INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

GOODNESS OF FIT IN INFANCY: MATERNAL AND INFANT TEMPERAMENTS AND THEIR RELATIONSHIP TO MATERNAL SENSITIVITY

by

LINDSAY S. DUNCKEL

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

2003

UMI Number: 3103104

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Sept. 5, 2003
Date

Chair of Examining Committee

Sept 8, 2003
Date

Executive Officer

Arietta Slade, PhD

Peter Vietze, PhD

Darlynne Devenny, PhD

Susan Engel, PhD

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

GOODNESS OF FIT IN INFANCY:
MATERNAL AND INFANT TEMPERAMENT AND THEIR RELATIONSHIP TO MATERNAL SENSITIVITY

by

Lindsay S. Dunckel

Advisor: Professor Herbert D. Saltzstein

The purpose of this research project was to explore the concept of *goodness of fit*, which holds that it is not just a child's temperament that contributes to the course of his or her development, but how that temperament fits with the child's environment, including the mother's temperament. This *fit* has been widely discussed but has lacked an operational definition. Within the immediate context of maternal-child interaction, fit is generally understood to be a compatible relationship between mother and child; however, the components of this compatibility are not fully understood.

The current study explores the relationship of maternal and infant temperaments to maternal sensitivity, used here as a marker of a good fit. Maternal sensitivity is conceptualized and measured in two ways: 1) as quality of interaction, which was assessed in a 20-minute observation, and 2) as maternal reflective functioning, or the extent to which the mother considers her own and her infant's mental states as underlying behavior, which was assessed in maternal interviews about highpoints and lowpoints in interaction with the infants. Maternal and infant temperaments were assessed by established questionnaires. Thirty-four mothers and their four- to seven-month-old infants participated in the study.

While hypothesized relationships between easy maternal and infant temperaments and high quality interaction were not supported, an intriguing finding was that mothers who rated themselves as much more difficult in temperament than their infants had higher quality interactions than did other mothers. Several explanations for this finding are put forth, including that difficult temperament has a different meaning when applied to mothers, or that mothers who scored themselves as more difficult than their infants are likely to hold themselves accountable for problems in interaction with their infants, which contributes to their higher quality interactions.

Additionally, mothers who breastfed their infants longer scored higher on reflective functioning; and when controlling for breastfeeding history, mothers who scored higher on reflective functioning touched their infants more. These findings may reflect a discomfort with infant dependence among mothers who are less able to understand their infants' behavior.

For my parents, Jeanette and Peter Dunckel,
who always believed I could and therefore made it possible

For my husband, Randy McKean,
who has stood by my side throughout

and

For my children, Sophia and Callum McKean,
who have given me my true knowledge of motherhood

with deepest love and gratitude
because together, we made this happen

Acknowledgements

This dissertation would not have been possible without the thirty-four mothers and infants who opened their homes and their hearts to me; to them I owe my boundless thanks.

A dissertation is created by group process. My thanks go out to the members of that group: Herb Saltzstein, my advisor, who guided me through this process with warmth, respect, and humor; Arietta Slade, a member of my committee, whose always supportive nature helped me articulate what I was thinking; Peter Vietze, whose keen editorial ability helped strengthen the final product; Gary Winkel, who consulted with me on statistics and seemed to be miraculously always available and able to help; Darlynne Devenny and Susan Engel, who added their insights in serving as outside readers; Linda Perotta, director of the C.U.N.Y. Graduate Center Child Development and Learning Center, and her staff, who took such marvelous care of my children thereby enabling me to frame and conduct much of this research project; Leslie Thomas, who provided reliability coding and delightful companionship during a lonely process; Lisa Bransten, my oldest friend, who, in her desire to see me finish, managed to recruit a major portion of my West coast sample; and Sylvie Goldman, Pamela Degotardi, and Faye Fried-Walkenfeld, who, having come through the other side of the motherhood/dissertation juggling act, provided mentorship.

Table of Contents

List of Tables

I. Introduction

Goodness of fit refers to the 'fit' between individual and environment; this fit is hypothesized to influence development, for development is an emergent process in which the characteristics of the individual and the environment are mutually influential. Goodness of fit theorists have hypothesized that a good fit leads to optimal development, while a poor fit leads to less optimal development. The concept of goodness of fit was first introduced in the developmental literature by Thomas and Chess (1977); according to their concept of goodness of fit, the most adaptive outcomes are achieved by individuals whose characteristics fit well with the demands of the social and physical environment in which they are developing.

