding scheme would appear to represent a different construct and/or classification system than the traditional WMCI coding scheme as none of the traditional classifications (balanced, disengaged, or distorted) were significantly more associated with the disrupted classification.

WMCI with diverse caregivers. A significant limitation of much of the literature on the psychometric properties of measures, including those of the WMCI, is the reliance on convenience samples and samples that are predominated by participants from the majority culture in the United States and these are thought to generalize to other cultures; a limitation that Minde, Minde, and Vogel (2006) attempted to address in their examination of attachment representations in South Africa. They administered the WMCI to 46 mothers from a black township. The measures were administered in the primary language of the mothers by individuals who had grown up in the area and who spoke the language fluently. The author coded the interview transcripts using the standard scoring

criteria of the WMCI. Approximately six months later, the same author coded the interview transcripts using the region-specific scoring criteria developed by five individuals who were 'experts' of the culture. When compared to the Attachment Q-Sort, a culturally-neutral observational assessment (van IJzendoorn & Sagi, 1999, as cited by Minde et al., 2006), the classifications based on the standard WMCI criteria did not significantly correlate. When the culture-specific WMCI criteria were utilized the representation classifications were significantly correlated with Attachment Q-Sort classifications. The results of this study would suggest that the WMCI scoring criteria may have a significant cultural-loading, which may affect the validity of the measure when used with individuals from different cultures. Additional studies should be conducted to determine the extent of the cultural-loading and the need for alternative scoring criteria for different cultural groups.

Minde et al. (2006) attempted to explore the utility of the WMCI coding scheme with a cultural group outside the United States or Western Europe, others have explored its utility with culturally diverse participants within the United States or Western Europe. For the most part, these studies have only minimally empirically explored the potential influence of culturally diverse groups on the coding of the WMCI. For example, Rosenblum, McDonough, Sameroff, and Muzik (2008) described their participants in the methods of their article, but failed to report any preliminary analyses to examine whether there were differences between the cultural groups represented in the study. Their study included primarily European Americans, but 16 percent of the sample identified as African American, and 5 percent identified as Asian, Latino, biracial, or "other." The omission of analyses to determine if there were significant differences between WMCI

scores is interesting, considering such analyses were reported for other demographic variables (adolescent mothers vs. adult mothers, and educational attainment). Similarly, Theran et al. (2005) provided self-reported ethnicity of their sample, which included primarily participants who were White/Caucasian (63%), Black/African-American (24%), Hispanic/Latina (5%), biracial (4%), with other ethnicities at or below two percent of the sample. Though the sample included ethnic diversity there were no reported preliminary or primary analyses to check for differences in WMCI classifications by ethnicity. Dayton et al. (2010) also had an ethnically diverse sample with 62 percent Caucasian, 26 percent African American and 12 percent multiracial/Latina/other persons of color. As with the other studies, they too did not examine or at least report any analyses to explore differences in WMCI coding scores or classifications. Given the highly verbal demands of producing a narrative for the WMCI, such an examination would be particularly important if the sample included individuals who though fluent in English may have a different primary language. Furthermore, the failure to examine the possibility of measurement invariance across ethnicity essentially treats the results as colorblind, when oral traditions, communication styles, and other factors associated with ethnicity may have a significant relationship to WMCI scores and classifications.

In a study of African-American mothers' IWMs of their relationship with one of their children, the authors found that 38 percent were classified as balanced, 36 percent were disengaged and 26 percent were distorted (Sokolowski, Hans, Bernstein, & Cox, 2007). This study found that mothers with disengaged representation classifications were significantly less responsive to their children, used less encouragement, and were more

withdrawn during a parent-child interaction than were mothers classified as either

balanced or distorted. These classification rates, particularly in the balanced and

disengaged classifications, are substantially different than the distribution of WMCI

classifications based on a systematic review of WMCI articles (Vreeswijk, Janneke,

Maas, & van Bakel, 2012). It is worth noting that the sample used by Sokolowski et al.

(2007) was not only different from many other studies using the WMCI in terms of

ethnicity, but also in terms of poverty-rates, educational attainment, exposure to

community violence, and living in a community with one of the largest public housing

projects in the nation.

Similar to the Sokolowski et al. (2007) study, Schechter et al. (2005) also

explored the use of the WMCI with an inner-city sample. This sample differed in

important ways, in that the participants were predominantly Hispanic American (88% of

the sample) with the remaining participants identifying as African American.

Additionally, this was defined as a clinical sample based on self-reported risk indicators

such as filing for restraining orders, history of child protective services investigation,

history of acting violently, and history of suicide attempts. In this sample WMCI

classifications were as follows: balanced – 17 percent, disengaged – 24 percent, and

distorted – 59 percent (Schechter et al., 2005).

To date, the published research on the use of the WMCI coding scheme and

WMCI classifications with mothers from ethnically diverse backgrounds represents a

significant limitation. Although many studies have included mothers from multiple

ethnicities, none of those available reported whether there were significant differences in

WMCI outcomes based on ethnicity. One study was available for review that focused

37

exclusively on mothers who identified as African American, and no attention was given to how the WMCI performed with this sample. Similarly, the only study with a predominantly Hispanic American sample did not compare WMCI outcomes to those from other studies; additionally, that study was limited by focusing on a clinically-referred convenience sample, suggesting limited generalizability of its findings. Focused research on measurement invariance across ethnicities is needed to determine if the WMCI coding scheme and classifications are reliable and valid for diverse populations.

    WMCI and other maternal characteristics. In addition to ethnic differences potentially influencing the overall classification on WMCIs, there is some evidence to support that other caregiver demographic variables also influence the quality of the narrative and related coding. Rosenblum et al. (2008) determined from their preliminary analyses, that adolescent mothers could not be included in their overall analysis because their scores on the WMCI were significant outliers when compared to the other mothers in the sample. Adolescent mothers were found to score on the very low end of the coding scheme; however, due to the small subset of adolescent mothers in the sample additional analyses to attempt to understand why they represented outliers was not conducted. Possibly related to why adolescent mothers were outliers, Rosenblum et al. (2008) found that higher maternal educational attainment was moderately, positively correlated with higher WMCI scores; however, after controlling for the effect of educational attainment, there was still a significant relationship between WMCI scores and caregiving behaviors. Other researchers have documented a similar relationship between maternal educational level and WMCI classification (Sokolowski et al., 2007).

Others have examined how a number of other maternal characteristics related to WMCI classifications. Huth-Bocks et al. (2011) found in their primary analyses that WMCI classifications of mothers in their sample significantly related to family income, with higher income predicting balanced classifications. Again, this finding is inconsistent within the literature, as Sokolowski et al. (2007) found no significant relationship between income and WMCI classification. Huth-Bocks et al. (2011) also found a main-effect for multiple-parent households predicting balanced classifications, though the effect size of this variable was small. Other researchers, in their preliminary analyses, found no significant relationship between WMCI classification and any of the demographic variables they collected, which included maternal age, maternal education, number of children in the home, severity of trauma exposure, or degree of social support (Schechter et al., 2005). It is worth noting that the lack of significance of some of these variables may be due to limited power due to relatively small sample size, or possible range restriction on some of these variables.

WMCI and mothers with psychopathology or trauma. There have been some studies attempting to identify mothers who are more likely than the general population to have non-balanced WMCI representations. One such study examined the attachment representations of mothers who were experiencing postnatal depression (Wood, Hargreaves, & Marks, 2004). The authors of this article found that among mothers with depression only approximately 25 percent were classified as having a balanced representation, compared to general population rates of around 60 percent for this same classification. The majority of mothers with depression were classified as distorted (50 percent). Another study designed to examine WMCI classifications and the influence of

social/contextual factors, had different findings (Huth-Bocks et al., 2011). This study also explored the possible relationship between maternal depression symptoms and WMCI classifications and found no significant relationship. Similarly, Sokolowski et al. (2007) found no significant difference in WMCI classification by maternal depression or anxiety symptoms. Whether these differences are due to the use of balanced/nonbalanced classifications as opposed to the standard three classification system used in the study by Wood et al. (2004) study cannot be determined without analyses beyond those provided by the authors.

In addition to testing for possible relationships between maternal depression and WMCI outcomes, studies have also focused on trauma-exposure and Posttraumatic Stress Disorder (PTSD) symptomology on WMCI classifications. The results of these studies indicate that it is not merely the exposure to trauma, but rather the severity of PTSD symptoms that meaningfully relates to WMCI classifications (Huth-Bocks et al., 2011; Schechter et al., 2005). Schechter et al. (2005) found that severity of trauma exposure did not predict WMCI classification; however, severity of PTSD symptoms did significantly predict a WMCI distorted classification. In another study, the researchers found a significant relationship between elevated Hostility symptom scores on the Brief Symptom Inventory (BSI) and WMCI classifications of either distorted or disengaged (Sokolowski et al., 2007). The results of these studies are somewhat mixed, but point to the importance of conducting preliminary data analyses to determine what, if any, relationship exists between maternal characteristics and WMCI scores and/or classifications.

WMCI and children with non-typical development. The stability of IWMs as measured using the WMCI has also been empirically examined. In a prospective study conducted in the Netherlands, Hall et al. (2014) examined the relationships between maternal IWMs, caregiving behaviors and infant attachment. Specifically, they were interested in the influence of preterm birth on IWMs. At approximately 6 months following birth, they conducted WMCIs with samples of mothers of full-term and preterm infants. They examined caregiving interactions at the same time, and then again at 24 months post-partum. Also, at 24 months post-partum, they examined infant attachment. Their analyses controlled for the significant medical differences between the full-term and pre-term infants, as well as for the significant maternal demographics between the two groups. The result of their study was a fully mediated model, where observed caregiving interactions accounted for 62.7 percent of the effect of maternal IWMs on infant attachment (Hall et al., 2014). Contrary to the initial hypothesis that there would be significant differences between mothers of full-term versus pre-term infants, no significant differences were found in the structural equation model by group type. The authors of this study did find that there was a strong relationship between overall IWM classification and infant attachment at 24 months. Of the mothers with non-disrupted IWMs, only 15 percent had an infant with an insecure attachment relationship; mothers with disrupted IWMs were much more likely to have an infant with an insecure attachment, 47 percent were classified as insecure (Hall et al., 2014). The results of this study demonstrated that among this sample, there was no significant difference in the relationship between maternal IWMs of mothers of full-term versus pre-term infants. It

did demonstrate that non-disrupted maternal IWMs are a strong predictor of the young child developing a secure caregiving attachment.

A significant limitation of this study is that it utilized a different coding scheme than the one developed by the original authors of the WMCI. These researchers utilized dichotomous classifications of disrupted versus non-disrupted IWMs according to an alternate coding scheme that was more recently developed and less utilized in the published literature (Crawford & Benoit, 2009). This is particularly problematic given the relatively low percentage of IWMs that were classified as disrupted, representing approximately 20 percent of both full-term and pre-term maternal IWMs. Therefore, it is unclear to what extent the results of this study are comparable to studies of the WMCI using the traditional coding scheme.

Another study explored WMCI classifications of mothers of children diagnosed with failure to thrive (FTT) and a comparison sample of mothers of children with typical development (Coolbear & Benoit, 1999). The results of this study were significant differences in WMCI classifications between the two groups. Of the mothers of children with FTT only 14 percent had balanced classifications versus 55 percent of mothers of children who were developing typically. Whether this difference is due to a child characteristic should be interpreted with extreme caution, since FTT is thought to be a medical problem that is developed as a result of both organic and environmental/relational difficulties. Specifically, FTT has been linked to neglectful parental care and disturbed relationship between the infant and parent (Coolbear & Benoit, 1999).

Summary of WMCI literature. A thorough review of the empirical literature on the WMCI revealed a number of ways to improve upon future research with this measure. The most apparent difficulty with the available literature on the WMCI, is the considerable variation in the coding schemes employed and subsequently the types of classifications made from the interview. Some studies have used a more recently developed coding scheme that results in dichotomous classifications of disrupted or non-disrupted classifications (e.g., Crawford & Benoit, 2009; Hall et al., 2014). Another study, utilized a completely different coding system, designed to measure parental reflective functioning, a distinct concept from the parental IWMs the WMCI was designed to measure (Rosenblum et al., 2008). Still other researchers utilized a dichotomous classification: balanced versus nonbalanced, keeping the original balanced classification from the WMCI coding manual and combining the disengaged and distorted classifications into the nonbalanced classification (Theran et al., 2005). The authors noted that this balanced/nonbalanced classification system was utilized because of the greater test-retest stability. Huth-Bocks et al. (2011) utilized this same balanced/nonbalanced classification system, after their initial results were nonsignificant for the three classifications from the WMCI coding manual.

Establishing Reliability and Validity of a Measure

In the preceding sections of this dissertation, I provided a review of studies exploring and testing the reliability and validity of the WMCI. Considering the types of decisions that are often made with measures of social science constructs, considerable attention must be given to these topics. The WMCI is no exception. Psychologists, clinical social workers, and other qualified mental health professionals, are bound by

their professional ethics to use measures for the purposes for which they were designed and demonstrate adequate reliability and validity (American Psychological Association, 2010, Standard 9.02). They also need to consider the limits of the measure. This is the purpose of conducting and publishing validation studies. When researchers conduct validation studies it is not really the assessment tool that is being validated, but rather the use of the measure with the sample and the population from which that sample was selected. Although it is beyond the scope of this dissertation, it is sufficient to emphasize that it is essential to always consider the reliability of the measurement prior to conducting validation studies (Nardi, 2006). If the measurement is consistent, then it is up to social scientists to review the available literature on the measure to determine if its use is valid for use with their particular population.

Validation studies are concerned with ensuring that a measure actually measures what it claims. In other words, do the items on the measure and the method of measurement accurately operationalize the construct of interest? Researchers typically discuss multiple forms of validity, it is this overarching idea of determining if the measure captures what it is designed to that is common to all of the forms; thus, some have asserted that there really then is only one type of validity – construct validity (Nunnally & Bernstein, 1994). Nonetheless, it is still important to examine the different types of criterion-related validity that largely comprise validation studies in the social sciences. In the sections that follow, I will provide an overview of three types of criterion-related validity that were utilized in this dissertation: concurrent validity, convergent validity, and discriminant validity. An additional type of criterion-related

validity, predictive validity, will not be discussed since this type of validity can only be examined in longitudinal studies and the design of this dissertation is cross-sectional.

Concurrent Validity. Concurrent validity is a type of criterion-related validity that is examined at, or around, the same time that the researcher collects data using the measure to be validated (Nardi, 2006). Concurrent validity is concerned with the ability of the measure to distinguish between people based on some external standard that is similar to the construct being measured. For example, with the WMCI, there are several ways to examine the concurrent validity of the WMCI. One of the most important contributions to the establishment of the concurrent validity of the WMCI, is to examine the relationship between the WMCI coding scheme scores and the WMCI classification made by the clinician/researcher. This comparison is under the umbrella of both construct and criterion-related validity. It is given that the WMCI classification is not independent of the WMCI coding scheme scores, since the classifications are not directly derived from the scores, but rather are based on clinician judgment, this comparison is important.

Another method to examine the concurrent validity would be to compare the classifications of the WMCI to the classifications on the AAI. If there was a high degree of correspondence between the classifications, then this would provide evidence of concurrent validity, since Zeanah et al. (1996) reported that they based the WMCI on the AAI. Of the literature reviewed, only two studies have utilized both the AAI and WMCI with maternal caregivers; however, one study failed to report the correspondence rates between the classifications obtained from each procedure (Coolbear & Benoit, 1999). The other study, supported the distinction that the AAI is a general measure of working models of relationships, and that the WMCI is a specific measure of the working model

of a parent-child relationship. The authors found that the WMCI-Disrupted classification fully mediated the relationship between the AAI and Strange Situation classifications (Crawford & Benoit, 2009). The paucity of studies directly comparing these two measures is likely a result of the very time consuming nature of conducting both interviews with participants, when they are essentially designed to capture the same construct; however, the establishment of other measures such as intelligence tests is equally, if not more time consuming, but is routinely done to validate new intelligence measures (Nunnally & Bernstein, 1994). Similarly to the time and expense invested in developing intelligence measures, similar expenditures are warranted for measures of parent-child relationship since like measures of intelligence, these measures may be used to make life-altering decisions for children.

  Convergent Validity. A very closely related type of criterion-related validity is convergent validity. Convergent validity is concerned with the relationship between the measured construct and those constructs that are closely-related to the measured construct (Trochim, 2006). One way of examining the convergent validity of the WMCI coding scheme is to compare it to other measures of constructs similar to that of the IWM of caregiving measured by the WMCI.

  With regard to the WMCI, the majority of the research on the convergent validity of the WMCI has focused on WMCI classifications and child attachment classifications made through parent-child behavioral observation procedures (e.g., Strange Situation, Crowell Procedure, Marshak Interaction Method). Based on attachment theory, parents IWM of their child should be related to the behavioral interaction between the parent and child. This relationship has been explored by some researchers who have found

concordance between WMCI classifications and parent-child interactions (e.g., Benoit et al., 1997; Forcada-Guex, Borghini, Pierrehumbert, Ansermet, & Muller-Nix, 2011). In their review of WMCI studies, Vreeswijk et al. (2012), note that the relationship between maternal IWMs and parent-child behavioral observations has been well-established among infants; there is insufficient research on this relationship with samples including older children. Additionally, one of the most overlooked tests of validity with the WMCI is the relationship between the quantitative scores given on the WMCI coding scheme and the WMCI classifications (balanced, distorted, disengaged or balanced/nonbalanced). From a thorough review of the published literature on the WMCI, it appears there are no published studies examining this most basic relationship. Such analysis may provide rough guidelines for using WMCI scores to distinguish between the different WMCI classifications. In combination, concurrent and convergent validity lend significant credibility to the validation of a social science measure.