This goodness of fit model has intuitive appeal. It essentially refers to compatibility between people and institutions, between people and environments, or just between people. For instance, it is likely that a highly active child in a traditional classroom setting would have trouble meeting the behavioral expectations of her teacher, who would want her to sit still at a desk for long periods of time. She would likely be considered a behavior 'problem' in such a classroom. This poor fit of the child's temperament with the classroom setting and teacher expectations may then have negative repercussions for that child, perhaps leading to poor academic outcomes. However, a highly active child in a less traditional classroom in which children are allowed to move about the classroom and engage in more hands-on learning activities may fare better because of a better fit of her temperament with the classroom environment and teacher expectations. Goodness of fit studies usually focus on childhood, with the child's temperament serving as the individual variable of interest. Goodness of fit has most

often been explored in school-age children and adolescents, with the quality of the fit between teachers and children, peers and children, and parents and children the focus of study (e.g., Lerner, 1983; Keogh, 1986; Galambos & Turner, 1999). Yet a clear definition of what constitutes a good fit has not emerged.

Goodness of fit is central to temperament theory because it is through the fit of the individual's temperament with the environment that the notion of temperament takes on a more flexible and useful meaning. Temperament theory arose from a study exploring the antecedents of behavioral problems in young children (Thomas, Chess, Birch, Hertzig, & Korn, 1963). In focusing on early individual differences, Thomas and Chess and colleagues (e.g., Thomas, Chess, & Birch, 1968) found that a certain constellation of infant temperamental characteristics (i.e., negative mood, intense expression of mood, a tendency to withdraw from new stimuli, low adaptability of behavior in the face of change, and arhythmicity in bodily functions) were generally taxing to parents. They labeled this constellation the 'difficult' temperament. However, difficult temperament was not in and of itself predictive of poor child outcome; only when difficult temperament was disruptive for the family, in other words had a poor fit with the family environment, did it predict later childhood referral for behavioral or mental health problems. Thus, temperament takes on meaning in its interaction with the environment.

Goodness of fit is best explained using a transactional model of developmental processes. A transactional approach to development necessitates the study of child development in context; from this perspective, it is only in considering an organism within its context over time that the mechanisms of change that lead to development can

be understood (Sameroff & Chandler, 1975; Seifer & Sameroff, 1986; Sameroff & Fiese, 1992; Wapner, 1993). This approach holds that it is inadequate to consider singular causal mechanisms in developmental explanation because development is a complex, dynamic, nonlinear process. Children exist within a family and cannot be studied without consideration given to that family context (Bronfenbrenner, 1979; Lerner, Castellino, Patterson, Villaruel, & McKinney, 1995). The interpretation of temperament and the fit of temperament within the environment determine the valence of that temperament. Child development is an emergent process; a good fit or a poor fit in a given environment helps to determine child outcome.

Despite its obvious relevance to a complex and complete notion of developmental processes, goodness of fit has not often been operationalized in temperament research. This is due, in part, to the lack of agreement in the literature on what constitutes a good or a poor fit and on what aspects of the environment should be measured in defining a good fit. When researchers have explored goodness of fit, their approaches have varied greatly. Goodness of fit has been studied in the home and in the classroom, as well as within romantic pairs. It has been discussed as the fit between temperaments of people in relationships and as the fit between individual temperament and other environmental demands. Goodness of fit has been explored in school-aged children, adolescents, and young adults, but rarely has it been addressed in infancy. This, despite the fact that it was first introduced by Thomas and Chess when studying infants nearly a quarter of century ago (Thomas & Chess, 1977).

Goodness of fit research is marked by inconsistency in the definition of a good fit. While all researchers appear to agree that goodness of fit refers to the relationship

between the individual's temperament and environmental demands, including the interpersonal environment, the operationalization of these environmental demands varies greatly. Chess and Thomas stated that a good fit is not simply a match between maternal and child temperaments, because a match could sometimes be beneficial and sometimes detrimental (Chess & Thomas, 1991). Therefore, a good fit has been alternatively conceptualized as a fit between child temperament and peer, teacher, and/or parental expectations or tolerances for temperament (e.g., Lerner, 1983); as a fit between child temperament and home organization (e.g., Peters-Martin & Wachs, 1984); as a relationship between key parental temperamental variables, such as flexibility, and child temperament (e.g., Galambos & Turner, 1999); and as parents' perceptions of similarity versus 'differentness' of the child from themselves (e.g., Grotevant, McRoy, & Jenkins, 1988). And though some significant relationships have been found, no single definition of a good fit has emerged. The pattern seen most frequently seems to be a relationship between more maternal flexibility and child 'difficulty,' which is related to better outcomes (Galambos & Turner, 1999; Doelling & Johnson, 1990), though this has not been studied in the infancy period.

In infancy, 'environment' is more constrained than at most other times of life; an infant tends to spend most time with family members or set caregivers in the home or childcare center, to have few interpersonal relationships, and is unable to move about on his or her own. Because of this constrained environment, infancy seems a natural starting point for exploring the concept of goodness of fit. Additionally, as Goldsmith and Campos (1982) point out, there are advantages to studying temperament in infancy in that the infant is less susceptible to a number of socialization influences which can later mask

underlying temperament, and that his or her behavioral expressions may be less strongly mediated by cognitive processes.

In exploring goodness of fit, researchers must decide on a marker of a good fit; that is, each research project seeking to define a good fit must include a measure that stands for that good fit. Such things as measures of school performance (e.g., Lerner, Lerner, & Zabski, 1985), of foster care placement success (e.g., Doelling & Johnson, 1990), of child cognitive development (e.g., Peters-Martin & Wachs, 1984), of psychosocial functioning (Windle, Hooker, Lernerz, East, Lerner, & Lerner, 1986), and of maternal acceptance of adolescents (Galambos & Turner, 1999) have been used as markers of a good fit. Using a marker of a good fit, researchers can begin to define the aspects of child and environment, including parental attributes, that are a good fit with one another. In other words, if a good fit leads to good outcomes, then in measuring the relationship of these good outcomes to attributes of children and parents, those attributes can be identified that, when they co-occur, lead to a good fit as measured by the marker.

What does a good fit of mother and child look like during the infancy period? Maternal sensitivity, defined as a mother's ability to read, interpret, and respond to her infant's cues, is key to the regulation of mother-infant interactions because it is crucial to maintaining the flow of social interaction. Maternal sensitivity can be considered a marker of a good fit: in mother-infant pairs in which there is a good fit of temperaments, maternal sensitivity is expected to be high. The compatibility of the pair that results from a good fit is expected to be seen in increased maternal sensitivity; likewise, the incompatibility of the pair that results from a poor fit is expected to be seen in decreased maternal sensitivity. To this point, however, research has not revealed what aspects of

maternal and infant temperament make up a good fit; what matches or mismatches of temperament are related to compatibility of the pair is unknown.

It is important to distinguish that using maternal sensitivity as a marker of a good fit does not imply that a good fit can be said to *cause* maternal sensitivity, or vice versa; again, a transactional perspective precludes assigning strict, linear causality. The infant may be eliciting and shaping maternal behavior, just as the maternal behavior may be eliciting and shaping infant behavior; both processes unfold over time, the typical interactive behavior of the pair emerging slowly (see Boukydis & Burgess, 1982, a description of a laboratory study of adult physiological response to infant cries from infants of different temperaments, for a discussion about some of the ways infant temperament might influence adult behavior). In the proposed research, maternal sensitivity, defined as a mother's ability to read, interpret, and respond to her infant's cues, is considered to be not a property of the mother, but of the mother-infant relationship. This is a departure from some others' use of the term maternal sensitivity. However, given that infant behaviors and characteristics can elicit maternal behavior and influence maternal ability to read and interpret infant behavior, maternal sensitivity seems to be fundamentally a characteristic of the mother-infant pair rather than of the mother alone.