Discriminant Validity. The other side of convergent validity is discriminant validity. Convergent validity is concerned with establishing that the measure captures the construct of interest, and discriminant validity is concerned with establishing that the measure is not diluted by constructs that it should not be measuring (Trochim, 2006). Again employing the example of the WMCI, a review of the literature on the WMCI would suggest that WMCI scores should not be strongly related to measures of maternal educational level, maternal psychiatric symptoms, or child psychiatric status. If there are non-significant or weak relationships between WMCI scores and these other measures, then it is possible to be more confident that the WMCI is measuring IWMs of the child rather than these other factors which may also influence caregiving. In the literature,

47

discriminant validity is often used rather liberally, with little regard for the discriminant criterion measure. I would assert that a more meaningful test of discriminant validity is to select criterion that are somehow, though distally related to the overall topic of interest. Such selectivity allows for greater confidence that the intended construct is being measured. Again, in reference to the WMCI, there is some evidence that psychiatric status does have a meaningful relationship to parenting behaviors (e.g., Benoit et al., 1997; Wood et al., 2004), but it is essential that the WMCI does not simply measure maternal or child psychiatric status. It must contribute something significantly different from these other constructs in order to be useful.

Problems with Measure Validation. In addition to the consideration of the multiple forms of construct and criterion-related validity, there are problems inherent in establishing the validity of a measure. As previously mentioned, each validation study is really a validation for the specific use of the measure with the specific research sample (Nunnally & Bernstein, 1994). Thus, each new use of a measure with a new population requires some degree of validation to ensure that the measure still measures the construct of interest accurately. One method of examining the validity of a measure across populations and uses is through the use of meta-analyses and measure reviews. One such systematic review was available in the literature for the WMCI (Vreeswijk et al., 2012). That review, with a number of cautions, noted that the WMCI is a "valid and useful clinical and research tool" for examining mothers' IWMs of the child and parenting behaviors (Vreeswijk et al., 2012, p. 326).

Other concerns related to criterion-related validity involve the criterion selected, range restriction, temporal relationship of predictor to criterion measure, and determining

48

adequate strength of the relationship (Nunnally & Bernstein, 1994). With respect to each of these concerns, there is rarely uniform agreement between researchers about the best measures to which to compare the measure of interest. This limitation can be overcome by multiple researchers conducting multiple studies and using different, but conceptually-related criterion measures. To some extent, this has been accomplished with the WMCI; however, most of the research has focused on establishing criterion-validity with other measures that are closely theoretically-related. These findings would be extended by establishing convergent validity with a variety of measures that measure parent perceptions of their child. A further concern with conducting research to establish a measure is the need to ensure that the sampling framework includes enough variation to avoid the problem of range restriction. This may particularly be a problem with the WMCI because there are a relatively limited number of items on the coding scheme, yet few of the studies have gone beyond examining the broad classifications obtained from the WMCI coding scheme to examine the Likert-type ratings from the coding scheme with the exception of the study by Sprang, Clark, and Bass (2005). The final limitation of conducting validity studies is determining how strong the relationship needs to be between the measure and selected criterion, or on the discriminant side how weak the relationship needs to be to provide evidence of criterion-related validity. On this topic there is relatively limited guidance, due to the number of variables that influence the significance of the relationship. Factors such as sampling framework, time between measurement of the construct of interest and the predictor criterion, individual factors, and measurement error to name only a few (Nunnally & Bernstein, 1994). For these

reasons, even modest correlations between a measure and the reference criterion, may be worthy of interpretation and support the validity of the measure.

Standards for Forensic Measures

Due to the significance of decisions made with forensic assessments, particularly in cases of forensic parenting capacity evaluations, it is essential that social scientists are providing the courts data from measures with established reliability and validity. Given that such data and testimony may be used as evidence for reunifying children with caregivers who formerly maltreated them or for pursuing the termination of parental rights, it is especially important that measures meet rigorous requirements for use. When measures have questionable psychometric properties and are used for clinical and/or forensic decision-making, then it potentially places the clients at significant risk. Potentially, clinicians may be inadvertently making recommendations based on pseudo-science in such cases. These well-intended recommendations may lead to inappropriate decisions that contribute to the well-established problem of repeated child maltreatment after reunification. Alternately, such recommendations may steer judges to terminate parental rights, thus permanently severing the parent-child relationship, when not necessary. Such implications necessitate the ongoing psychometric evaluation of social science measures, to ensure appropriate use and decision-making. Regardless of how well-established a measure may be, it is essential, as with other types of assessment, that interpretations, conclusions and recommendations are informed by a convergence of data sources rather than on any one piece of data.

The literature has outlined various criteria to aid in the determination if a measure has acceptable reliability and validity for forensic use to minimize the

misapplication/misinterpretation of measures. A number of different criteria for acceptable forensic measures have been published in this area (e.g., Otto, Edens, & Barcus, 2000). From a psycho-legal perspective, the case law established by *Daubert v. Merrell Dow Pharmaceuticals, Inc.* upheld by the U.S. Supreme Court established the *Daubert* standard for testimonial admissibility (Yanez & Fremouw, 2004). The *Daubert* standard consists of: (1) the technique must be assessed as helpful by reviewing the history of the techniques use with the population and validity of the technique; (2) the technique has been subjected to peer review and generally found favorable; (3) the technique has a known potential error rate and there are standards for the use of the technique; and (4) the measure is generally accepted by the scientific community. These standards are inter-related to some extent; however, the failure of a technique or measure to meet all of these standards means it does not have the necessary properties to be admitted into testimony.

At the current time, one would necessarily conclude that the WMCI coding scheme does not currently meet the *Daubert* standard. One could convincingly argue that the WMCI coding scheme does currently meet the last criterion, in that it is widely used by early childhood mental health professionals. The WMCI representation classifications of balanced or unbalanced, or balanced, distorted or disengaged would seem to meet the criteria for the first standard as it relates to concurrent and predictive validity and inter-rater reliability. Considerably fewer publications have documented the reliability or validity of the WMCI coding scheme and there are not published factor analyses of the WMCI coding scheme to establish its psychometric properties. The third criteria for the WMCI coding scheme is marginally met through the publication of a relatively loosely

defined administration and interpretation manual; however the manual does not include technical information regarding the structure, reliability or validity. The authors of the measure have attempted to standardize the training for the measure by providing training in the administration and scoring of the WMCI to professionals and requiring them to reach acceptable inter-rater agreement on training interviews.

In their article reviewing parent-child relationship assessment tools that possess acceptable empirical support for forensic use, Schmidt et al. (2007) cite the WMCI as having "established" stability and concurrent validity; however, they fail to specify that, at best, only the global classification of balanced versus nonbalanced would seem to meet such empirical criteria. In their article providing empirical support for the current assessment protocol, Sprang et al. (2005) cite that the WMCI coding scheme has not been psychometrically established in the existing literature. They cited the need for future studies to establish the WMCI's psychometric properties and factor structure. The purpose of the present study is to begin to address this gap in the literature.

Research Questions

The primary research question to be answered by this study is: Do the eight quality/content items of the WMCI coding scheme constitute a psychometrically reliable and valid measure of the IWM of caregiving? The following questions/hypotheses will be tested to answer this:

1. What is the relationship between maternal and/or child demographic variables and WMCI item and total scores?

2. Do certain WMCI item scores accurately discriminate between WMCI descriptive classifications of balanced, distorted, or disengaged?

3. What factor structure of the eight content/quality items of the WMCI coding scheme best represents these items empirically and conceptually/theoretically? Does the factor structure of the WMCI coding scheme possess acceptable reliability?

4. Are WMCI descriptive classifications meaningfully predicted by mean score differences in factor score(s) on the WMCI coding scheme? It is hypothesized that WMCI factor scores will be highest for balanced classifications, and lowest for disengaged classifications, with distorted classification scores between these.

5. Do the WMCI factor scores demonstrate convergent validity with the CAPI and PSI-SF?

   a. It is hypothesized that WMCI factor scores will have significant, negative correlations with CAPI scores.

   b. It is hypothesized that WMCI factor scores will have significant negative correlations with PSI-SF scores.

6. Do the WMCI factor scores demonstrate discriminant validity with the demographic variables of the sample?

Chapter Two: Method

Data Source and Sampling Framework

The participants in this study were part of a larger ongoing study of individuals who were evaluated at a speciality program designed to provide family court-appointed evaluations to develop recommendations for families who have open maltreatment cases. The dataset included data on approximately 1500 children and their caregivers. The majority of cases included the child's data and those of the biological mother, and a minority of cases included biological father's data in addition to the biological mother's or singly. A smaller number of cases did not include either biological parent, but instead included relative caregivers or adoptive parents who had maltreated. All participants included in this analysis provided their informed consent to have their de-identified data used for research at the time of their evaluation, the protocol was approved by both the University of Kentucky Internal Review Board, and the relevant state Internal Review Board.

The speciality program evaluates families who have open, substantiated cases of abuse or neglect and takes referrals from the courts and state's child welfare agency. Certain criteria must be met before a referral is accepted: (a) the children cannot have been seen at the clinic before; (b) the children must have a medical card (with rare exceptions); (c) parent(s) cannot be testing positive for illicit substances; (d) parents must be visiting regularly with their children; and (e) parent(s) must be assessed as no or minimal risk to the program clinicians (e.g., no current orders of protection between parents, no recent history of violence toward others). The majority of the families

evaluated were from within 75 miles of the specialty program; however, cases came from anywhere in the state.

The evaluations provided by the specialty program are multidisciplinary, multidimensional, and multi-method. Empirical support for the assessment protocol and measures is published in the professional literature (Sprang et al., 2005). The typical assessment is completed over the course of 6-8 hours of direct assessment at the clinic. The assessment protocol includes: individual interviews with each family member (for young children this is a play evaluation), observations of each parent with each of the children, a family interaction observation, WMCI interviews with each parent about each child, foster parent interviews, and completion of psychometrics by parents (self-report, relational and child-report), the children (self-report) and foster parents (child-report). Psychometrics and procedures utilized vary somewhat by referral questions, clinical assessment of child and parent distress or fatigue, and parental cognitive functioning or reading ability.

Parents could deny the use of their data for research purposes without consequence and no information regarding consent for research was included in the evaluation report or feedback to the referring child welfare agency or the courts. Those parents who were under the age of 18 provided their own consent to participate in the evaluation and consent was also secured from their legal guardian, if either denied consent, then that parent's data were not included in the database.

The research assistants who completed the data report form from the family evaluation assigned a case number to each child, with children from the same family receiving consecutive identification numbers in the database. The log of case numbers to

child names is maintained in a log book which is kept in a locked cabinet at the program in a location separate from the SPSS database. The original court evaluation report is kept in accordance with HIPPA privacy and confidentiality guidelines, and in accordance with state and University of Kentucky policies. Data coding and entry accuracy was checked at random by having a second member of the study personnel code and enter the same case to check for errors.

Parents ranged in age from 14 to 72 years-old at the time of the evaluation. The majority of parents in the database were mothers, or heterosexual couples, with a minority of same-sex couples, or cases with fathers only being evaluated. Reflective of the population of Kentucky, the majority of participants were White/Caucasian.

For the purposes of this dissertation only biological mothers of children 12 years-old, or younger, at the time of the family assessment are included in these analyses. Most of studies of the WMCI have focused on mothers of infants, and to a lesser extent children under the age of 5 years-old; however, I decided to include children through age 12 for a couple of reasons: (1) the WMCI is routinely used as part of the program assessment protocol for children of all ages; (2) including children up through this age maximized the sample size for analyses including the PSI-SF, which is normed for children up to this age. Although the majority of the sample was Caucasian, all ethnicities were included since preliminary analyses indicated no significant difference in WMCI scores by dichotomized ethnicity (white or person of color).

In cases where a biological mother had more than one child included in the database, then only one of her mother-child dyads were included in the sample. Decisions about which mother-child dyad to include were based on a number of considerations: (1)

cases where the mother-child dyad had not been administered the WMCI, or WMCI item scores were missing, were eliminated; (2) cases where there was no substantiated maltreatment by the mother with the specific child were eliminated; (3) since the majority of cases assessed were with younger children, older children were over-selected to balance this initial age bias.

Biological fathers, other biological primary caregivers, step-parents, foster parents, adoptive parents, and non-biologically related same-sex partners were excluded from analysis because of the lack of published WMCI research including them.

Based on these inclusionary and exclusionary criteria, and data cleaning described in the results section, 403 different mothers who had maltreated were included in the sample for this dissertation. Table 1 presents a summary of demographic variables of the mothers in this sample. Table 2 presents a summary of demographic variables of the children in this sample. The total sample was utilized for all analyses, with the exception of the factor analyses. For the purposes of the exploratory factor analysis and confirmatory factor analysis, the sample was randomly split into two groups, using SPSS's Select Cases, Random sample of cases.

Instrumentation

   Demographic variables. I also extracted a number of demographic variables from the program database to allow for accurate description and analysis of the sample. For mothers these variables included: age, race/ethnicity, mental health symptom severity, substance-related disorder, and type(s) of maltreatment with which the mother had been charged. For both age and race/ethnicity, maternal self-reports at the time of the evaluation were utilized. For maternal substance-related disorder, welfare caseworker

Table 1

Demographic Composition of Mothers in the Sample

Race/Ethnicity	
White	87%
Person of Color	13%
African American	(11%)
Multiethnic	(1.5%)
Hispanic	(0.5%)
200% or more Below Poverty	
Yes	92%
No	8%
Employment	
Unemployed	62%
Full-time	25%
Part-time	12%
Maternal Age	
Mdn	27 years-old
M (SD)	28.31 (7.15)
Years of Education	
Mdn	11 years
M (SD)	11.03 (2.04)

N = 403.

Table 2

Demographic Composition of Children in the Sample

Race/Ethnicity	
White	74.4%
Person of Color	25.3%
Multiethnic	(13.9%)
African American	(10.2%)
Other Ethnicity	(1.0%)
Hispanic	(0.2%)
Child Sex	
Male	56%
Female	44%
Child Age	
Mdn	5 years-old
M (SD)	5.37 (3.56)
Child Placement	
Foster/Adoptive Care	59.1%
Kinship	19.1%
Biological Parent(s)	15.4%
Residential Treatment	5.0%
Group Home	1.2%
Months Out-of-Home	
Mdn	9 months
M (SD)	11.16 (9.00)

N = 403.

report, mental health or substance-abuse provider report, and/or self-report were utilized to determine the presence or absence of a substance-related disorder. A substance-related disorder was considered present if any of these sources confirmed it. Type of maltreatment charge was based on child welfare caseworker reports and court documents and will be classified as: neglect (physical and/or medical), and abuse (physical and/or sexual with, or without, neglect charges). This dichotomy was made due to the smaller sample size in all categories other than neglect.

For mental health symptom severity, the results of the Brief Symptom Inventory (Derogatis, 1993) were utilized for mothers. The BSI is a 53-item self-report inventory on which individuals use a 5-point Likert-type rating for various psychiatric symptoms within the past week. The BSI validity index was utilized to remove cases that do not meet the validity criteria for the measure. The BSI Global Severity Index (GSI) was selected for use in this study based on a review of the literature which supports that the BSI is an appropriate measure of general psychological distress, but has questionable reliability with disorder-specific subscales (Skeem et al., 2006). A review of the BSI supported that the GSI is the most internally consistent index from this measure, with internal consistencies across studies ranging from .90 - .97 (Mohammadkhani, Dobson, Amiri, & Hosseini Ghafari, 2010).

Child demographic variables extracted included: age, race, mental health diagnosis, and length of time the child has been living in foster/kinship care. Age reported was the child's chronological age in years at the time of the family evaluation. Race was dichotomously coded, Caucasian or Person of Color, this was done due to the

largely homogenous racial composition of the state and of those referred for evaluations

at the speciality program. Child mental health diagnosis was based on child welfare

caseworker report and/or report by the child's therapist (if child has been seeing a

therapist) and was coded as either positive for any mental health diagnosis or negative

for a mental health diagnosis. Length of time living in foster/kinship care was rounded to

the nearest month at the time of the evaluation; it should be noted that this variable

represents only the amount of time the child has lived out-of-home since the most recent

removal. In this sample, some children had prior removals, but this information is not

included in the existing dataset.

Working Model of the Child Interview. The primary instrument for this study

was the Working Model of the Child Interview Coding Scheme (WMCI: Zeanah et al.,

1996). It is a semi-structured interview and coding scheme developed to measure parents'

internal working model of her child and relationship to her child. The interview is

designed to have parents "reveal as much as possible in a narrative account of their

perceptions, feelings, motives and interpretations of a particular child and their

relationship to that child" (Zeanah et al., 1996, Introduction section, para. 2).

The WMCI (Zeanah et al., 1996) interview procedure takes between 45 – 90

minutes to complete with an individual parent about one of the parent's children. A

shorter version of the WMCI interview is also acceptable if by clinician judgment they

have enough information on which to make a determination. At the program, an

abbreviated version of the WMCI was developed, which typically takes approximately 20

– 30 minutes to complete. This abbreviated version of the WMCI was used at the

discretion of the evaluation team leader; reasons for using this version were related to

reducing the likelihood of fatigue during the evaluation (particularly if the parent(s) had numerous children) or reducing time/economic burden (e.g., using the abbreviated version to prevent the parents from needing to return to the clinic on another day particularly when they had traveled some distance). A recent study, utilizing the speciality project data, demonstrated that clinicians scored abbreviated WMCIs similarly to full WMCIs on all of the content/quality items of the measure (Gustman & Sprang, working manuscript).

The probes of the WMCI cover a number of different content areas including: thoughts about their child's early development, personality, concerns and hopes for their children, favorite memories, thoughts about when the parent learned she was going to have this child, etc. It also includes probes about how the parent responds to the distressed child, setbacks with the child, and worries for her child. The probes are designed to elicit a comprehensive narrative about the parent and her thoughts and feelings about her relationship with the child and about the child himself. At some point, all of the items are answered during the interview; however, the interviewer is trained to go along with the flow of the narrative and to add follow-up probes that encourage elaboration to more fully understand the narrative. The interviews are typically audio and/or video recorded for the clinician to reference for the coding/scoring of the narrative (Rosenblum, Zeanah, McDonough, & Muzik, 2004).