In infancy, assessment of the mother's sensitivity within the domain of mother-child interactions provides a primary means of assessing goodness of fit. Mother-infant play interactions are an excellent setting in which to observe maternal behavioral sensitivity. As has been noted by a number of infant researchers, particularly Daniel Stern (1977, 1983, 1985, 1990, 1992, 1995), mother-infant play interactions tend to be

good indicators of the overall quality of the relationship and differentiate well among types of mother-infant pairs, such as differentiating between pairs in which the mother has been diagnosed with depression and those in which the mother has not been so diagnosed. Play interactions are generally agreed to be those in which both partners are focused on the social behaviors of the other, and respond to these with social behaviors of their own. Stern describes the typical structure of mother-infant play by likening it to a dance, each episode with its own tempo. The play periods consist of engagement episodes, in which the pair interact, and time-out episodes which allow the pair to regulate the stimulation level of the interaction; indeed, engagement episodes end when an upper or lower limit for stimulation or affect is reached or is about to be reached. The regulation of this interaction by the pair is generally well established by the fourth month. Sometimes there are 'misregulations' in this dance; these are moments when upper or lower limits are breached. This is most often the result of intrusive maternal behavior, such as following the infant's gaze as the infant tries to disengage by averting his or her gaze. From a goodness of fit perspective, the regulation of this 'dance' of mothers and infants during play interaction is closely related to the fit of the pair.

Maternal interactive behavior relates to only the behavioral aspect of maternal sensitivity. As defined here and in some of the literature (i.e., Ainsworth, Blehar, Waters, & Wall, 1978), maternal sensitivity includes both psychological (reading and interpreting infants' cues) and behavioral (responding to infants' cues) components. Indeed, Slade (2002) asserts that maternal sensitive responsiveness depends on a mother's ability to accurately read the emotions, intentions, thoughts, and other internal states underlying her infant's behavior. Thus, a measure of maternal capacity for understanding the link

between internal states and behavior is necessary in order to create a more complete picture of maternal sensitivity.

By analyzing maternal use of mental state terms in describing her interactions with her infant, the extent to which a mother considers her own and her infant's mental state (beliefs, desires, comprehension, emotions, etc.) and the mutual influences of those mental states (such as how an infant can interpret the mother's joy at witnessing a new accomplishment and how this can heighten the infant's own pleasure in the accomplishment) can be assessed. The capacity to reflect upon ones' own and other's internal experience has been called "reflective functioning" by Fonagy and colleagues (e.g., Fonagy, Target, Steele, & Steele, 1998). A mother high in assessed reflective functioning will use more sophisticated descriptions of mental states; this ability to read and interpret her infant most likely underlies her behavioral sensitivity, since responding to an infant's cues is based on reading and interpreting those cues.

Temperament, maternal sensitivity, and goodness of fit thus fall together in a dynamic system, that will, over time and child development, experience the perturbations and phase shifts endemic to dynamic systems. Despite such complexity, certain patterns of maternal and infant temperament emerge as a good fit, using maternal sensitivity as a marker of goodness of fit. The identification of these patterns within a proscribed period in infancy is the purpose of this research.

Review of the Relevant Theory and Research Related to Goodness of Fit in Infancy

Goodness of Fit

Conceptual Definition

Goodness of fit refers to the compatibility of an individual's temperament with characteristics of the environment; these environmental characteristics include parental temperament. The concept of goodness of fit grew out of the temperament research of Thomas, Chess, and colleagues (e.g., Thomas, Chess, Birch, Hertzig, & Korn, 1963). Their New York Longitudinal Study (NYLS) was originally focused on individual differences in disposition. It soon became evident that although temperamental dispositions appear early in infancy, they are not independently predictive of child outcome. Their functional significance in child development only becomes clear within the environmental context of the child. The interplay between individual temperament and specific environment was labeled the goodness of fit of the child's temperament within the given context, most often the family (Thomas & Chess, 1977). The concept builds on Schneirla's (1957) idea of a *circular function*; because of temperamental individuality, people evoke differential responses from others in their environment and these responses, in turn, function as feedback which influences further interactions and therefore, serve to shape development. Children showing a "difficult" temperamental constellation in infancy did not necessarily develop behavioral disorders later in childhood: it depended on how their difficult temperament fit into their home environment (Thomas & Chess, 1977).

The New York Longitudinal Study illustrates how goodness of fit determines problem behavior. The study included two samples, one white, middle-class sample of

children of professional parents, and the other New York City Puerto Rican children of working-class parents. While similar temperaments were found in both samples, the parents' reactions to those temperaments differed. While arhythmicity in sleep patterns was a poor fit in the White, middle-class sample where there tended to be strong demands for rhythmic sleep patterns, almost no Puerto Rican parents viewed arhythmicity as a problem because they tended to structure the household around the baby rather than the reverse. Instead, high activity was a problem for the Puerto Rican parents, who tended to live in smaller apartments with fewer, safe play areas available either inside or outside their apartments. A large percentage of the Puerto Rican children who developed clinical problems in the first nine years of life were diagnosed as exhibiting problematic motor activity. This same high activity level was not troublesome to the White, middle-class sample (Thomas, Chess, Sillen, & Mendez, 1974). Thus, it is the fit of the child's temperament within the context of environment (physical environment, culture, family, parent-child interaction, etc.) which determines what will be problematic and impinge on development.