Following the conclusion of the interview, the clinician completes the WMCI Coding Scheme (Zeanah et al., 1996). These items are named: Richness of Perceptions, Openness to Change, Intensity of Involvement, Coherence, Acceptance, Caregiving Sensitivity, Infant Difficulty (for accuracy in this study renamed Child Difficulty), and

Fear for Safety (Zeanah et al., 1996). The WMCI Coding Manual includes all but two of these items, Infant/Child Difficulty and Fear for Safety, as items representing the quality of the mother-child relationship. These other two items are designed to measure the content of the interview. Each feature/item is assigned a Likert-type rating ranging from 1 – 5 (1 = *None*, 2 = *Limited*, 3 = *Moderate*, 4 = *Considerable*, and 5 = *Extreme*). The coding manual provides descriptive guidelines for each point on the scale in reference to the particular feature being rated. For most of the features a higher rating represents a more balanced narrative, with the exception of the two content items, which are reverse scored. Clinicians also complete ratings of the affective tone of the interview; however, these were not included in the present study because the affective items have been less consistently described in the published literature on this measure, with some reports on just four affective items and others with seven affective items. Unlike the qualitative/content items of the WMCI, these items would not be reasonably expected to load into a meaningful factor/scale. Additionally, the affective items have not been found to be significant predictors in other studies and it is likely that such items would have potentially strong gender and cultural loadings.

After rating the qualitative/content items of the WMCI, the clinician is then prompted to assign an overall representation classification to the interview. For the purposes of this study only the three primary classification types were utilized. In the manual each overall classification also has between 2 – 4 subtypes within each classification. The manual does not specify how the eight quality/content item scores figure into the representation classification. It instead provides guidelines about what features should be present for each classification both in terms of features of the narrative

about the child and the caregiving role (Zeanah et al., 1996). The authors do specify that *balanced* representations are associated with higher scores on at least some of the qualitative/content items. This classification is characterized by understanding both the positive and challenging aspects of the child and the parent's relationship with the child. The narrative conveys a sense that the parent is invested in the relationship, understands the child as a developing and changing individual, and has some richness of detail. *Distorted* representations are associated with inconsistencies within the narrative, feeling overwhelmed by the child, not viewing the child as an individual, and one of several types of distortion (Zeanah et al. 1996, distorted representations, para. 1). Often the caregiver seems preoccupied or distracted by other concerns other than the child. She may view herself as a victim of the child's behaviors and neediness, and be disappointed that the child is not meeting her needs. The narrative also lacks an understanding of how the mother's behavior has an influence on the child and be overly self-concerned (Zeanah et al., 1996.) *Disengaged* representations are associated with lack of emotional/personal involvement with the child and relationship, lack of flexibility to accommodate changes, and incoherence (Zeanah et al., 1996, disengaged representations, para. 1). Consistent with the research by others, the representation classifications will also be examined as *balanced* or *nonbalanced* (collapsing the distorted and disengaged classifications into the latter classification: e.g., Theran et al., 2005).

Child Abuse Potential Inventory. The Child Abuse Potential Inventory (CAPI: Milner, 1986), is a 160-item self-report measure designed to identify parents who are at-risk for or alleged of physically abusing their children. The Abuse scale is comprised of 77 items and according to the manual has internal consistencies of .92 for control groups

and .95 for abuse groups. The author recommended a 215-point cutoff, with scores at or above this level being described as potential abusers. The overall classification rate using this cutoff is 73.8 percent correct for parents who have physically abused, and 99.1 percent for control group parents. Correct classification rates drop substantially to 28 percent correct with a mixed sample of sexually abusive and physically abusive parents (Holden, Willis, & Foltz, 1996). A study by Ayoub and Milner (as cited by Yanez & Fremouw, 2004) suggested there is a moderate relationship between parents who neglect and CAPI scores. The measure includes three validity scales: faking good, faking bad, and random response. An invalidating score on any of these scales will result in the protocol being excluded from the analysis. A review of parental capacity measures found the CAPI was the instrument that scored highest on their *Daubert* standard rating system (Yanez & Fremouw, 2004).

Parenting Stress Index-Short Form. The Parenting Stress Index, Short-Form *(*PSI-SF: Abidin, 1995) is a 36-item self-report measure of parenting stress. It is comprised of three, 12-item subscales: Parental Distress, Parent-Child Dysfunctional Interaction and Difficult Child. It also includes a Defensive Responding scale that indicates a parent may be minimizing problems. A factor analysis of this scale, conducted by Haskett, Ahern, Ward, and Allaire (2006) found that the latter two scales (Parent-Child Dysfunctional Interaction and Difficult Child) were most accurately represented by one factor rather than the two proposed by the measure's author. Therefore, for the purposes of this study, a mean score was computed from the *T*-scores for these two subscales, to produce a Childrearing Stress score (term borrowed from Haskett et al., 2006). Childrearing Stress scores have been found to have a significant negative

correlation with observational measures of parenting sensitivity and positive correlations with a measure of family violence (Haskett et al., 2006). The construct validity of the PSI-SF has further been established among parents with low incomes and maltreating families (Holden & Banez, 1996; Whiteside-Mansell et al., 2007). In a review of parenting capacity measures, the PSI-SF was found to meet the *Daubert* standard for testimonial admissibility, though the authors of the article caution that the most reliable and validated scale is the Total Stress scale, with a reliability of .90 (Yanez & Fremouw, 2004). The Childrearing Stress score addresses their cited concerns regarding the unstable factor structure of the separate child and parent domain scores.

Chapter Three: Results

In this section, I present the results of the study. Prior to conducting analyses to answer the research questions, results from data screening are reported. Next, preliminary analyses for the sample are presented. Then, the results of analyses testing each of the research questions are presented.

Data Screening

Data were screened for accuracy of data entry, missing values, and univariate and multivariate outliers. Decisions regarding treatment of problematic cases and outliers were decided on a case-by-case basis. To examine distributions, the data were analyzed for skewness and kurtosis and these statistics were reported for appropriate variables.

From the specialty program dataset there were a total of 420 different biological mothers who had data for WMCI items. In accordance with Tabachnick and Fidell (2007), I conducted data screening and preliminary analyses to ensure appropriateness for factor analysis. I examined the total sample on the eight WMCI items. In the dataset, there were 11 cases with incomplete or missing data on one or more of the WMCI items. Due to the relatively low number of missing values, these cases were removed from the analysis rather than imputing estimates. Then, I conducted and examined the descriptive statistics and the distributions of each of the WMCI items. As expected with Likert-type items, there were no univariate outliers. Examination of the skewness and kurtosis of the WMCI items revealed that two of the items, Child Difficulty and Fear for Safety, were both significantly negatively skewed; however, these variables were not transformed. I made this decision following Tabachnick and Fidell's (2007) recommendation not to transform Likert-type items, particularly in the case of an already existing measure.

Additionally, since both items were negatively skewed, it was acknowledged that these items were more likely to load on a factor together (Tabachnick & Fidell, 2007).

I next conducted analyses to screen the dataset for multivariate outliers. This was accomplished by using SPSS to calculate Mahalanobis distance. Cases were removed when the Mahalanobis distance exceeded the chi-square critical value with 8 degrees of freedom (one degree of freedom for each of the WMCI items) at the $p < .001$ value, as recommended by Kline (2005). This statistic was re-examined following the removal of each case. This procedure resulted in the removal of an additional six cases. After I completed data cleaning, a total of 403 cases remained for analysis. Due to missing values on some demographic variables, percentages may not add up to 100%.

Descriptive Statistics

I examined other descriptive statistics, beyond the previously reported mother and child demographic variables. All mothers in the sample had been charged with at least one form of child maltreatment for which they currently had an open child welfare case. The primary child maltreatment charges were for neglect (78.4%), physical abuse (14.4%), medical neglect (4.3%), and sexual abuse (1.7%). Of the mothers in this dataset, 16.6% were charged with both a form of neglect and abuse. When maltreatment was dichotomized, then 76.1% of mothers were charged with neglect and 23.9% were charged with some form of abuse with, or without, neglect charges. It is important to note that the child maltreatment charge often does not reflect the research definition of that type of child maltreatment. Often neglect charges are more readily legally substantiated, based on the information provided to child protective services, though the report may include

elements of abuse (Bross, 1987). Thus, it is likely that cases with neglect substantiated, without accompanying abuse, do not represent a pure neglect sample.

At the time of the court-appointed evaluation, 34% of mothers in this sample had a substance-related disorder, and 64% had no diagnosed or treated substance-related disorder; this information was unknown for 2.5% of mothers. The majority of mothers in the sample had been victims and/or perpetrators of domestic violence (60.9%). An additional 19.5% were suspected of being, or having been, in domestically violent relationships, and 19.5% reported no domestic violence history. In this high-risk sample, number of exposures to different types of traumatic events was relatively high with a median of 5 types of traumatic events, and a range of 0 – 29 self-reported types of traumatic life events ($M = 5.86$, $SD = 4.66$).

With regard to additional child demographic variables in this sample of mother-child dyads, mental health diagnostic status was examined. At the time of evaluation, 58.2% of the children had one or more mental health diagnoses.

Descriptive statistics were also performed on the measures utilized in this study. Table 3 displays the means, standard deviations, skewness, and kurtosis statistics for the eight WMCI items, CAPI scores, and PSI-SF scores. After examination of the three validity indices for the CAPI, out of 384 cases with CAPI scores, only 151 cases remained for analysis. Eleven cases were removed due to a random response pattern and another 222 cases were removed from the analysis due to invalidating scores on the Faking Good index. The Faking Good index is a consistent response pattern of under-endorsing items that would typically be endorsed at a higher level by most individuals. It is an index of socially-desirable responding. No cases were significant for the Faking Bad

index. In reference to the Parenting Stress Index, Short-Form (PSI-SF), there were 145

cases without PSI-SF scores. Due to invalidating scores on the Defensive Responding

index another 75 cases were removed from the sample. Therefore, a total of 183 cases

remained for analyses utilizing the PSI-SF. There were no univariate outliers on any of

the measurement variables. Skewness and kurtosis indices indicated departure from a

normal distribution, but the values did not exceed the criteria set forth by Kline (2005)

with skewness values above an absolute value of 3, or kurtosis values above an absolute

value of 10.

Table 3

Means, Standard Deviations, Skewness and Kurtosis for the WMCI Items, CAPI and PSI

Variable	*M*	*SD*	Skewness	Kurtosis
WMCI: Richness of Perceptions	3.01	0.95	1.81	-3.36
WMCI: Openness to Change	3.01	0.84	1.94	-0.09
WMCI: Involvement	3.24	0.94	-0.70	-3.09
WMCI: Coherence	3.16	0.83	2.00	0.16
WMCI: Sensitivity	3.00	0.97	2.14	-2.21
WMCI: Acceptance	3.61	0.87	-0.88	-1.83
WMCI: Child Difficulty (R)	3.82	0.97	-3.24	-2.53
WMCI: Fear for Safety (R)	4.00	0.86	-3.17	-3.02
CAPI Total Score	188.34	96.50	0.69	-2.52
PSI Total Score	57.09	30.68	-2.11	-3.28
PSI Childrearing Stress	55.52	27.10	1.41	-2.50

Note: WMCI items *N* = 403. CAPI item *N* = 151. PSI items *N* = 175. (R) indicates a
reverse scored item. Skewness and kurtosis are standardized by dividing their statistic
by their standard error.

In order to determine what, if any, significant relationships exist between the

demographic variables of the sample and the WMCI items, Spearman's *rho* correlations

were calculated. For dichotomized demographic variables, the relationship to WMCI

items was explored using Hotelling's *T*.

Table 4 shows the correlations between interval/continuous child and mother variables and the WMCI items. Although some of the variables did significantly correlate with certain WMCI items, all significant correlations were weak; the strongest correlation was a weak, negative correlation between age of the child and the WMCI Child Difficulty item, with $r = -.24$. As the age of the child increased, clinician ratings of a mother viewing her child as difficult to parent increased (resulting in lower WMCI Child Difficulty scores since this item is reverse scored). These demographic variables, though statistically, significantly related to certain WMCI items, are of minimal concern as they account for approximately 6% of the items' variance, at most.

To examine the relationship between mothers' dichotomous demographic variables and WMCI items, Hotelling's T was used to control for Type I error rates. For mothers, only diagnosed substance-related disorder was significantly related to any of the WMCI items, Hotelling's $T = 25.61$, $F(8, 225) = 2.47$, $p = .01$; however, tests of between-subjects effects revealed no significant mean differences at $p < .05$ for any of the WMCI items by presence or absence of a substance-related disorder. No significant differences existed on WMCI items by mothers' exposure to domestic violence, Hotelling's $T = 6.69$, $F(8, 225) = 0.65$, $p = .74$. No significant differences were found on WMCI items by mothers' dichotomized ethnicity, Hotelling's $T = 15.13$, $F(8, 225) = 1.45$, $p = .18$. Nor were there significant differences by dichotomized child maltreatment charge, Hotelling's $T = 5.82$, $F(8, 225) = 0.57$, $p = .80$.

The same procedure, using Hotelling's T was utilized to explore possible relationships between child dichotomous demographic variables and WMCI items. Only mental health diagnosis was significantly related to any of the WMCI items, Hotelling's

Table 4

Correlations, Means, and Standard Deviations of Demographic Variables and WMCI Items

Variable	M (SD)	Richness of Perceptions	Openness to Change	Intensity of Involvement	Coherence	Caregiving Sensitivity	Acceptance	Child Difficulty	Fear for Safety
Child Age	5.37 (3.56)	.15**	.05	.07	.03	.02	.00	-.24**	-.11*
Months Out of Home	11.16 (10.18)	-.12*	-.05	-.07	-.05	-.03	-.09	-.05	-.01
Mother's Age	28.31 (7.15)	.05	-.05	.01	-.08	-.03	-.07	-.15**	-.09
Mother's Education (Years)	11.03 (2.04)	.04	.04	-.02	.00	-.05	-.01	.03	.01
Mother's Traumatic Events	5.86 (4.66)	.03	.06	.06	.12*	.13*	.09	-.04	-.09
Mother's BSI:GSI	53.84 (10.87)	-.01	.01	-.02	.00	.00	-.05	-.11*	-.14**

*Note: N*s range from 284 to 403 due to missing data. * $p < .05$, ** $p < .01$

$T = 23.57$, $F(8, 225) = 2.27$, $p = .02$. Tests of between-subjects effects revealed a significant mean difference on the WMCI Child Difficulty item. The mean difference was approximately half a point on this item; mothers of children with a mental health diagnosis were rated as perceiving their child as more difficult ($M = 4.12$) than mothers of children without a mental health diagnosis ($M = 3.63$). There was no meaningful relationship between child gender and any of the WMCI items, Hotelling's $T = 3.49$, $F(8, 225) = 0.35$, $p = .95$. Child ethnicity was also not significantly related to any WMCI item, Hotelling's $T = 7.86$, $F(8, 225) = 0.75$, $p = .65$.

After considering these minimally significant relationships between demographic variables and the eight WMCI items, I decided not to control for their influence. Additionally, none of the demographic variables significantly related, even weakly, to a majority of the WMCI items. Therefore, it is unlikely that any of these relationships influenced the factor scores of primary interest to this study. In order to test this possibility, the WMCI factor scores obtained from the exploratory and confirmatory factor analysis will be examined for significant relationships to these demographic variables as part of validity testing.

Inferential Statistics

WMCI items predict WMCI classifications. I utilized logistic regressions in order to evaluate which WMCI items meaningfully predict the different WMCI descriptive classifications. The first logistic regression was used to determine how well the WMCI items predicted the WMCI balanced versus nonbalanced classification. Then, since the regression for the dichotomized classification was significant, I conducted a

second logistic regression to explore how well each of the WMCI items predicted the WMCI disengaged versus distorted classifications.

Prediction of WMCI balanced classification. A direct logistic regression analysis was performed on WMCI classification (balanced versus nonbalanced) as the outcome and the eight WMCI items as predictors. Analysis was performed using SPSS LOGISTIC. Of the 403 cases with WMCI item scores, 20 cases did not have WMCI classifications. Data imputation was not utilized due to the few number of missing case values and the adequate sample size. Data for 380 cases were available for this analysis: 207 (54.5% of the total sample) were balanced and 173 were nonbalanced.

A test of the full model with all eight predictors against the constant-only model was statistically significant, χ^2 (8, $N = 383$) = 296.65, $p < .001$. Additionally, Negelkerke's R^2 of .72 further indicated that the WMCI items as a group moderately strongly distinguished between mothers with WMCI balanced and WMCI nonbalanced classifications. The model was able to correctly classify 90% of mothers with WMCI balanced classifications and 85% of those with WMCI nonbalanced classifications, for an overall success rate of 88%. The Hosmer and Lemeshow test for goodness-of-fit, at χ^2 (8) = 6.52, $p = .59$, supported that the predicted classifications of the full model did not significantly differ from the observed, clinician-determined classifications.

Table 5 shows the logistic regression coefficient, Wald test, and odds ratio for each of the predictors. Using a .05 criterion of statistical significance, Richness of Perceptions, Coherence, and Caregiving Sensitivity items had significant partial effects. Results indicated that for every one point increase in Coherence when other variables were held constant, a mother's WMCI narrative was 4.76 times more likely to be

classified as balanced. For every one point increase in Caregiving Sensitivity, when other

variables were held constant, a mother's WMCI narrative was 4.13 times more likely to

be classified as balanced. Finally, for every one point increase in Richness of Perceptions

when other variables were held constant, a mother's WMCI narrative was 2.14 times

more likely to be classified as balanced. As displayed in Table 4, the other variables were

not statistically significant at the .05 level; however, their odds ratios above 1.0 except

for the Child Difficulty item indicated that they contributed to the overall model

prediction. The eight WMCI items as a whole, and particularly the Richness of

Perceptions, Coherence, and Caregiving Sensitivity items significantly distinguished

between WMCI classifications of balanced or nonbalanced.

Table 5

Logistic Regression Predicting WMCI Balanced Classification from the WMCI Items

Predictor	B	Wald χ^2	p	Odds Ratio
Richness of Perceptions	0.76	9.87	.002	2.14
Openness to Change	0.57	3.58	.06	1.77
Intensity of Involvement	0.51	3.47	.06	1.67
Coherence	1.56	22.44	<.001	4.76
Caregiving Sensitivity	1.42	24.25	<.001	4.13
Acceptance	0.47	2.75	.10	1.60
Child Difficulty	-0.02	0.02	.90	0.98
Fear for Safety	0.39	3.25	.07	1.47

Note: N = 383. Nonbalanced correct classification = 84.7%. Balanced correct
classification = 90.3%.

Prediction of WMCI nonbalanced classifications. Due to the significant model,

with the eight WMCI items accurately predicting WMCI balanced versus nonbalanced

classifications, a second logistic regression was conducted. This regression was

performed to determine which items, if any, reliably predicted between the two WMCI

nonbalanced classifications (distorted versus disengaged). Of the 173 nonbalanced

74

classifications, 100 were classified as distorted (26.3% of the total sample) and the remaining 73 (19.2% of the total sample) were classified as disengaged.

A test of the full model with all eight predictors against the constant-only model was statistically significant, χ^2 (8, N = 176) = 69.92, $p < .001$. Additionally, Negelkerke's R^2 of .44 indicated that the WMCI items, as a group, moderately distinguished between mothers with WMCI distorted and WMCI disengaged classifications. The model was able to correctly classify 81% of mothers with WMCI distorted classifications and 74% of those with WMCI disengaged classifications, for an overall classification rate of 78%. The Hosmer and Lemeshow test for goodness-of-fit, at χ^2 (8) = 3.85, $p = .87$, supported that the predicted classifications of the full model did not significantly differ from the observed, clinician-determined classifications.

Table 6 shows the logistic regression coefficient, Wald test, and odds ratio for each of the predictors. Using a .05 criterion of statistical significance, only the WMCI Intensity of Involvement and Richness of Perceptions items significantly predicted between the two WMCI nonbalanced classifications. For every one point increase in Intensity of Involvement, a mother's narrative was 6.03 times more likely to be classified as distorted instead of disengaged. For every one point increase in Richness of Perceptions, a mother's WMCI narrative was 2.40 times more likely to be classified as distorted instead of disengaged.

Table 6

Logistic Regression Predicting WMCI Disengaged Classification from the WMCI Items

Predictor	B	Wald χ^2	p	Odds Ratio
Richness of Perceptions	0.88	7.62	.006	2.40
Openness to Change	-0.76	4.96	.03	0.47
Intensity of Involvement	1.80	23.86	<.001	6.03
Coherence	-0.43	2.07	.15	0.65
Caregiving Sensitivity	0.05	0.03	.87	1.05
Acceptance	-0.22	0.50	.48	0.80
Child Difficulty	-0.11	0.35	.56	0.90
Fear for Safety	-0.03	0.01	.91	0.98

Note: N = 176. Disengaged correct classification = 73.7%. Distorted correct classification = 81.0%.

Factor structure of the WMCI. In order to provide evidence for construct validity, I used both exploratory and confirmatory factor analysis. These methods allowed for an exploration of how the WMCI coding scheme items related to each other to represent the intended construct of IWMs of caregiving.

Exploratory factor analysis. Consistent with the procedures outlined by Brown (2006), the first step in establishing the usefulness of the WMCI coding scheme was exploring the structure of the construct through the use of exploratory factor analysis (EFA). I used EFA to determine the number of factors that best fit the WMCI qualitative/content items. I used SPSS 22 FACTOR to conduct the EFA.

I conducted the EFA with principal-axis factoring extraction (common factor analysis) as it is a less biased solution over principal components analysis (PCA) and it tends to estimate correlations better and produce smaller residuals (Nunnally & Bernstein, 1994). I had a few reasons for this making this extraction selection. The first reason was the directionality of the relationships between items and constructs. Most often in the social sciences there is an underlying construct that measure developers are

attempting to operationalize through items; this is the direction tested in common factor analysis, going from construct to items. PCA assumes the other direction; the items create the construct being measured. The second reason I selected common factor analysis was the type of variance it measures. If I had elected to use PCA, then all three types of variance (shared, unique, and error variance) would be measured together, which often inflates variance estimates. Of more interest when individuals are developing, or testing, a measure is the examination of the shared variance of items, which is accomplished through the use of common factor analysis (Beavers et al., 2013). The final reason I chose common factor analysis over PCA was mathematically-driven. Both common factor analysis and CFA rely on the same mathematical model (Harrington, 2009). This allowed for more direct confirmation testing. I assert multiple reasons for selecting common factor analysis over PCA, I acknowledge that frequently the results obtained by both methods are quite similar (Worthington & Whittaker, 2006).

Prior to splitting the total sample for this study, the WMCI items were examined for normality by calculating their skewness and kurtosis (See Table 3). Using SPSS, the total sample was randomly split in half, one half for use in the EFA and the other half for use in the follow-up CFA. After the random split, a total of 202 cases were used for the EFA. The descriptive statistics for this data subset are presented in Table 7. Data were re-examined for multivariate outliers, using Mahalanobis distance. All cases remained below the chi-square critical value of 26.13. Conservatively, the sample size utilized for this EFA is fair according to Comrey and Lee (1992) or large according to Kline (2005). Based on the participants-to-items ratio of 25:1, the sample size exceeds most proposed

ratios and the ratio of 3:1 most common in the social sciences (Worthington & Whittaker, 2006).

Table 7

Correlations, Means and Standard Deviations of WMCI Items for EFA Sample

WMCI Items	*M (SD)*	Open	Involv	Coh	Sens	Accept	Diff	Safe
Richness	3.00 (0.96)	.58*	.63*	.54*	.52*	.52*	.08	-.16*
Openness (Open)	2.99 (0.83)		.50*	.58*	.58*	.55*	.16*	-.12*
Involvement (Involv)	3.24 (0.93)			.46*	.63*	.63*	.11	-.23*
Coherence (Coh)	3.11 (0.85)				.54*	.56*	.13*	-.11
Sensitivity (Sens)	2.99 (0.99)					59*	.16*	-.15*
Acceptance (Accept)	3.56 (0.87)						.25*	-.11
Child Difficulty (Diff)	3.78 (1.00)							.23*
Fear for Safety (Safe)	4.05 (0.90)							

*Note: N = 202, * p < .05*

In order to examine the factorability of these data, I examined multiple criteria. First, correlations between the items were examined using Pearson's *r*. Ideally, each item would have a correlation at or above .30 with at least one of the other WMCI items (Beavers et al., 2013). As displayed in Table 7, six of the eight WMCI items had correlations within this range. Child Difficulty and Fear for Safety were the exceptions with their highest correlation with each other at .26. Given that an aim of this EFA is to examine the WMCI Coding Scheme as it is currently used, I decided to retain these two items despite their weak item correlations.

Next, in order to ensure that the determinant of the matrix was significantly different from zero, Bartlett's Test of Sphericity was examined. The determinant of the

correlation matrix was .043. Bartlett's test of sphericity, $\chi2$ (623, $n = 202$) = 623.10, $p <$

.001. When this test is significant, there are linear combinations of the items and the

items are factorable. Then, the Kaiser-Meyer-Olkin (KMO) measure of sampling

adequacy was examined to ensure that the items have enough shared variance to be

appropriate for factor analysis, reflected by a value at or above .60 (Tabachnick & Fidell,

2007). The KMO measure of sampling adequacy of .86 also supported the factorability of

the WMCI items.

To determine the number of factors to be retained, I utilized a combination of

theoretical/conceptual and empirical approaches. Examination of factor eigenvalues using

the Kaiser criterion, retaining factors with values greater than 1.0, and Cattell's scree plot

supported a two-factor solution (Beavers et al., 2013). This solution also fits with the

WMCI Coding Scheme manual which separates these items into six quality items and

two content items. Consistent with EFA results in the social sciences, the two-factor

solution accounted for 64.26% of the total variance (Beavers et al., 2013). I also

attempted to examine a both a unidimensional and three-factor solution. The three-factor

solution could not be successfully extracted within 100 iterations, making this solution

untenable. The unidimensional model resulted in significantly less total variance

accounted for in the solution, lower communalities, and weaker factor loadings.

Therefore, the two-factor solution was retained as both the best data-driven and

conceptual solution.

Initially, the two-factor solution was run using principal axis factoring with

promax rotation to allow for an oblique solution. This allowed the two factors to correlate

and also provided a factor correlation matrix, to determine if the oblique rotation was

necessary. Examination of the factor correlation matrix revealed that the two factors were not significantly correlated with each other, $r = -.02$. Following the recommendations of Tabachnick and Fidell (2007), due to the factors being uncorrelated the EFA was run an additional time using varimax rotation since the orthogonal rotation allows for more direct interpretation.

Table 8

WMCI Items, Factor Loading, Communality Estimates, and Corrected Item-Total Correlations

| | Factor Loadings | | | |
	WMCI Quality	WMCI Content	h^2	Item-Total Correlation
WMCI Quality Items:				
Richness of Perceptions	**.74**		.56	.86
Openness to Change	**.74**		.55	.89
Intensity of Involvement	**.78**		.62	.82
Coherence	**.71**		.50	.89
Caregiving Sensitivity	**.77**		.59	.89
Acceptance	**.77**		.61	.88
WMCI Content Items:				
Child Difficulty (R)	.21	**.58**	.38	.64
Fear for Safety (R)	-.19	**.45**	.24	.68

Note: $N = 202$. (R) indicates a reverse scored item. Factor loadings below an absolute value of .15 were suppressed.

Examination of the communalities of the items on the extracted factors, see Table 8, indicated that two items, Child Difficulty and Fear for Safety were potentially problematic. Both of these items communalities were below an absolute value of .40, with Fear for Safety being especially problematic at $h^2 = .24$. In both cases, these values indicate that these items are not well accounted for by the extracted factors. As previously discussed, I retained the items in order to further explore the WMCI Coding Scheme in its currently published form.

The WMCI items loaded onto two factors. Using a cross-loading criteria of .32, none of the items from either factor loaded onto more than one factor. Factor 1 was comprised of the six WMCI items described in the manual as the narrative quality items, so Factor 1 was named WMCI Quality. WMCI Quality had an eigenvalue of 3.88 and accounted for 48.54% of the total variance. All of the items on the WMCI Quality factor loaded at or above .71 (50% overlapping variance) and are considered "excellent" (Comrey & Lee, 1992). For WMCI Quality, the simple factor score (sum of item scores) mean was 18.90 ($SD = 4.32$, skewness = 2.59, kurtosis = -0.72), and a Cronbach's α = .88.

Factor 2 was comprised of the two WMCI items described in the manual as the narrative content items, and was therefore named WMCI Content. WMCI Content had an eigenvalue of 1.26 and accounted for an additional 15.72% of the total variance. As anticipated by the lower item-correlations and communalities of these two items, the rotated factor loadings of these two items were lower than those on the WMCI Quality factor, see Table 8. The Child Difficulty item factor loading of .58 was considered good, and the .45 factor loading of the Fear for Safety item was fair (Comrey & Lee, 1992). Still, both WMCI Content items loaded above .32 suggesting that they were appropriate for retention in the model. For WMCI Content, the simple factor score mean was 7.83 ($SD = 1.49$, skewness = -1.61, kurtosis = 1.96). The reliability of the WMCI Content factor, Cronbach's α = .37, indicated that this factor was highly unstable.

Table 8 provides a summary of the factor loadings, communalities, and item-total correlations for the eight WMCI variables on this two-factor EFA solution.

Confirmatory factor analysis. A confirmatory factor analysis (CFA) was

conducted to evaluate the model fit of the model from the two-factor EFA solution. The

other half of the sample, randomly split by SPSS provided the remaining 201 cases that

were not used for testing the EFA. The descriptive statistics for these cases are presented

in Table 9.

Table 9

Correlations, Means and Standard Deviations of WMCI Items for CFA Sample

WMCI Item	M (SD)	Open	Involv	Coh	Sens	Accept	Diff	Safe
Richness	3.02 (0.95)	.59*	.65*	.54*	.60*	.49*	-.04	-.07
Openness (Open)	3.02 (0.86)		.55*	.50*	.61*	.57*	.09	-.08
Involvement (Involv)	3.25 (0.95)			.46*	.66*	.61*	.04	-.21*
Coherence (Coh)	3.20 (0.81)				.54*	.55*	.04	-.07
Sensitivity (Sens)	3.01 (0.95)					.61*	.08	-.13
Acceptance (Accept)	3.65 (0.87)						.31*	.03
Child Difficulty (Diff)	3.87 (0.94)							.22*
Fear for Safety (Safe)	3.96 (0.81)							

*Note: N = 201, * p < .05*

Full CFA model. A CFA with a two-factor solution using robust maximum

likelihood (ML) estimation was performed using the AMOS 22 program (see Figure 1).

Two pairs of errors on the WMCI Quality factor were allowed to covary, since they

loaded on the same factor and modification indices indicated that allowing these

covariances would improve model fit. Several goodness-of-fit indices were used to

evaluate the model fit of the solution. The chi-square test statistic was significant χ^2 (18,

$n = 201) = 56.48$, $p < .001$; however, the low ratio (3.14) of the chi-square to degrees of freedom indicated a good model fit (Kline, 2005). The Root Mean Square Error of Approximation (RMSEA) of .10 (90% Confidence Interval: .07 - .13) indicated that the model had poor fit in terms of a parsimonious solution. The Comparative Fit Index (CFI) of .94 was close to the .95 cut-off suggested by (Hu & Bentler, 1999, as cited in Worthington & Whittaker, 2006). The Tucker-Lewis Index (TLI) of .91 was below this same threshold. Kline (2005) has suggested that values greater than .90 are acceptable for both the CFI and TLI. Additionally, the Standardized Root Mean Square Residual (SRMR) value of .07 was also acceptable (Kline, 2005). For the WMCI Quality items, all had acceptable factor loadings at or above .71 with the exception of the Coherence item which was still acceptable at .69. The WMCI Content items were more problematic with factor loadings of .50 for the Fear for Safety item, and .44 for the Child Difficulty item. Overall, the two-factor solution had borderline fit, with the high RMSEA value and low factor loadings on the WMCI Content factor being the most problematic.

Modified CFA model. Given the consistently poor performance of the two items on the WMCI Content factor, I also tested a unidimensional factor solution on the six-item WMCI Quality factor (see Figure 2). As with the initial model, item errors were allowed to covary if the modification index indicated that it would substantially improve the overall model. One such covariance was added between the Richness of Perceptions and Acceptance items. Since this model simply removes the WMCI Content factor, and that the original model did not allow for correlation between the two factors, I again refer to Figure 1. The statistics reported in the figure do not change with the elimination of the WMCI Content factor. However, this modified model significantly changes the

goodness-of-fit indices. The chi-square test statistic was not significant χ^2 (8, $n = 201$) =

12.82, $p = .12$ which indicated a good model fit. The RMSEA of .06 (90% Confidence

Interval: .00 - .11) indicated that the model had an acceptably parsimonious solution

(Harrington, 2009). Next comparative fit indices were examined; the CFI of .99 and the

TLI of .99 were above the .95 minimum value suggested for these indices. Finally, the

SRMR of .02 was well within acceptable limits. With the exception of the Coherence

item, all other items had loadings above .71, indicating that the factor accounted for a

substantial amount of the variance in the original items.

Figure 1. The EFA Hypothesized 8-Item, Two-Factor Confirmatory Factor Analysis
Model of the WMCI Coding Scheme.
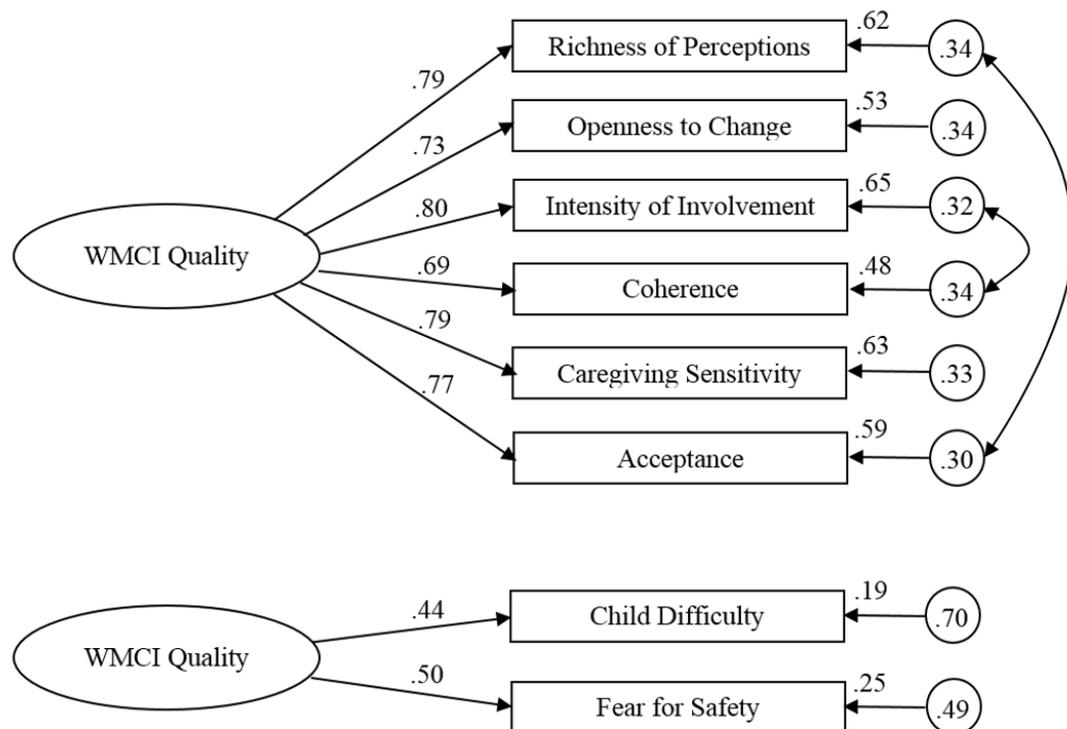
Note. N = 201. χ^2 (18, *n* = 201) = 56.48, *p* < .001. RMSEA = .10. CFI = .94. TLI = .91.
SRMR = .07.