Seifer and Sameroff (1986) advanced the notion of goodness of fit by shifting the emphasis slightly from a child-centered to a relationship-centered definition:

> Rather than employ the goodness-of-fit as a characteristic of families that are more likely to produce well-adjusted children, we employ it as a fundamental property of the system under study. To elaborate, there are three basic components in the model: (1) child behavioral style, (2) parent qualities. . . , and (3) goodness-of-fit between parent and child. Where Thomas and Chess equate temperament with behavioral style (thereby locating it in the child), we propose to equate temperament with goodness-of-fit (locating it within the family system). . . With respect to temperament, we propose modifying the current question, "How does child A behave most of the time?" (child-centered definition) to "What is the nature of the behavior-style relationship with the parent?" (relationship-centered definition). (p. 33)

This line of thinking makes it very clear that temperament is only understood when contextualized; the notion of goodness of fit between parent and child is intrinsic to the conceptualization of temperament.

Measuring Goodness of Fit

Windle and Lerner (1986b) note that "while the literature associated with the 'goodness of fit' model attests unequivocally to the significance of temperament in psychosocial functioning, it does so without necessarily providing strong empirical support for the model" (p.110). There are few studies that directly test the goodness of fit model. Operationalizing the 'demands of the environment' has proven difficult.

Goodness of fit has primarily been tested in classroom settings where it has been shown to play a role in adolescent achievement and adjustment (e.g., Lerner, 1983; Lerner, Lerner & Zabski, 1985; Keogh, 1986). In classroom studies, demands of the environment are typically measured through teachers' and/or peers' ratings of tolerance for student behavior; this is then compared directly to students' ratings of their own behavior in order to assess whether or not the child meets expectations. Teacher-rated academic ability and objective academic testing serve as the outcome variables against which this 'fit' is tested. The goodness of fit model has also been tested as a one-to-one match among dating pairs where it is possible to predict the temperament of one of the pair from the score of the other, indicating that at least young adults tend to date others of similar temperament; for that study, the existence of a dating relationship was taken as evidence of a good fit (Windle & Lerner, 1984).

In one study that looked at the interaction of temperament, environment, and infant cognitive performance over the first year of life, evidence was found for the impact

of a specific fit. Highly active infants scored higher on a test of their cognitive abilities when their home environment was scored as 'organized', leading the researchers to speculate that active infants may be more sensitive to existing environmental stimulation, including physical and temporal organization (Peters-Martin & Wachs, 1984). This study indicates that a good fit would be an active infant in an organized home; for that study, infant cognitive performance was used as evidence of a good fit. Because of the primacy of the mother-infant relationship, it would be important to evaluate the fit of the infant temperament not just with the physical environment, but with maternal characteristics, which provide a critical part of the 'environment' in the first seven months, when the infant is likely to depend on adults for moving around in her environment.

Chess and Thomas, who pioneered the idea of goodness of fit, emphatically state that goodness of fit is not a simple match of temperamental characteristics between parent and child (Chess & Thomas, 1991). They bolster their argument with some case-study examples of apparent matches of temperamental characteristics that led to poor outcomes and apparent mismatches of temperamental characteristics that led to good outcomes. However, while a one-to-one match of temperamental characteristics would be a simplistic definition of goodness of fit, the idea of compatibility can be borrowed from the social psychology literature on marital adjustment (e.g., Logan, Kern, Curlette, & Trad, 1993; Lipetz, Cohen, Dworin, & Rogers, 1970). Certain characteristics complement one another to create a good fit, while others match one another to create a good fit. Patterns of characteristics are formed that create a good fit.

As part of their theory of temperament, Buss and Plomin (1984) outline the matches and mismatches they suspect would create a good fit between child and parent

temperaments, saying that the "interaction of the child's temperament with that of a parent is a source of harmony or disturbance in the relationship" (p.156). They seem to have determined these matches and mismatches intuitively or from informal observation. They postulate that a match on activity and sociability would make a good fit. They also postulate that a match on emotionality when both parent and child are highly emotional (this construct refers to negative emotionality) would be a cause of disturbance, so that it would be a better fit to have a mismatch in the case where one member of the pair is high on emotionality.