Figure 2. Modified Unidimensional Confirmatory Factor Analysis Model of the WMCI
Quality Items.
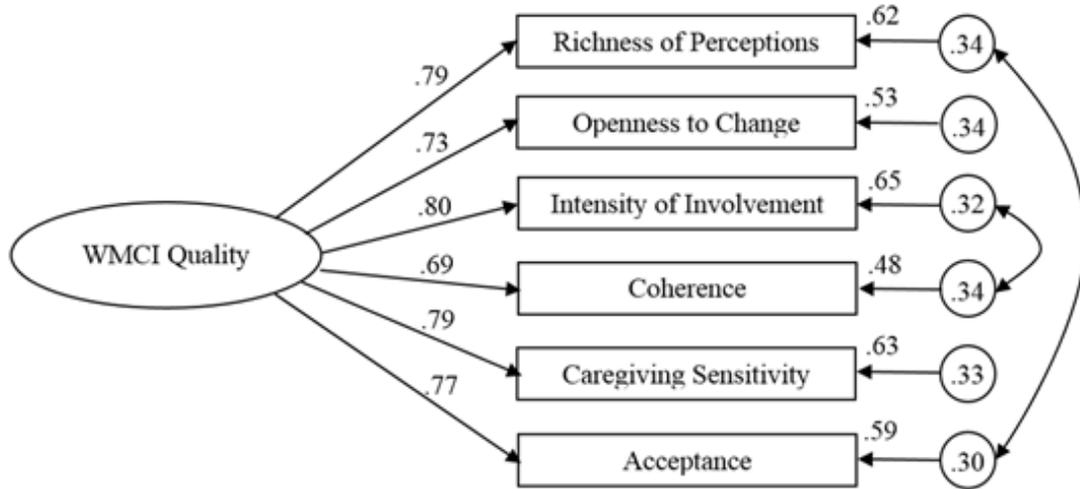
Note: $N = 201$. χ^2 (8, $n = 201$) = 12.82, $p = .12$. RMSEA = .06. CFI = .99. TLI = .99.
SRMR = .02.

Construct validity of the WMCI factor scores. Total factor scores were created

by summing the items that load onto the factors. Since the WMCI is summarized using a

single descriptive classification, and since the WMCI Content factor was unstable as an

independent factor, a WMCI Total Score was calculated by adding the WMCI Quality

and Content factor scores together. This tested the WMCI in its existing form. Given the

better model fit of the WMCI Quality factor without the influence of the WMCI Content

factor, this factor was also examined by itself. The relationship between WMCI Total

Score and WMCI Quality factor score and WMCI classification (balanced, disengaged or

distorted) was then tested by conducting a one-way ANOVA.

A one-way analysis of variance (ANOVA) was conducted to determine if there

were significant mean differences in WMCI Total Scores by type of WMCI

classification. The analysis was conducted using the three WMCI classifications:

balanced, distorted, and disengaged. The ANOVA indicated significant differences across

the three classifications, F (2, 377) = 176.02, $p < .001$, $\eta^2 = .48$. Games-Howell post-hoc

analyses indicated that all three WMCI classifications differed significantly from each

other on WMCI Total Score with $p \leq .02$ for all pair-wise comparisons. WMCI balanced

classifications had the highest mean score (M = 29.79, SD = 3.73, range 22 – 40). WMCI

distorted classifications had mean WMCI Total Scores that were significantly lower (M =

23.95, SD = 2.78, range 18 – 31). Of the three WMCI classifications the disengaged

classifications had mean WMCI Total Scores that were significantly lower than both

balanced or distorted classifications (M = 22.84, SD = 2.49, range 17 – 29). These results

confirmed my hypothesis that balanced representations would have the highest factor

scores, followed by distorted representations, and disengaged representations with the

lowest factor scores.

Due to the lack of reliability of the WMCI Content factor during the factor

analysis, a second one-way analysis of variance was conducted to determine if WMCI

Quality scores accounted for more variance in the three WMCI classifications, again

using a one-way ANOVA. To control for Type I error, significance levels were set at $p <$

.025. The ANOVA indicated significant differences in mean WMCI Quality scores by

WMCI classification, $F(2, 377) = 212.13$, $p < .001$, $\eta^2 = .53$. The results supported that

WMCI Quality scores accounted for more variance in the mean differences on WMCI

classifications than did the WMCI Total Score, which included the WMCI Content items.

Games-Howell post-hoc tests indicated that mean WMCI Quality scores significantly

differed on all three WMCI classifications, all at $p < .001$. WMCI balanced classifications

had the highest mean WMCI Quality scores (M = 21.89, SD = 3.38, range 15 – 30).

WMCI distorted classifications mean WMCI Quality scores were significantly lower (M

= 16.41, *SD* = 2.42, range 12 – 23). Similar to the previous one-way ANOVA, WMCI

disengaged classifications had mean WMCI Quality scores that were significantly lower

than both of the other WMCI classifications (*M* = 14.74, *SD* = 2.27, range 9 – 21). As

with the WMCI Total scores, WMCI Quality mean scores significantly differed by

WMCI classification in the order hypothesized from highest to lowest: balanced,

distorted, and disengaged.

Convergent validity of the WMCI scores. The convergent validity of the WMCI

factor(s) will be established with this population if there exist moderate or higher

Pearson's *r* correlations with PSI-SF Childrearing Stress or Total Stress scores and/or

with the CAPI Total Abuse score. If significant correlations exist, additional independent

samples *t*-tests comparing those above and below the respective Clinically Significant

scores on the CAPI and PSI-SF scales on WMCI factor scores will be performed.

Correlation with the CAPI. In order to evaluate the convergent validity of the

WMCI factor structure, Pearson's *r* correlations were conducted with the WMCI Total

score, WMCI Quality score and CAPI Total score. Due to the lack of reliability of the

WMCI Content factor, this score was not used independently in these analyses. It was

hypothesized that there would be moderate, or stronger, negative correlations between

WMCI scores and the CAPI. The correlation between the WMCI Total Score and the

CAPI Total Score was *r* = -.22, which was significant at *p* < .01. The correlation between

the WMCI Quality Score and the CAPI Total Score was *r* = -.18, *p* = .03. As child abuse

potential increased on the CAPI, WMCI Total Scores and WMCI Quality Scores

decreased. In both cases this represented a weak, negative correlation, which was in the

predicted direction, but lower than the hypothesized strength. Follow-up independent-

samples *t*-tests were conducted with CAPI scores dichotomized as clinical or non-clinical. For WMCI Total Score there was a significant mean difference based on CAPI classification, $t(149) = 2.06$, $p = .04$, $d = .34$. For cases that were clinically significant on the CAPI Total Score the mean WMCI Total Score was significantly lower than for non-clinical cases ($M = 26.70$, $SD = 4.18$ versus $M = 28.24$, $SD = 4.78$). For WMCI Quality Score there was no significant mean difference based on CAPI classification, $t(149) = 1.88$, $p = .06$.

  Correlation with the PSI-SF. To further evaluate the convergent validity of the WMCI factor structure, the same analyses were conducted with the WMCI Total Score, WMCI Quality score, PSI-SF Total Score, and PSI-SF Childrearing Stress Score. As with the CAPI, I hypothesized that there would be moderate, or stronger, negative correlations between the WMCI scores and the PSI-SF scores. The correlation between PSI-SF Total Score and WMCI Total Score of $r = -.16$, $p = .04$ was significant. The correlation between PSI-SF Childrearing Stress Scores and WMCI Total Scores of $r = -.19$, $p = .01$ was significant. In both cases, the correlations were weak, negative correlations. As PSI-SF scores increased, WMCI Total Scores decreased. The correlation between PSI-SF Total Scores and WMCI Quality Scores was non-significant, $r = -.11$, $p = .14$. Similarly the correlation between PSI-SF Childrearing Stress Scores and WMCI Quality Scores was non-significant, $r = -.13$, $p = .10$. As with the CAPI results, the correlational relationships between WMCI scores and PSI-SF scores were in the hypothesized direction, but weaker than hypothesized. In the case of the WMCI Quality scores, these weak correlations were not statistically significant at $p < .05$.

Since only the WMCI Total scores were significantly correlated, follow-up *t*-tests

were only conducted for this variable and the two PSI-SF scores. There were significant

mean differences in WMCI Total scores for PSI-SF Total scores that were dichotomized

as clinically significant versus nonclinical, $t(173) = 2.48$, $p = .01$, $d = .38$. PSI-SF Total

scores that were clinically significant ($M = 25.58$, $SD = 3.85$) had mean WMCI Total

scores that were significantly lower than nonclinical cases ($M = 27.17$, $SD = 4.36$). There

were significant mean differences in WMCI Total scores for PSI-SF Childrearing Stress

scores that were dichotomized as clinically significant versus nonclinical, $t(173) = 3.07$, p

$< .01$, $d = .51$. The mean WMCI Total score for clinically significant PSI-SF Childrearing

Stress ($M = 25.06$, $SD = 3.43$) was significantly lower than nonclinical PSI-SF

Childrearing Stress cases ($M = 27.15$, $SD = 4.38$).

Table 10

Correlations between Demographic Variables and WMCI Total Score and WMCI Quality
Factor Score

Demographic Variable	WMCI Total	WMCI Quality
Child Age	.00	.08
Months Out of Home	-.10	-.08
Maltreatment Severity	-.04	-.05
Mother's Age	-.08	-.03
Mother's Education (Years)	.01	.01
Mother's Traumatic Events (Count)	.06	.09
BSI: Global Symptom Index	-.05	.00

Note: *N*'s ranged between 284 and 403. No correlations were statistically significant at
$p < .05$.

Discriminant validity of WMCI scores. To test for discriminant validity,

analyses were conducted examining relationships between demographic variables of the

sample and both the WMCI Total score and WMCI Quality factor score. Since some of

the demographic variables were significantly related to specific WMCI item scores, I

wanted to determine to what extent these variables related to the WMCI factor scores. An examination of the variables: child age, time living in out-of-home care, mother's age, mother's educational level, and mother's number of traumatic life events were all examined. None of the Pearson's *r* correlations were significant between any of these variables and either WMCI Total score or the WMCI Quality factor score (see Table 10). Tests of mean differences were run for dichotomous demographic variables, including: child's ethnicity, mother's ethnicity, child mental health status, child exposure to domestic violence, type of maltreatment charge, and mother's substance-related disorder status. No significant mean differences were found for any of these variables on either WMCI Total scores or WMCI Quality factor scores.

Chapter Four: Discussion

In this chapter, I discuss the results of this dissertation. First, I provide a discussion of the descriptive and preliminary analyses. Next, I interpret and discuss the results for each research question. Then, I provide a discussion for the research and practice implications of the findings of this study. Finally, I discuss the limitations and propose future directions for research with the WMCI.

Demographics and Preliminary Analyses

When considering the interpretation of the results of this paper, it is essential to understand the sample on which these results were based. The sample was comprised of clinic-referred biological mother-child dyads with open child maltreatment cases. Due to the maltreatment, the majority of these mothers were not living with their children at the time of the data collection. Additionally, mothers included in this study were nearly all living below federal poverty levels and were Medicaid eligible. The sample was rather homogenous in terms of ethnicity as well; all participants were from the same Mid-South state and reflective of the state's demographics were predominantly White/Caucasian.

In addition to the child maltreatment risk, this sample is also more typical of a clinical, rather than general population, sample in terms of a number of other risk factors. Mothers in this sample self-reported high levels of exposure to potentially traumatic life experiences, averaging five types of exposure. A significant number of the mothers had been in domestically violent relationships, either currently or historically. Consistent with a traumatized, clinical sample, substance-related disorders were common.

Correlational analysis and tests of mean differences on these maternal demographic variables and the eight WMCI items revealed that there were few

91

significant correlations. None of the maternal demographic variables significantly correlated with multiple WMCI items, and the strongest relationships were weak correlations. Therefore, there were no attempts to control for the influence of these variables on WMCI items.

Correlational analyses and tests of mean differences were also performed on child characteristics and the WMCI items. Similar to the relationships between maternal characteristics, there were few and weak correlations or mean differences between child characteristics and WMCI scores. The WMCI item that was most related to a few demographic characteristics was the Child Difficulty item. It was significantly correlated with the age of the child, with mothers of older children more often being rated by clinicians as viewing their child as more difficult than mothers of younger children. As may be expected, mothers of children with a mental health diagnosis were rated as viewing their child as more difficult than mothers of children without a mental health diagnosis. The specificity of the relationship to this one WMCI item, with no additional relationships to the other WMCI items would suggest that this child characteristic is only influencing the variable to which it would be expected to relate. Given the weak relationship between child characteristics and WMCI item scores, and the conceptual sense between those relationships that existed, this study did not attempt to control for the influence of these relationships in the remaining analyses.

The literature has only occasionally examined the relationship between demographic characteristics of samples and the WMCI; the few that have looked at these relationships have not done so at the item level, but have instead focused on the overall WMCI descriptive classification (Benoit et al., 1997; Sokolowski et al., 2007; Vreeswijk

et al., 2012). The analyses from the present study contribute to the literature and demonstrated negligible influence of maternal and child demographic characteristics on WMCI item scores. In this sample, WMCI item scores were minimally related to any demographic variables. Thus, variables such as maternal educational level, employment status, the length of time the child had been removed from the home, etc., did not systematically influence WMCI item ratings or measure outcomes. From these results, it can be inferred that WMCI items and scores are not biased by these demographic variables. This outcome also provided some evidence of discriminant validity for the WMCI factor scores, as they were not significantly correlated with any of the demographic variables.

An unanticipated finding from the present study, was the distribution of the different types of WMCI narrative classifications. In reviews of articles that used the WMCI, there were significant differences in the distribution of WMCI classifications by the type of sample utilized – clinical versus nonclinical mothers (Vreeswijk et al., 2012). The three studies that examined WMCI classifications for mothers with diagnosed psychopathology or their own history of abuse, revealed that they were significantly less likely to have balanced classifications than nonclinical mothers. Across clinical studies, mothers' classifications were as follows: 34% balanced, 23% disengaged, and 43% distorted. WMCI classifications for mothers without clinical difficulties were: 53% balanced, 21% disengaged, and 26% distorted (Vreeswijk et al., 2012). In this dissertation study of mothers who had maltreated their children, the WMCI classification rates were more similar to studies of mothers without clinical difficulties. This finding may be attributable to the length of time between the time of the maltreatment and the

93

time of the evaluation at the speciality program. It may also reflect the effect of the case

plans that parents had been completing prior to the evaluation. Another, less optimistic

possibility was that mothers' narratives were influenced by socially-desirable responding,

given the purpose of the evaluation. This finding may also be an artifact that the majority

of the mothers in the study were not currently the primary caregiver to their child, since

most had been removed due to maltreatment some months prior; however, that

interpretation seems less likely as there was no significant linear relationship between

WMCI scores and length of time children had been removed from their mothers.

WMCI Items and WMCI Classifications

In addition to beginning to address the research gap on relationship between

demographic characteristics and WMCI item scores, this dissertation also began to

address another research gap. One of the current limitations of the WMCI manual and

literature is the lack of a clearly defined connection between clinicians' WMCI coding

scheme ratings and their overall WMCI descriptive classification. Although the manual

provides guidelines regarding specific items, ultimately the WMCI descriptive

classification is not directly derived from the WMCI item scores. This paper attempted to

address this gap by regressing the eight WMCI quality/content items on the type of

classification.

The first logistic regression was designed to explore which WMCI items

significantly predicted WMCI narrative balanced or nonbalanced classifications; this

decision was consistent with much of the literature that has used these two classifications

as their sole variable related to the WMCI (e.g., Coolbear & Benoit, 1999). The results of

the logistic regression were significant. The eight WMCI items did significantly predict

94

these classifications with considerable accuracy (88% overall). Three of the WMCI items significantly predicted whether a narrative was classified as balanced. These items were: Richness of Perceptions, Coherence, and Caregiving Sensitivity. In the case of all three of these items, the higher the rating, the more likely that the WMCI is classified as balanced. Based on these results, researchers and clinicians should pay particular attention to these items when attempting to assign a descriptive classification to the WMCI. Further, it would suggest that clinicians were consistently using the WMCI Coding Scheme items to guide the WMCI classification.

Since the WMCI items were able to accurately distinguish between the basic balanced versus nonbalanced classifications, the next step was to determine if the item scores could distinguish between the two nonbalanced classification types: distorted and disengaged. Again the model with all eight WMCI items was significant and accurately classified 78% of the nonbalanced cases. For this logistic regression, two items significantly distinguished between the subtypes: Intensity of Involvement and Richness of Perceptions. These findings are conceptually consistent with the definitions of the two nonbalanced classifications (Zeanah et al., 1996). Specifically, the disengaged classification is characterized by emotional aloofness and distance in the mother's description of her relationship with her child. When a higher score was given on the Intensity of Involvement item, the WMCI classification was significantly more likely to be a distorted classification. In a similar manner, higher scores on the Richness of Perceptions item also were more indicative of a distorted instead of a disengaged classification. These findings made sense conceptually, based on the description in the manual; they also provided the first empirical evidence supporting this distinction.

As a whole, these results provided evidence that researchers and clinicians using this measure should attend to WMCI item-level responses and the combinations of these ratings when determining the overall WMCI narrative classification. In cases where the clinician gave *moderate* or higher ratings on the Richness of Perceptions, Coherence, and Caregiving Sensitivity items, then a WMCI balanced classification was likely indicated. A *moderate* or higher rating on Richness of Perceptions without similar ratings on the other two items likely indicated a WMCI distorted classification, particularly if accompanied by a *moderate* or higher rating on the Intensity of Involvement item. A pattern of lower ratings on these four items made a WMCI disengaged classification more likely.

Factor Analysis of the WMCI Items

The next focus of this dissertation study was the psychometric evaluation of the WMCI using factor analytic procedures. A review of the literature revealed only one previous study that utilized factor analysis with the WMCI (Sprang et al., 2005). Multiple sources emphasize the need for measures to be subjected to factor analytic procedures as a routine process for developing and content validating a measure (Schmitt, 2011; Worthington & Wittaker, 2006). Factor analysis allows researchers to determine how items on a measure hold together to capture underlying constructs. In this case, measurement of the construct of the internal working model of caregiving. Additionally, it allows for empirical testing of items, which often then informs decisions regarding the retention, revision and omission of specific items.

Exploratory Factor Analysis of the WMCI. The results of the exploratory factor analysis of the eight WMCI items best supported a two-factor solution. In addition

to being the best empirical solution, it was also consistent with prior research and the

WMCI manual (Sprang et al., 2005; Zeanah et al., 1996). Six of the items loaded onto

Factor 1, which was named WMCI Quality, since this is how the items are referenced in

the WMCI manual (Zeanah et al., 1996). Those items were: Richness of Perceptions,

Openness to Change, Involvement, Coherence, Caregiving Sensitivity, and Acceptance.

The Quality factor accounted for the most variance in the model. It was a clean factor

with none of the six items substantially cross-loading onto the second factor. Each of the

items had excellent factor loadings onto the WMCI Quality factor and the factor was

acceptably reliable.

Factor 2 in the model was named WMCI Content, also in keeping with the name

used in the WMCI manual (Zeanah et al., 1996). The WMCI Content factor had two

items: Child Difficulty and Fear for Safety. This factor was problematic for a number of

reasons. The items loaded less well on the WMCI Content factor, than did the items on

the WMCI Quality factor. The Fear for Safety item only loaded as fair on this factor.

Additionally, the communalities for these items were also problematic with values below

.40. Consistent with guidance from the literature, factors with fewer than three items tend

to be unreliable as was the case with the WMCI Content factor which had very poor

internal reliability (Harrington, 2009). The lack of stability of this factor makes it of little

research or clinical utility. Further, due to the weak or nonsignificant correlations

between the items on the WMCI Content factor and the items on the WMCI Quality

factor, the two factors were not significantly correlated.

If the purpose of this study were to develop the WMCI coding scheme, then by

every indication, the items from the WMCI Content factor would have either been

revised or removed; however, since the primary purpose of this dissertation was to examine the psychometric properties of the WMCI Coding Scheme in its current form, the items were retained. Thus, the orthogonal two-factor structure was retained for the next step of the factor analytic process – confirmatory factor analysis.

Confirmatory Factor Analysis of the WMCI. A confirmatory factor analysis (CFA) of the WMCI Coding Scheme two-factor solution was examined using the other half of the sample. The results of the CFA confirmed the factor structure from the EFA, with most of the goodness-of-fit indices within acceptable ranges, with the exception of the RMSEA. Consistent with the EFA, the two-item WMCI Content factor and items continued to be problematic with lower path coefficients. I concluded that the problems that existed with model fit were likely attributable to this factor and those items.

Therefore, I ran a second CFA with only the WMCI Quality factor and the six items that loaded on that factor. By eliminating the WMCI Content factor and items, the goodness-of-fit for the model was excellent. These results suggested that when utilizing the WMCI Coding Scheme, there is empirical support to utilize the WMCI Quality factor and items separately from the WMCI Content factor and items. Since the two factors are orthogonal this modified model did not alter the retained WMCI Quality factor or its items.

If researchers or clinicians insist on the continued use of the WMCI Content factor, then in order to reduce the influence of this factor and its instability, I would suggest combining the item scores from that factor with the WMCI Quality factor into a WMCI Total Score. When making that decision, it is important to acknowledge the reasons for doing this and accept the poorer fitting, less reliable model. Using a simple

summative factor score solution with all eight items, reduces the influence of the two

WMCI Content items, and still allows them to contribute to the overall score and

decisions made with the measure. In practice, I would generally recommend against this

since it makes the measure less psychometrically sound. Since I am primarily interested

in the use of this measure as a part of parenting capacity evaluations, if this measure is

used, then there is an ethical obligation to use the measure in its most robust form. For

such purposes, my recommendation would be to only use the WMCI Quality items and

factor.

WMCI Factor Scores and WMCI Descriptive Classifications

One of the central concepts of the WMCI Coding Scheme is that it guides the

researcher or clinician to making an overall descriptive classification of a mother's

WMCI narrative response. Therefore, I examined the relationship between WMCI factor

scores and the WMCI descriptive classification given to mothers WMCI narratives. I

decided to do this first with a WMCI Total Score which was a simple summative factor

score for all eight of the items of the WMCI. This option was selected over examining the

WMCI Quality and WMCI Content factors separately for reasons already discussed. The

mean WMCI Total Scores did significantly differ based on type of WMCI classification

(balanced, distorted, or disengaged).

The mean WMCI Total Score differences were ordered in the way that was

conceptually sound with the descriptions of the three WMCI classifications (Zeanah et

al., 1996). In this sample, WMCI Total Scores ranging from 32 – 40 were all classified as

balanced. Scores at or below 21 were all either distorted or disengaged. Although there

was considerable overlap between distorted and disengaged total scores, the mean scores

of disengaged classifications were significantly lower than distorted classifications. Thus, not only did mean WMCI Total Scores significantly differ by type of WMCI descriptive classification, but also the WMCI Total Score ranges also provided some indication of whether the WMCI descriptive classification was likely balanced or nonbalanced.

These same analyses were repeated utilizing the WMCI Quality factor score instead of the WMCI Total Score. As with the WMCI Total Scores, there were significant mean differences in WMCI Quality scores by each of the three descriptive classifications. The WMCI Quality scores had a slightly larger effect size than the WMCI Total scores. This finding would support that the use of the WMCI Quality factor alone may be superior to the two-factor solution.

The mean WMCI Quality scores were highest for the balanced classification, then the distorted classification, with lowest mean scores for the disengaged classification. WMCI Quality scores from 24 – 30 were all classified as balanced narratives. On the opposite extreme, scores below 15 were all classified as nonbalanced. Additionally, scores below 12 were all classified as disengaged. As with the WMCI Total Scores, the WMCI Quality scores were useful in distinguishing between types of WMCI descriptive classifications. Unlike with the WMCI Total Score, the WMCI Quality scores better distinguished between the two nonbalanced classifications, with ranges that did not overlap as much. These results are consistent with those of the factor analyses, supporting the reliability and utility of the WMCI Quality scores over the combined WMCI Total Score.

Convergent and Discriminant Validity of the WMCI

The final aim of this dissertation was to examine the convergent validity of the WMCI Coding Scheme with other measures of parent-child relationship and/or potential for maltreatment. In order to evaluate this, correlational analyses were conducted between WMCI Total Scores, WMCI Quality scores, CAPI Total Scores, PSI-SF Total Scores, and PSI-SF Childrearing Stress scores. The hypothesized moderate or stronger negative correlations between WMCI scores and scores on these other two measures were not found. There were weak negative relationships between WMCI scores and CAPI Total Scores. When the clinical cut-off score for the CAPI was used to identify high abuse potential cases from nonclinical cases, there were significant mean differences in WMCI Total Scores in the predicted direction. WMCI Quality scores did not significantly differ by CAPI clinical cut-offs.

The weak negative correlations between the WMCI factor scores and the CAPI are perhaps more understandable given that the constructs being measured are rather distally related. Understandably, there would be a number of other factors that may influence mothers' abuse potential aside from how she perceives her relationship with her child. Additionally, the reliance on different raters, self-report for the CAPI and clinician-rated for the WMCI also likely contributed to the lower correlation. The low correspondence rate between different raters of even essentially parallel forms is a common finding in social science measurement (e.g., Behavior Assessment System for Children, Second Edition: Reynolds & Kamphaus, 2004).

Similarly, there were significant, weak negative correlations between WMCI Total Scores and PSI-SF Total Scores and PSI-SF Childrearing Stress Scores,

respectively. Again, the correlations between scores were in the expected direction, but were not at the level of significance initially expected. No significant correlation was found between WMCI Quality scores and either of the PSI-SF scores. Given that the PSI-SF is specific to the mother's rating of her relationship with one of her children, in the same way that the WMCI is specific to the mother's relationship with one of her children, it was anticipated that the correlations with this measure would have been stronger than those with the CAPI (which is not dyad specific). Correlations with the PSI-SF were essentially identical to those with the CAPI.

These results provide tentative support for convergent validity between the WMCI Coding Scheme and self-report measures of parenting stress, parent-child relationship, or child abuse potential. A stronger interpretation of these results would support the conclusion that these are distinct constructs that are minimally related to each other. Given this lack of convergence, I would recommend suspending the use of the WMCI coding scheme for clinical decision-making until further research is conducted to establish convergent validity and clinical utility. Alternatively, the WMCI coding scheme could continue to be utilized, only to the extent that the results converge with results of other formal and informal measures.

In order to examine for discriminant validity, the relationships between WMCI Total Scores and WMCI Quality scores and the sample demographic variables were explored. Since none of the demographic variables significantly correlated with the WMCI Total Score or WMCI Quality factor scores, this provided evidence that the WMCI scores are not significantly related to these variables. Additionally, none of the categorical demographic variables had significantly different mean scores on either of the

WMCI scores. This too provided evidence of a lack of relationship between demographics of the sample and WMCI scores. These results would support the conclusion that in this sample, WMCI factor scores are not unduly influenced by demographic variables that would not be expected to have meaningful relationships to the internal working model of caregiving construct being measured by the WMCI.

Potential Clinical and Forensic Implications

From the results of this study, I concluded that the items from the WMCI Coding Scheme did behave in the way they are conceptualized in the manual (Zeanah et al., 1996). A strength of the current coding scheme is the internal reliability and construct validity of the WMCI Quality factor and the six items that load on this factor. The internal consistency is well within the range for acceptable use as a research tool, and approaches the internal consistency alpha of .90 or higher suggested for clinical diagnostic measures (Sattler, 1998). Conversely, the WMCI Content factor and its two items were highly unstable, correlated poorly, or nonsignificantly, with the WMCI Quality factor and the items on that factor. In addition to the factor analytic concerns with the quality items, there were additional concerns raised by the significant relationship between these items, particularly the Child Difficulty item, and other variables such as the age of the child, whether the child has a mental health diagnosis, etc. Thus, my recommendation would be to utilize the WMCI Quality factor and items in future research and practice, and giving minimal consideration to the WMCI Content factor and items. Based on these results, current best practice would be to completely exclude these items with this and similar populations of biological mothers with histories of child maltreatment. I would be unable to justify the continued use of these items given their

poor reliability and their lack of significant relationship to any of the WMCI descriptive classifications.

This dissertation provided additional empirical support for the use of the WMCI coding scheme for mothers of children from infancy through 5 years-old as previously cited in the literature (Vreeswijk et al., 2012). This study also provided initial empirical support for the use of the WMCI coding scheme with mothers of older children. In these analyses, there were no significant WMCI factor score differences by child age up through 12 years-old. Since the age of a child was only significantly (and weakly) correlated with a single item, Child Difficulty, and that item is not on the WMCI Quality factor recommended for further research and clinical use, the results provide evidence of the utility of the measure from birth through 12 years-old.

Additionally, the literature on the WMCI has clearly established that the measure is test-retest reliable pre- to postnatal (e.g., Benoit et al., 1997), no published studies had examined the possible influence of time the dyad has lived apart on WMCI scores. This study is the first to report that there was no significant correlation between WMCI item scores and the length of time that a child had been placed with caregivers other than the biological mother; however, it is important to emphasize that all mothers in the study had visitation schedules with their children at the time of the evaluation. This result was consistent with the finding of utility of the WMCI prenatally and provides initial evidence for use of the WMCI with mothers who may not be the primary caregiver/custodian of their child.

When examining the results of this study against the *Daubert* standards for forensic use of parental capacity measures (Yanez & Fremouw, 2004), this study

provided evidence both for and against the use of the WMCI coding scheme. This study demonstrated that with the removal of the WMCI Content factor and items, the WMCI Quality factor and items were generally reliable. The items significantly related to each other and held together as one WMCI Quality factor. Additionally, the significant mean differences of WMCI Quality factor scores by type of WMCI narrative classification: balanced, distorted, or disengaged, provided additional evidence of concurrent and construct validity. The lack of significant relationships between maternal and child characteristics and WMCI scores provided some additional evidence of discriminant validity. Despite all of these findings, of considerable concern when examining the convergent validity of this measure, there were only weak, or insignificant correlations with the other measures. This finding suggested the need for additional research examining how the WMCI Quality factor relates to other aspects of the larger parental capacity construct. Given the life changing decisions that are made during the course of forensic evaluations, current best practice would be to either exclude the WMCI scores, or ensure that the results are only considered as one piece of a robust, comprehensive, multi-informant, multi-method evaluation.

Acknowledging the limitations of the convergent validity of the WMCI Quality factor in the present study, the WMCI may be useful in the evaluation of parenting interventions. Since the questions that are part of the WMCI interview schedule are questions that easily blend with an intake interview for parenting concerns, obtaining a baseline WMCI score and classification would be consistent with most intake protocols. Then after completing the parenting intervention, a post-test WMCI interview schedule could be completed, scored and again classified to measure changes in a mother's

perception of her relationship with her child as a result of the intervention. Such procedures would be particularly interesting with interventions such as Parent-Child Interaction Therapy (McNeil & Hembree-Kigin, 2010), since this intervention involves live coaching of parenting behaviors that aim to develop more positive parenting scripts. Changes from baseline to post-intervention on WMCI Quality scores would provide additional support for the validity of the WMCI coding scheme.

Limitations and Directions for Future Research

One of the limitations of this research was the reliance on a sample of court-referred mothers whom had open, substantiated child maltreatment cases at the time of the evaluation. Since the sample was not a random sample of mothers, or more ideally a stratified random sample of mothers, then the generalization of the findings from this studies to other mothers needs to be made judiciously. Additionally, this sample was quite homogenous on a number of other demographic variables including socioeconomic status, relationship to the child, and ethnicity, which also limits generalization. Simultaneously, the focus on court-referred mothers who had maltreated provided important psychometric data regarding the WMCI Coding Scheme with this special population.

It should be noted that though there were no significant differences in WMCI item scores by dichotomous maternal demographic variables (presence of substance-related disorder, ethnicity, exposure to domestic violence, or type of maltreatment charge), the lack of significance on some of these items may be due partly to inadequate power due to unequal sample sizes. For example, approximately 85% of the sample was

White/Caucasian, with then only approximately 15% Persons of Color. Furthermore, by dichotomizing some of these variables some sensitivity to differences may be lost.

Another limitation of this study was that all of the participants and clinicians were from one translational research site. Thus, this study provided potentially useful and rich information for that site and their clinicians, but generalization is again a concern. Potentially, future research could address this limitation by contacting individuals who have been trained on the WMCI by the authors of the measure and other trainers and requesting data from individuals who have been trained and are using the measure for research and practice. This would allow for the exploration and validation of scoring by location, type of setting, and sampling population.

There are many important next steps to take with the WMCI Coding Scheme. Importantly, additional research needs to be conducted to refine, solidify and more formally manualize this measure. In this dissertation, the WMCI Quality factor score had good internal consistency and appeared to represent a relatively strong factor, the WMCI Content factor was unreliable. If both factors are critical to the operationalization of the construct of internal working models of caregiving, then additional items need to be developed to more accurately and completely capture this construct.

An initial step to potentially improving the reliability of both factors would be to use the existing WMCI manual to construct multiple, simple items from the rather complex and cumbersome current items. For instance, consider the Child Difficulty item. In its current form, this item, as with all of the other WMCI items, is really a compound item. The first sentence for the Child Difficulty item reads, "This [item] is used to rate the caregiver's perception of the infant[/child] as difficult to care for and to relate to..."

(Zeanah et al., 1996, p. 10). Already, this statement could serve as the basis for two items, one that rated the caregiver's perception of caregiving difficulty, and another that rated the caregiver's perception of ability to relate to or understand the child's needs. The item description continues by emphasizing how "burdensome" the child is to the caregiver, potentially the basis for a third item. Potentially, through the generation of these multiple items, each current item could become a factor or subscale. Alternatively, additional factor analytic procedures would assist in decisions about which items to retain, reconstruct, or remove. This process would likely substantially increase the number of items for both the WMCI Quality and Content factors, and likely improve the reliability of the items and factors.

One of the considerable advantages of refining the WMCI Coding Scheme is the existence of interviews to test new items and scoring. It has been recommended in the literature that researchers and clinicians audio and/or video record WMCI interviews (Rosenblum et al., 2004). This allows for the testing of alternate scoring protocols and piloting of new items with interviews that have already been subjected to the original WMCI Coding Scheme. Such testing would require considerable researcher and clinician time, but it would drastically reduce the overall time needed, since new interviews would not need to be conducted, at least in the early stages of modifying the coding scheme.

In order for the WMCI to be more useful in both research and practice, there is also the need to develop national and/or regional norms for the WMCI. Ideally, this process would involve collecting a stratified random sample of caregivers, administering the WMCI and then having trained clinicians score and classify the caregiver interviews. This dissertation was able to provide descriptive statistics for this subpopulation of

caregivers; however, without norms making meaning of those scores beyond a descriptive level is difficult. There is no reference group of typical caregivers for comparison purposes and additional interpretation.

Another potential future direction for research with the WMCI Coding Scheme is to continue to build on the convergent and predictive validity of this measure. The majority of the literature on the WMCI has focused on the WMCI classifications, rather than the WMCI scores. Such studies are limited in the diversity of inferential analyses possible with the measure when the scores are converted from scale/continuous to categorical variables. There has been considerable attention given to the correspondence of WMCI classifications to classifications on measures of observed parent-child interactions. These studies could easily be extended by re-examining these relationships using the WMCI factor scores instead. Further, few studies have attempted to establish the WMCI coding scheme's convergent validity with measures that are not clinician rated. It is unclear to what extent the observed convergence is a result of shared rater versus shared/related construct(s). This dissertation attempted to begin to address this limitation; however, the correlations between WMCI factor scores and other measures of conceptually-related constructs was relatively weak.

In addition to more research to further refine the WMCI coding scheme, additional research should also focus on the forensic outcomes of cases where this tool was utilized as part of the evaluation. Given the previously cited frequency of repeated maltreatment after children are reunified with their biological parents, research should focus on maltreatment recidivism rates for families who participated in parent capacity evaluations similar to the one described. Lower repeated maltreatment rates

among families where the evaluation recommendations were followed by the judge would support the effectiveness of these court-appointed evaluations. At a more basic level, studies could focus on the relationship between WMCI Quality scores and the ultimate evaluation recommendations. For instance, I hypothesize that mothers with the lowest WMCI relationship quality scores, would be more likely to have recommendations that suggested no reunification, or that had significantly more and complicated recommendations to complete prior to considering reunification. Further, prospective studies could focus on the social, behavioral, emotional, and academic development of children from parenting dyads assessed to see if WMCI Quality scores had predictive validity for any of these areas of development. All of these avenues would contribute to the sparse translational research literature on the utility of the WMCI coding scheme for forensic and clinical purposes. If no significant relationships are found between the WMCI scores and any of the aforementioned variables, then there would be little empirical rationale for the continued use of the WMCI for these types of evaluations.