While goodness of fit has been conceptualized as the compatibility of child temperament with environmental demand, measuring environmental demand has proven difficult. While some studies focus on the fit of parental characteristics with child characteristics, others have sought to measure environmental demands as something other than parental characteristics. Some researchers create a 'demand' score by measuring parent or teacher expectations about or tolerances for children's temperament Windle, Hooker, Lernerz, East, Lerner, & Lerner, 1986). In a study which incorporated parent demand scores in investigating older adolescents' adjustment, researchers found that adolescents' individual temperament ratings (self-rated in this case) generally exceeded the corresponding parent demand score; for instance, while parents generally rated low quality of mood as difficult to tolerate, most adolescents rated their mood as in the middle to upper range (Windle, et al., 1986). As a result, little support was found for the goodness of fit model, but, because of the difficulties with variance in the scores, it seems that it was not a good test.

Another study which used a demand score to investigate goodness of fit looked at the impact on foster care placement success of the relationship of child temperament with both foster mother expectations and foster mother temperament (Doelling & Johnson, 1990). The children in the study were already placed with the foster mothers and had a mean placement length of eight months at the time of the study. Family caseworkers assessed placement success with a scale developed for the study; scores on this scale served as the marker for goodness of fit. Foster mothers of children five to seven years old were assessed for their own temperament, their perceptions of their foster child's temperament, and their expectations for child temperament. The expectations scale was a re-wording of the child temperament scale, adapted to say, "I expect my child to smile often" versus "My child smiles often."

For this expectations scale, any dimension where the child's temperament rating met or exceeded foster mothers' expectations was labeled a match; matches were predictive of more successful placement. A mismatch in only one dimension, child mood (more negative than expected), was predictive of poorer placement outcome. Foster mother to foster child temperament matches were also considered, but on just two dimensions for the mothers: flexibility-rigidity and approach-withdrawal. The relationship of these two dimensions in the foster mothers with all the child temperament dimensions were considered. The combination of an inflexible foster mother and a negative mood child was found to be predictive of low placement success, and therefore considered to be a poor fit. This combination was also predictive of lower foster mothers' satisfaction ratings, and higher family conflict. Additionally, pairings of inflexible mothers with children of low eating and sleeping rhythmicity, and high sleep

activity were predictive of poorer placement outcome and lower mother satisfaction ratings (Doelling & Johnson, 1990). Although the evidence for the matching of temperaments and of comparing temperaments with expectations in determining interaction patterns is compelling in this study, the findings would have been greatly strengthened if foster mother expectations had been assessed prior to the placement of the child.

These studies provide some support for the notion that child temperament alone is not a determinant of family interaction patterns. A study of goodness of fit focused on early infancy could help to determine whether goodness of fit begins influencing interactions as early as the first half of the first year of life.

The Concept of Temperament

Conceptual Definition

In the psychological literature, temperament is used to refer to basic infant and child personality characteristics; the characteristics that constitute temperament are a matter of much debate. Bell (1968) pioneered the idea that not only do parents influence children, but children influence parents. He argued that parental behavior is organized within repertoires in the areas of social response and child control. The child contributes to parental behaviors by activating parental behavior repertoires, affecting the level of response, and through differentially reinforcing the parental behavior which has been evoked. Temperament theory, originally espoused by Thomas and Chess (e.g., 1977), accommodates the idea that children's characteristics influence parenting behavior. Temperament theory holds that early developing personality characteristics can be observed in infancy and that these characteristics influence, and are in turn influenced by,

parental interactive behavior. One of the essential tenets of temperament theory is that similar parenting has differential effects on different children. A corollary to this is that different children can evoke different parenting. Varying dimensions of early temperament have been associated with later aggressive behavior, intellectual development, physical and mental health, rate of development, child abuse, and ability to cope with stressful life events (Sanson & Rothbart, 1995). Temperament is generally thought to act as a mediator of the influences of the environment on the individual (Peters-Martin & Wachs, 1984).