Conclusions

This dissertation explored the psychometric properties of the WMCI coding scheme (Zeanah et al., 1996) a measure of internal working models of caregiving, under the umbrella of attachment theory. This and similar measures have been utilized in forensic evaluations of parenting capacity and child custody within the United States (Heinze & Grisso, 1996; Main et al., 2011; Rivas et al., 2009; Sprang et al., 2004). The published literature on the WMCI coding scheme with regard to establishing its psychometric properties was lacking considering the implications of its current uses (Sprang et al., 2005).

This dissertation study was an attempt to explore and begin to establish the reliability and validity of the WMCI Coding Scheme with a sample of court-referred biological mothers who had maltreated their children. The results were mixed. When evaluated through factor analysis, the WMCI coding scheme was best represented by an uncorrelated two-factor solution. Only one of the factors, the WMCI Quality factor possessed acceptable reliability. The items on this factor were related to each other and the relationship quality factor as outlined in the WMCI manual (Zeanah et al., 1996). There were also significant mean differences in WMCI factor scores for the WMCI narrative classifications, providing some support for construct validity. Further, the results of this study indicated minimal, to no, linear relationships between maternal or child demographic variables and any of the WMCI factor scores. There were some weak relationships between demographic variables and specific WMCI items; however, most of these were with the Child Difficulty item.

The Child Difficulty item and the Fear for Safety items were the most problematic in these analyses. These two items only weakly correlated with other items on the WMCI coding scheme, and as such ended up on their own independent factor – WMCI Content. This factor, because of the problems with its items and its composition of only two items contributed to the instability of the WMCI Content factor. Based on the results of the analyses of this dissertation, these items and the WMCI Content factor do not have the empirical support necessary to warrant their continued use.

Even with the removal of the WMCI Content factor and items, additional research is necessary to further establish the convergent validity of the WMCI coding scheme. The WMCI Quality factor did not relate as strongly to other measures of different, but related,

parent-child relationship constructs. Thus, in this dissertation study there was only minimal evidence for convergent validity.

Additional research should be conducted to refine the WMCI coding scheme to more reliably and validly capture this important attachment theory construct. Such research should include the development and testing of new items as well as the simplification of the current items using factor analytic approaches. Other researchers are encouraged to not only report the overall WMCI classifications in their studies, but also the WMCI Coding Scheme item/factor scores. By doing so, researchers will be further building the empirical base for the WMCI Coding Scheme. In its present form, with similar populations, researchers may wish to focus on only the six items that loaded on the WMCI Quality factor. Only after these additional steps have occurred, should the WMCI Coding Scheme results be utilized for clinical or forensic decision-making. Until such time, its use is most appropriate as a qualitative semi-structured clinical interview rather than as a psychometrically sound measurement tool.

References

Aber, J., Slade, A., Berger, G., Bresgi, I., & Kaplan, M. (1985). *The parent development interview.* Unpublished manuscript.

Abidin, R. F. (1990). *Parenting Stress Index (Rev. ed.).* Odessa, FL: Psychological Assessment Resources.

Abidin, R. F. (1995). *Parenting Stress Index (3rd ed.).* Odessa, FL: Psychological Assessment Resources.

Ainsworth, M. D. S. (1979). Infant – mother attachment. *American Psychologist, 34,* 932-937. DOI: 0003-066X/79/3410-0932

Ainsworth, M. D. S. (1969). Object relations, dependency, and attachment: A theoretical review of the infant-mother relationship. *Child Development, 40,* 969-1025. DOI: 10.2307/1127008

American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct.* Washington, DC: American Psychological Association.

Baer, J. C., & Martinez, C. D. (2006). Child maltreatment and insecure attachment: A meta-analysis. *Journal of Reproductive and Infant Psychology, 24,* 187-197. DOI:10.1080/02646830600821231

Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (1993). A psychometric study of the Adult Attachment Interview: Reliability and discriminant validity. *Developmental Psychology, 29,* 870 – 879. DOI: 10.1037/0012-1649.29.5.870

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. J. (2013). Practical considerations for using exploratory factor analysis in

educational research. *Practical Assessment, Research & Evaluation, 18*(6). Available online: http://pareonline.net/getvn.asp?v=18&n=6

Benoit, D., Parker, K. C. H., & Zeanah, C. H. (1997). Mothers' representations of their infants assessed prenatally: stability and association with infants' attachment classifications. *Journal of Child Psychology and Psychiatry, 38,* 307-313. DOI:0021-9630/97

Benoit, D., Zeanah, C. H., Parker, K. C. H., Nicholson, E., & Coolbear, J. (1997). Working Model of the Child Interview: Infant clinical status related to maternal perceptions. *Infant Mental Health Journal, 18,* 107-121. DOI:10.1002/1097-0355(199721)18.1.107

Bosmans, G., Braet, C., Van Leeuwen, K., & Beyers, W. (2005). Do parenting behaviors predict externalizing behavior in adolescence, or is attachment the neglected 3rd factor? *Journal of Youth and Adolescence, 35,* 373-383. DOI: 10.1007/s10964-005-9026-1

Bowlby, J. (1958). The nature of the child's tie to his mother. *The International Journal of Psychoanalysis, 39,* 350-373. PMid: 13610508

Bowlby, J. (1973). *Attachment & Loss, Volume II: Separation: Anxiety and Anger.* USA: HarperCollins.

Bowlby, J. (1977a). The making and breaking of affectional bonds. I. Aetiology and psychopathology in the light of attachment theory. The fiftieth Maudsley Lecture. *The British Journal of Psychiatry, 130,* 201-210. DOI:10.1192.bjp.130.3.201

Bowlby, J. (1977b). The making and breaking of affectional bonds. II. Some principles of

    psychotherapy. The fiftieth Maudsley Lecture. *The British Journal of Psychiatry,*

    130, 421-431. DOI:10.1192/bjp.130.5.421

Bowlby, J. (1984). Violence in the family as a disorder of the attachment and caregiving

    systems. *American Journal of Psychoanalysis, 44,* 9-27.

    DOI:10.1007/BF01255416

Bowlby, J. (1988). *A secure base: Parent-child attachment and healthy human*

    development. London: Routledge.

Briere, J., & Elliott, D. M. (2003). Prevalence and psychological sequelae of self-reported

    childhood physical and sexual abuse in a general population sample of men and

    women. *Child Abuse & Neglect, 27,* 1205-1222.

    DOI:10.1016/j.chiabu.2003.09.008

Bross, D. C. (1987). Defining child abuse and neglect from a legal perspective. In D. C.

    Bross & L. F. Michaels (Eds.)*Foundations of child advocacy: Legal*

    representation of the maltreated child (pp.71-82). Lakewood, CO: Bookmakers

    Guild.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York, NY:

    Guilford Press.

Browne, C. & Winkelman, C. (2007). The effect of childhood trauma on later

    psychological adjustment. *Journal of Interpersonal Violence, 22,* 684-697. DOI:

    10.1177/0886260507300207

Budd, K. S., Clark, J., & Connell, M. A. (2011). *Best practices in forensic mental health assessment: Evaluation of parenting capacity in child protection.* New York, NY: Oxford University Press.

Budd, K. S., Poindexter, L. M., Felix, E.D., & Naik-Polan, A. T. (2001). Clinical assessment of parents in child protection cases: An empirical analysis. *Law and Human Behavior, 25,* 93 – 108. DOI: 10.1023/A:1005696026973

Carlson, E. A., Sroufe, A., & Egeland, B. (2004). The construction of experience: A longitudinal study of representation and behavior. *Child Development, 75,* 66-83. DOI: 10.1111/j.1467-8624.2004.00654.x

Ciciolla, L., Gerstein, E. D., & Crnic, K. A. (2014). Reciprocity among maternal distress, child behavior and parenting: Transactional processes and early childhood risk. *Journal of Clinical Child & Adolescent Psychology, 43,* 751-764. DOI:10.1080/15374416.2013.812038

Cohen, P., Brown, J., & Smailes, E. (2001). Child abuse and neglect and the development of mental disorders in the general population. *Development and Psychopathology, 13,* 981-999. PMid:11771917

Cohn, D. A., Cowan, P. A., Cowan, C. P., & Pearson, J. (1992). Mothers' and fathers' working models of childhood attachment relationships, parenting styles, and child behavior. *Development and Psychopathology, 4*, 417-431. DOI: 10.1017/S0954579400000870

Collins, N. L., & Read, S. J. (1994). Cognitive representations of attachment: The structure and function of working models. In K. Bartholomew & D. Perlman

(Eds.) *Attachment processes in adulthood* (pp. 53-90). London: Jessica Kingsley Publishers.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis (2nd ed.).* Hillsdale, NJ: Lawrence Erlbaum Associates.

Connell, C. M., Vanderploeg, J. J., Katz, K. H., Caron, C., Saunders, L., & Tebes, J. K. (2009). Maltreatment following reunification: Predictors of subsequent child protective services contact after children return home. *Child Abuse & Neglect, 33,* 218-228. DOI: 10.1016/j.chiabu.2008.07.005

Coolbear, J., & Benoit, D. (1999). Failure to thrive: Risk for clinical disturbance of attachment? *Infant Mental Health Journal, 20,* 87-104. DOI:10.1002/1097-0355(199921)20.1.87

Crawford, A., & Benoit, D. (2009). Caregivers' disrupted representations of the unborn child predict later infant-caregiver disorganized attachment and disrupted interactions. *Infant Mental Health Journal, 30,* 124-144. DOI: 10.1002/imhj.20207

Cyr, C., Euser, E. M., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2010). Attachment security and disorganization in maltreating and high-risk families: A series of meta-analyses. *Development and Psychopathology, 22,* 87-108. DOI:10.1017/S0954579409990289

Dayton, C. J., Levendosky, A. A., Davidson, W. S., & Bogat, G. A. (2010). The child as held in the mind of the mother: The influence of prenatal maternal representations on parenting behaviors. *Infant Mental Health Journal, 31,* 220-241. DOI: 10.1002/imhj.20253

De Bellis, M. D., Baum, A. S., Birmaher, B., Keshavan, M. S., Eccard, C. H., Boring, A.

    M., … Ryan, N.D. (1999). Developmental traumatology part I: Biological stress

    systems. *Biological Psychiatry, 45,* 1259-1270. DOI:10.1016/S0006-

    3223(99)00044-X

Derogatis, L. R. (1993). *Brief Symptom Inventory: Administration, scoring and*

    procedure manual (4th Ed.). Minneapolis, MN: National Computer Systems.

Doll, E. J. (1989). Review of Parenting Stress Index, 2nd ed. *Professional School*

    Psychology, 4, 307-312. DOI: 10.1037/h0090661

Dollberg, D., Feldman, R., & Keren, M. (2010). Maternal representations, infant

    psychiatric status, and mother-child relationship in clinic-referred and non-

    referred infants. *European Journal of Child & Adolescent Psychiatry, 19,* 25-36.

    DOI:10.1007/s00787-009-0036-5

Dykas, M. J., & Cassidy, J. (2011). Attachment and the processing of social information

    across the life span: Theory and evidence. *Psychological Bulletin, 137,* 19-46.

    DOI:10.1037/a0021367

Etelson, E. (2007). Do real feminists attachment parent? *Journal of Prenatal & Perinatal*

    Psychology & Health, 21(4)*,* 363-385.

Fang, X., Brown, D. S., Florence, C. S., & Mercy, J. A. (2012). The economic burden of

    child maltreatment in the United States and implications for prevention. *Child*

    Abuse & Neglect, 36, 156-165. DOI: 10.1016/j.chiabu.2011.10.006

Fenning, R. M., Baker, J. K., Baker, B. L. & Crnic, K. A. (2007). Parenting children with

    borderline intellectual functioning: A unique risk population. *American Journal*

on Mental Retardation, 112, 107-121. DOI: 10.1352/0895-8017(2007)112[107:PCWBIFJ2.0.CO;2

Finger, B., Hans, S. L., Bernstein, V. J., & Cox, S. M. (2009). Parent relationship quality and infant-mother attachment. *Attachment & Human Development, 11,* 285-306. DOI:10.1080/14616730902814960

Forcada-Guex, M., Borghini, A., Pierrehumbert, B., Ansermet, F., & Muller-Nix, C. (2011). Prematurity, maternal posttraumatic stress and consequences on the mother-infant relationship. *Early Human Development, 87,* 21-26. DOI:10/1016/j.earlhumdev.2010.09.006

George, C. (1996). A representational perspective of child abuse and prevention: Internal working models of attachment and caregiving. *Child Abuse & Neglect, 20,* 411-424. DOI: 10.1016/0145-2134(96)00016-6

George, C., Kaplan, N., & Main, M. (1984). *The attachment interview for adults.* Unpublished manuscript, University of California, Berkeley.

George, C., & Solomon, J. (1996). Representational models of relationships: Links between caregiving and attachment. *Infant Mental Health Journal, 17,* 198-216. DOI:10.1002/(SICI)1097-0355(199623)17:3<198

Green, J., & Goldwyn, R. (2002). Annotation: Attachment disorganization and psychopathology: New findings in attachment research and their potential implications for developmental psychopathology in childhood. *Journal of Child Psychology and Psychiatry, 43,* 835-846. DOI: 10.1111/1469-7610.00102

Gustman, B. D., & Sprang. G. (Working manuscript). Equivalence of an abbreviated Working Model of the Child Interview with the full version.

Hall, R. A. S., Hoffenkamp, H. N., Tooten, A., Braeken, J., Vingerhoets, A. J. J. M., & van Bakel, H. J. A. (2014). Longitudinal associations between maternal disrupted representations, maternal interactive behavior and infant attachment: A comparison between full-term and preterm dyads. *Journal of Child Psychiatry and Human Development, 73.* DOI:10.1007/s10578-014-0473-3

Harrington, D. (2009). *Confirmatory factor analysis: Pocket guides to social work research methods.* New York, NY: Oxford University Press.

Haskett, M. E., Ahern, L. S., Ward, C. S., & Allaire, J. C. (2006). Factor structure and validity of the Parenting Stress Index-Short Form. *Journal of Clinical Child and Adolescent Psychology, 35,* 302-312. DOI:10.1207/s15374424jccp3502_14

Heinze, M. C., & Grisso, T. (1996). Review of instruments assessing parenting competencies used in child custody evaluations. *Behavioral Sciences & the Law, 14,* 293-313. DOI:10.1002/(SICI)1099-0798(199622)14.3

Hildyard, K., & Wolfe, D. A. (2002). Child neglect: Developmental issues and outcomes. *Child Abuse & Neglect, 26,* 679-695. DOI:10.1016/S0145-2134(02)00341-1

Holden, E. W., & Banez, G. A. (1996). Child abuse potential and parenting stress within maltreating families. *Journal of Family Violence, 11,* 1-12. DOI:10.1007/BF02333337

Holden, E. W., Willis, D. J., & Foltz, L. (1989). Child abuse potential and parenting stress: Relationships in maltreating parents. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1,* 64-67. DOI: 10.1037/1040-3590.1.1.64

Huth-Bocks, A. C., Theran, S. A., Levendosky, A. A., & Bogat, G. A. (2011). A social-contextual understanding of concordance and discordance between maternal

prenatal representations of the infant and infant-mother attachment. *Infant Mental Health Journal, 32,* 405-426. DOI:10.1002/imjh.20304

Johnson-Reid, M. (2003). Foster care and future risk of maltreatment. *Children and Youth Services Review, 25,* 271-294. DOI:S0190-7409(03)00012-4

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Lyons-Ruth, K., Bronfman, E. & Parsons, E. (1999). Maternal disrupted affective communication, maternal frightened of frightening behavior, and disorganized infant attachment strategies. In J. I. Vondra & D. Barnett (Eds.), *Atypical attachment in infancy and early childhood among children at developmental risk* (pp. 172 – 192). *Monographs of the Society for Research in Child Development, 64*(3).

Macfie, J., McElwain, N. L., Houts, R. M., & Cox, M. J. (2005). Intergenerational transmission of role reversal between parent and child: Dyadic and family systems internal working models. *Attachment & Human Development, 7,* 51-65. DOI:10.1080/14616730500039663

Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. *Monographs of the Society for Research in Child Development, 50,* 66-104. DOI:10.1111/1540-5834.ep11889989

Manly, J. T., Kim, J. E., Rogosch, F. A., & Cicchetti, D. (2001). Dimensions of child maltreatment and children's adjustment: Contributions of developmental timing and subtype. *Development and Psychopathology, 13,* 759-782. PMid:11771907

May-Chahal, C., & Cawson, P. (2005). Measuring child maltreatment in the United

 Kingdom: A study of the prevalence of child abuse and neglect. *Child Abuse &*

 Neglect, 29, 969-984. DOI:10.1016/jchiabu.2004.05.009

Mayseless, O. (2006). Studying parenting representations as a window to parents'

 internal working model of caregiving. In O. Mayseless (Ed.) *Parenting*

 representations: Theory, research and clinical implications (pp. 3-40)*.* New

 York, NY: Cambridge University Press. DOI: 10.1017/CBO9780511499869.002

McBride, C., & Atkinson, L. (2009). Attachment theory and cognitive-behavioral

 therapy. In J. H. Obegi & E. Berant (Eds.). *Attachment theory and research in*

 clinical work with adults (pp. 434-458)*.* New York, NY: Guilford Press.

McFarland-Piazza, L., Hazen, N., Jacobvitz, D., & Boyd-Soisson, E. (2012). The

 development of father-child attachment: associations between adult attachment

 representations, recollections of childhood experiences and caregiving. *Early*

 Child Development and Care, 182, 701-721.

 DOI:10.1080/03004430.2011.573071

McNeil, C. B., & Hembree-Kigin, T. L. (2010). *Parent-child interaction therapy* (2nd

 ed.). New York, NY: Springer.