Temperament theory was a response to the heavily environmental and psychoanalytic biases of American psychology in the 1950s, which tended to hold maternal behavior accountable for child outcome (Buss & Plomin, 1984). In an effort to emphasize the early appearance of individual differences and to outline how these differences impact development, Thomas and Chess began the New York Longitudinal Study (NYLS) in 1956. They found that during the infancy period, parental behavior is most strongly influenced by children who have the temperamental constellation of the 'easy' or 'difficult' child (Thomas & Chess, 1977). A difficult child is one who is predominantly negative in mood, intense in mood expression, withdrawing from new stimuli, low in adaptability, and unpredictable in rhythms of bodily function, such as eating, sleeping, and eliminating. An easy child, in contrast, is one who is predominantly positive in mood, approaching new stimuli, highly adaptable, and highly predictable. These temperamental constellations are based on observations of nine temperamental dimensions. These personality characteristics have an influence on maternal behavior and vice-versa.

Recent research into genetic aspects of temperament seems to support the conception of temperament that grew out of the NYLS. A study of 302 pairs of twins, split almost evenly between monozygotic and dizygotic, included maternal-report and laboratory measures of temperament (Goldsmith, Lemery, Buss, & Campos, 1999). Using a behavior genetics approach, the researchers found that most dimensions of temperament are best modeled by incorporating additive genetic, shared environmental, and non-shared environmental influences, with the genetic influences contributing a surprisingly small amount (Goldsmith, et al., 1999). Shared environmental influences include such things as maternal personality; non-shared environmental influences presumably include such things as individual relationship qualities, security of attachment, or evocations of behavior due to one's characteristics. The research was not designed to evaluate the components of these factors, only their relative contributions in determining temperament. The resulting model of the determinants of temperament closely mirrors the mediated process Thomas and Chess first described.

Varying Approaches

There is no universally accepted definition of temperament, no predominant measure for temperament. Because of this, and the fact that temperament research has been going on for more than 60 years, the field of temperament research is a disparate one. In an attempt to synthesize this disparate field, Goldsmith and Campos (1982) offer an overarching *theory* of temperament. In attempting to synthesize the field, they first describe the four most prevalent approaches to temperament: Thomas and Chess' *style approach,* Buss and Plomin's *criterial approach,* Rothbart's *psychobiological approach,* and Brazelton's *neonatal approach.* These are outlined briefly below.

Thomas and Chess (e.g., 1977) describe temperament as behavioral style. In their view, temperament consists of nine dimensions; this list has been criticized by other researchers who argue that it includes different levels of analysis and suffers from overlap. But Thomas and Chess's style approach to temperament remains the dominant approach; other approaches were formulated in reaction to it.

Buss and Plomin's criterial approach defines temperament as a set of inborn personality traits that appear early in life, which include emotionality, activity, and sociability (EAS theory). These dimensions are the ones that fit the five criteria specified by this approach: that the trait be heritable, stable, predictive of adult personality, adaptive in the evolutionary sense, and found in other animals.

The psychobiological approach acknowledges individual differences in reactivity and self-regulation as the components of temperament; in this view, temperament is tied to the nervous system and its response systems (Rothbart, 1981). The measurable components of the systems are the focus of this approach, including intensity, threshold, latency of response, rise time of response, and recovery time.

Brazelton's Neonatal Behavioral Assessment Scale (NBAS) has been used to consider the earliest measurable infant qualities, essentially neurophysiological ones. In Brazelton's neonatal approach, these earliest infant characteristics are the foundation of later infant temperament (Goldsmith & Campos, 1982). Much research has looked at the relationship of these early assessments to later assessments of infant temperament (Brazelton, Koslowski, & Main, 1974).

Goldsmith and Campos (1982) attempt to create an overarching theory of temperament that draws on these four dominant approaches. They argue that a behavior-

based approach to temperament, like that of Thomas and Chess, is best for both a theoretical and a practical reason. Both reasons arise from a consideration of the purpose of infant temperament. Its purpose, they argue, is to "increase the predictability of the infant's behavior and accordingly modify caretaker behavior. This latter function is an example of the socially communicative role of temperament. For example, the caretaker who is sensitive to the signals of a characteristically very irritable (i.e., low threshold for distress) infant will, in all likelihood, attempt to avoid exposing the infant to irritating stimuli" (Goldsmith & Campos, 1982, pp.179-180). Thus, the theoretical and the practical reason for embracing a behavior-based approach to temperament is that behavior is easier for a social partner to read, and for an observer to measure, than are the underlying physiological phenomena which are determinant of temperament in the other approaches (Goldsmith & Campos, 1982). Goldsmith and Campos sum up their theory of infant temperament by saying that temperament is an individual differences construct, and not a theory of normative development; that temperament is a dispositional construct; that in infancy, the dimensions of temperament are affect-related, or emotional in nature, in that their primary purpose is social-communicative; and that differences in temperament are based on individual differences in the intensity and timing with which the dimensions of temperament are expressed (Goldsmith & Campos, 1982). To date, this is the most solid, comprehensive, theoretically supported description of infant temperament.