Minde, K., Minde, R., & Vogel, W. (2006). Culturally sensitive assessment of attachment

 in children aged 18-40 months in a South African township. *Infant Mental Health*

 Journal, 27, 544-558. DOI:10.1002/imhj.20106

Milner, J. S. (1986). *The child abuse potential inventory manual (2nd Ed.).* Webster, NC:

 Psytec Corp.

Mohammadkhani, P., Dobson, K. S., Amiri, M., Ghafari, F. H. (2010). Psychometric properties of the Brief Symptom Inventory in a sample of recovered Iranian depressed patients. *International Journal of Clinical and Health Psychology, 10,* 541-551. DOI:10.1037/t00789-000

Moss, E., Smolla, N., Cyr, C., DuBois-Comtois, K., Mazzarello, T., & Berthiaume, C. (2006). Attachment and behavior problems in middle childhood as reported by adult and child informants. *Development and Psychopathology*, 18, 425 – 444. DOI: 10.1017/S0954578406060238

Nardi, P. M. (2006). *Doing survey research: A guide to quantitative methods.* Boston: Pearson Education.

Newland, L. A., Coyl, D. D., & Freeman, H. (2008). Predicting preschoolers' attachment security from fathers' involvement, internal working models, and use of social support. *Early Child Development and Care, 178,* 785-801. DOI: 10.1080/03004430802352186

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory (3rd ed.).* New York, NY: McGraw-Hill, Inc.

Oppenheim, D., & Koren-Karie, N. (2002). Mothers' insightfulness regarding their children's internal worlds: The capacity underlying secure child-mother relationships. *Infant Mental Health Journal, 23,* 593-605. DOI:10.1002/imhj.10035

Otto, R. K., Edens, J. F., & Barcus, E. H. (2000). The use of psychological testing in child custody evaluations. *Family and Conciliation Courts Review, 38,* 312-340. DOI:10.1111/j.174-1617.2000.tb00578.x

Prather, W. (2007). Trauma and psychotherapy: Implications from a behavior analysis

   perspective. *International Journal of Behavioral Consultation and Therapy, 3,*

   555-570.

Ramsauer, B., Lotzin, A., Quitman, J. H., Becker-Stoll, F., Tharner, A., & Romer, G.

   (2014). Insightfulness and later infant attachment in clinically depressed and

   nonclinical mothers. *Infant Mental Health Journal, 35,* 210-219.

   DOI:10.1002/imhj.21446

Reizer, A., & Mikulincer, M. (2007). Assessing individual differences in working models

   of caregiving: The construction and validation of the Mental Representation of

   Caregiving Scale. *Journal of Individual Differences, 28,* 227-239.

   DOI:10.1027/1614-0001.28.4.227

Reynolds, C. R., & Kamphaus, R. W. (2004). *The behavior assessment system for*

   children (2nd ed.). Bloomington, MN: Pearson.

Rivas, E. M., Handler, L., & Sims, C. R. (2010). Adult attachment measures and their

   potential utility in custody cases. *Journal of Child Custody, 6,* 25-37.

   DOI:10.1080/15379410902894833

Rosenblum, K. L., McDonough, S. C., Sameroff, A. J., & Muzik, M. (2008). Reflection

   in thought and action: Maternal parenting reflectivity predicts mind-minded

   comments and interactive behavior. *Infant Mental Health Journal, 29,* 362-376.

   DOI:10.1002/imhj.20184

Rosenblum, K. L., Zeanah, C., McDonough, S., Muzik, M. (2004). Video-taped coding

   of Working Model of the Child Interviews: A viable and useful alternative to

verbatim transcripts? *Infant Behavior & Development, 27,* 544-549.

DOI:10.1016/j.infbeh.2004.04.001

Sattler, J. M. (1998). *Clinical and forensic interviewing of children and families:*

Guidelines for the mental health, education, pediatric, and child maltreatment

fields. La Mesa, CA: Sattler Publisher.

Schechter, D. S., Coots, T., Zeanah, C. H., Davies, M., Coates, S. W., Trabka, K. A., ...

Myers, M. M. (2005). Maternal mental representations of the child in an inner-

city clinical sample: Violence-related posttraumatic stress and reflective

functioning. *Attachment & Human Development, 7,* 313-331.

DOI:10.1080/14616730500246011

Schmidt, F., Cuttress, L. J., Lang, J., Lewandowski, M. J., & Rawana, J. S. (2007).

Assessing the parent-child relationship in parenting capacity evaluations: Clinical

applications of attachment research. *Family Court Review, 45,* 247-259.

DOI:10.1111/j.1744-1617.2007.00141.x

Schmitt, T. A. (2011). Current methodological considerations in exploratory and

confirmatory factor analysis. *Journal of Psychoeducational Assessment, 29,* 304-

321. DOI: 10.1177/0734282911406653

Scott, K. L., Wolfe, D. A., & Wekerle, C. (2003). Maltreatment and trauma: Tracking the

connections in adolescence. *Child and Adolescent Psychiatric Clinics of North*

America, 12, 211-230. DOI: 10.1016/S1056-4993(02)00101-3

Skeem, J. L., Schubert, C., Odgers, C., Mulvey, E. P., Gardner, W., & Lidz, C. (2006).

Psychiatric symptoms and community violence among high-risk patients: A tests

of the relationship at the weekly level. *Journal of Consulting and Clinical Psychology, 74,* 967-979. DOI:10.1037/0022-006X.74.5.967

Skowron, E. A., & Woehrle, P. (2012). Child maltreatment. In N. A. Fouad, J. A. Carter, & L. M. Subich (Eds) *APA handbook of counseling psychology, Vol. 2: Practice, interventions, and applications* (pp. 153-180)*.* Washington, D.C.: American Psychological Association.

Slade, A., Belsky, J., Aber, J. L., & Phelps, J. L. (1999). Mothers' representations of their relationships with their toddlers: Links to adult attachment and observed mothering. *Developmental Psychology, 35,* 611-619.

Slater, R. (2007). Attachment: Theoretical development and critique. *Educational Psychology in Practice, 23,* 205-219. DOI:10.1080/02667360701507285

Sprang, G., Clark, J. J., & Bass, S. (2005). Factors that contribute to child maltreatment severity: A multi-method and multidimensional investigation. *Child Abuse & Neglect, 29,* 335-350. DOI:10.1016/j.chiabu.2004.08.008

Sprang, G., Clark, J., Kaak, O., & Brenzel, A. (2004). Developing and tailoring mental health technologies for child welfare: The Comprehensive Assessment and Training Services (CATS) Project. *American Journal of Orthopsychiatry, 74,* 325-336. DOI:10.1037/0002-9432.74.3.325

Sokolowski, M. S., Hans, S. L., Bernstein, V. J., & Cox, S. M. (2007). Mothers' representations of their infants and parenting behavior: Associations with personal and social-contextual variables in a high-risk sample. *Infant Mental Health Journal, 28,* 344-365. DOI:10.1002/imhj.20140

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics (5ᵗʰ ed).* Boston: Allyn & Bacon.

Theran, S. A., Levendosky, A. A., Bogot, G. A., & Huth-Bocks, A. C. (2005). Stability and change in mothers' internal representations of their infants over time. *Attachment & Human Development, 7,* 253-268. DOI:10.1080/14616730500245609

Trochim, W. M. (2006). *The research methods knowledge base (2ⁿᵈ ed.).* Retrieved from http://www.socialresearchmethods.net/kb/measure.php

U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. (2010). *Child Maltreatment 2009.* Retrieved from: http://www.acf.hhs.gov/programs/cb/stats_research/index.htm#can.

van der Kolk, B. A. (2003). The neurobiology of childhood trauma and abuse. *Child and Adolescent Psychiatric Clinics of North America, 12,* 293-317. DOI:10.1016/S1056-4993(03)00003-8

van Ijzendoorn, M. H., Goldberg, S., Kroonenberg, P. M., & Frenkel, O. J. (1992). The relative effects of maternal and child problems on the quality of attachment: A meta-analysis of attachment in clinical samples. *Child Development, 63,* 840-858. DOI:0009-3920/92/6304-0019

van IJzendoorn, M. H. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the adult attachment interview. *Psychological Bulletin, 117,* 387-403. DOI: 10.1037/0033-2909.117.3.387

Vreeswijk, C. M. J. M., Janneke, A., Maas, B. M., & van Bakel, H. J. A. (2012). Parental representations: A systematic review of the Working Model of the Child Interview. *Infant Mental Health Journal, 33,* 314-328. DOI:10.1002/imhj.20337

Whiteside-Mansell, L., Ayoub, C., McKelvey, L., Faldowski, R. A., Hart, A., & Shears, J. (2007). Parenting stress of low-income parents of toddlers and preschoolers: Psychometric properties of a short form of the Parenting Stress Index. *Parenting: Science and Practice, 7,* 27-56. DOI: 10.1207/s15327933par0701_2

Wong, W. C. W., Leung, P. W. S., Tang, C. S. K., Chen, W., Lee, A., & Ling, D. C. (2009). To unfold a hidden epidemic: Prevalance of child maltreatment and its health implications among high school students in Guangzhou, China. *Child Abuse & Neglect, 33,* 441-450. DOI:10.1016/j.chiabu.2008.02.010

Wood, B. L., Hargreaves, E., & Marks, M. N. (2004). Using the Working Model of the Child Interview to assess postnatally depressed mothers' internal representations of their infants: A brief report. *Journal of Reproductive and Infant Psychology, 22,* 41-44. DOI:10.1080/02646830310001643058

Worthington, R. W., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34,* 806-838. DOI: 10.1177/0011000006288127

Yanez, Y. T., & Fremouw, W. (2004). The application of the *Daubert* standard to parental capacity measures. *American Journal of Forensic Psychology, 22* (3), 5-28.

Zeanah, C. H. (1996). Beyond insecurity: A reconceptualization of attachment disorders
of infancy. *Journal of Consulting and Clinical Psychology, 64,* 42-52.
DOI:10.1037/0022-006X.64.1.42

Zeanah, C. H., & Anders, T. F. (1987). Subjectivity in parent-infant relationships: A
discussion of internal working models. *Infant Mental Health Journal, 8,* 237-250.
DOI: 10.1002/1097-0355(198723)8:3

Zeanah, C. H., & Benoit, D. (1995). Clinical applications of a parent perception interview
in infant mental health. *Child and Adolescent Psychiatric Clinics of North
America, 4*(3)*,* 539-554.

Zeanah, C. H., Benoit, D., Barton, M. L., & Hirshberg, L. (1996). *Working Model of the
Child Interview Coding Manual.* Unpublished manuscript. Division of Infant,
Child and Adolescent Psychiatry, Louisiana State University School of Medicine,
New Orleans, LA.

Zeanah, C. H., Keener, M. A., Anders, T. F., & Vieira-Baker, C. C. (1987). Adolescent
mothers' perceptions of their infants before and after birth. *American Journal of
Orthopsychiatry, 57,* 351-360. DOI:10.1111/j.1939-0025.1987.tb03544.x

Zeanah, C. H., Boris, N. W., Heller, S. S., Hinshaw-Fuselier, S., Larrieu, J. A., Lewis,
M., … Valliere, J. (1997). Relationship assessment in infant mental health. *Infant
Mental Health Journal, 18,* 182-197. DOI:10.1002/(SICI)1097-0355(199722)18:2

Ziberstein, K. (2006). Clarifying core characteristics of attachment disorders: A review of
current research and theory. *American Journal of Orthopsychiatry, 76,* 55-64.
DOI:10.1037/0002-0432.76.1.55

<div align="center">**Vita**
BRIAN D. GUSTMAN</div>

EDUCATION:

Education Specialist in School Psychology   2011
University of Kentucky, Lexington, KY

Master of Science in Education   2005
University of Kentucky, Lexington, KY

Bachelor of Arts in Psychology   2003
Bethel College, Mishawaka, IN

GRADUATE CERTIFICATES:

Graduate Certificate in Developmental Disabilities   2008
Interdisciplinary Human Development Institute, University of Kentucky

PROFESSIONAL WORK EXPERIENCE:

School Psychologist & Mental Health Service Provider   2013 - 2015
Denver Public Schools, Denver, CO

School Psychologist & On-site Program Supervisor   2011 - 2013
Weld County School District 6, Greeley, CO

Therapist/Clinician, Temporary Licensed Psychological Associate   2010 - 2011
University of Kentucky, Center on Trauma and Children

School Psychologist (APA-approved pre-doctoral internship)   2009 - 2010
Millcreek Township School District, Erie, PA

CLINICAL EXPERIENCE:

Pre-Doctoral Psychology Intern (APA-approved Internship)   2009 - 2010
Sarah A. Reed Children's Center, Erie, PA

Graduate Student Clinician   2008 - 2009
University of Kentucky, Child & Adolescent Trauma Treatment Institute

Pre-Doctoral School Psychology Intern   2008 - 2009
Fayette County Public Schools, Lexington, KY

Advanced Practicum Student   2007 - 2008
University of Kentucky, Comprehensive Assessment and Training Services

Advanced Practicum Student – School-based Mental Health   2007
Grant County Public Schools, Dry Ridge, KY

Advanced Practicum Student – School-based Mental Health   2006
Fayette County Public Schools, Lexington, KY

School Psychology Practicum Student   2005 - 2006
Scott County Public Schools, Georgetown, KY

RESEARCH EXPERIENCE:

Senior Research Assistant, Data Analyst   2010 - 2011
for Dr. Ginny Sprang, Center on Trauma and Children, University of Kentucky

Research Assistant, Data Analyst   2008 - 2009
for KIDS NOW Initiative, Department of Special Education & Rehabilitation Counseling,
University of Kentucky

Research Assistant   2007 - 2008
for Dr. Rich Gilman, Department of Educational, School, & Counseling Psychology,
University of Kentucky

SUPERVISORY & TRAINING EXPERIENCE:

On-Site Program Supervisor   2012 - 2013
Weld County School District 6, Greeley, CO

Graduate Student Research Coordinator   2010 - 2011
Center for the Study of Violence Against Children, Intramural Research Forum,
University of Kentucky

Co-Supervisor   2009 - 2010
for Ed.S. practicum student, Sarah A. Reed Children's Center, Erie, PA

TEACHING EXPERIENCE:

Graduate Statistics Teaching Assistant   2009
for Dr. Angela Miller, Department of Educational, School, & Counseling Psychology,
University of Kentucky, Lexington, KY
Intermediate Statistics & Multivariate Statistics

Instructor   2006 - 2008
Department of Educational, School & Counseling Psychology, University of Kentucky
EDP 203: Teaching Exceptional Learners in the Regular Education Classroom

PUBLICATIONS:

Sprang, G., Staton-Tindall, M., **Gustman, B.**, Freer, B., Clark J. J., Dye, H., & Sprang,
K. (2014). The impact of trauma exposure on parenting stress in rural America.
Journal of Child & Adolescent Trauma, 6, 287-300. DOI:
10.1080/19361521.2013.836585

Whitt-Woosley, A., Sprang, G., & **Gustman, B. D.** (2014). Lives at risk: Uncovering
factors associated with fatal child maltreatment. *Children and Youth Services*
Review, 47, 307-313. DOI: 10.1016/j.childyouth.2014.10.007

Wiese, H. J. C., Wilson, J. F., Peters, J. C., Yingling, K. S., **Gustman, B. D.**, & Diener,
L. M. (2006). Assessment of classroom intervention on interpreter use with
patients of limited English proficiency. *Annals of Behavioral Science and Medical*
Education, 12, 89-91.

CONFERENCE PRESENTATIONS:

Gustman, B. (April 2011). *Exploring the psychometric properties of the Working Model*
of the Child Interview. Paper presented at the University of Kentucky, Center for
the Study of Violence Against Children Intramural Research Forum, Lexington,
KY.

Whitt-Woosley, A., **Gustman, B.**, & Sprang, G. (April 2011). *Lives at risk: Uncovering*
factors associated with fatal child maltreatment. Paper presented at the University
of Kentucky, Center for the Study of Violence Against Children Intramural
Research Forum, Lexington, KY.

Truesdale, M., **Gustman, B.**, Aldarondo, J., & Hubbard, K. (April 2011). *Comorbidity, symptom overlap or misdiagnosis: ADHD symptoms with traumatized children.* Poster presented at the University of Kentucky, Center for the Study of Violence Against Children Intramural Research Forum, Lexington, KY.

Brown, S., & **Gustman, B.** (April 2011). *Equivalence of an abbreviated WMCI with the standard WMCI.* Poster presented at the University of Kentucky, Center for the Study of Violence Against Children Intramural Research Forum, Lexington, KY.

Gustman, B. D., Sprang, G. & Risk, H. M. (November 2010). *Collaborating with public schools to meet the psychoeducational needs of trauma-exposed children and youth.* Poster presented at the Annual Meeting of the International Society for Traumatic Stress Studies (ISTSS), Montreal, QC, CA.

Gustman, B. D., Pierce, L., Missall, K. & Grisham-Brown, J. (February 2009). *Classroom environmental predictors of social skills among preschool-aged children.* Poster presented at the Annual Meeting of the National Association of School Psychologists (NASP), Boston, MA.

Pierce, L., **Gustman, B.**, Missall, K. & Grisham-Brown, J. (February 2009). *Poster presented at the Annual Meeting of the National Association of School Psychologists (NASP), Boston, MA.*

Gustman, B. D. (March 2008). *Child and adolescent trauma: Implications for school psychologists.* Paper presented at UK Student Association of School Psychology Meeting, Lexington, KY.

PROFESSIONAL LICENSURE/CERTIFICATION/MEMBERSHIPS:

Nationally Certified School Psychologist,   2012 – Present
National Association of School Psychologists

Licensed School Psychologist, *Colorado Department of Education*   2011 – Present

Certified School Psychologist, *Kentucky Department of Education*   2011 - 2015

Certified School Psychologist, *Pennsylvania Department of Education*   2009 - 2016

Temporary Licensed Psychological Associate,   2010 - 2011
Kentucky Board of Examiners of Psychology

Member of *National Association of School Psychologists*   2005 – Present

Student member of *International Society for Traumatic Stress Studies*   2008 - 2011

Student affiliate of *American Psychological Association*   2005 - 2009

Certified School Suicide Prevention Specialist,   2014 - Present
American Association of Suicidology