Measuring Temperament

Shortly after Thomas, Chess, and colleagues began publishing their temperament findings from the New York Longitudinal Study (e.g., Thomas, Chess, Birch, Hertzig, &

Korn, 1963), a pediatrician named William B. Carey began to develop a questionnaire which could assess the nine dimensions of temperament that had been defined. Carey's Infant Temperament Questionnaire (ITQ; Carey, 1970) was the first questionnaire with which temperament could be briefly assessed. The ITQ was a parent-report instrument; though others were developed (e.g., Lerner, Palermo, Spiro, & Nesselroade, 1982), the ITQ, in its revised form, has remained the most used parent-report measure of temperament. Parent-report measures are popular because they are efficient and rely on the parent for information about the child: the parent has more extended, direct experience with the infant than any outside rater would. Over the years, other methodologies for measuring temperament have emerged, such as behavioral observations (e.g., Rothbart, 1981) and laboratory assessments of reactivity (e.g., Wilson & Matheny, 1986). As a result, much debate over the interpretation and meaning of temperament when measured in these different ways has arisen. Rothbart (1982) argues that temperament should continue to be studied by a variety of means, including parent report, in-home observations, and controlled laboratory settings, because each method has the potential to increase our knowledge about temperament.

Some researchers have critiqued temperament as being a construct of maternal perception and not a measure of actual child characteristics when measured with a parent-report instrument (e.g., Bates, 1980). Some argue that child temperament is most reliable when rated by observers either in the lab or in the home (e.g., Rothbart, 1981). Others defend the use of maternal report measures as the only way of tapping infant behavior over time, arguing that a maternal report measure which is behaviorally anchored does not include maternal perception but only maternal report of infant behavior (e.g.,

Thomas, Chess, & Korn, 1982). Furthermore, rater bias is reduced in most parent-report measures of infant temperament through the reverse scoring of some items and the randomization of questions assessing a single dimension. Kagan (1982) argues that parents are able to report accurately about infant behavior, particularly when that behavior is extreme, as in an infant displaying a difficult temperament. Indeed, a correlation between mothers' and fathers' temperament ratings has been found in studies using different rating scales and was found to be particularly strong ($r = .58$) for a subset of scales making up the difficult rating (Bates, Freeland, & Lounsbury, 1979). Some researchers question why, when there is a discrepancy between parent-rated and observer-rated infant temperament, the error is assumed to lie in the parents' ratings and not in the observers' ratings after their short-term exposure to the child (Carey, 1983). Because there is no adequate, standardized, observation rating technique, researchers must use one of several different observation protocols. Most of these are brief, allowing for a single observation time with the infant, and are unmatched to the temperament dimensions in parent-report measures; therefore, when discrepancies between parent report and observer report of infant temperament arise, it is difficult to separate discrepancies due to measurement construct from discrepancies due to parental perception interfering with accurate reporting of child temperament.

Thomas, Chess, and Korn (1982) state that even if parental distortion is a part of temperament ratings, this, too, will contribute to child development in a meaningful way and therefore merits study. Seifer and Sameroff (1986) suggest that subjective aspects of parent-report measures of temperament become central rather than be weeded out because temperament should be conceptualized as located within the interaction, rather

than within the child. Bates and Bayles (1984) found support for the idea that maternal temperament ratings consist of both objective and subjective components, as shown by the relationship of maternal temperament ratings to both observer variables and maternal personality variables.

A study which considered the convergence between laboratory assessment and maternal report of infant temperament involved mothers and infants at four time-points in the first year (Carnicero, Perez-Lopez, Salinas, & Martinez-Fuentes, 2000). Evidence of a convergence between the maternal report and laboratory assessments was found at 6 and 12 months (Carnicero, et al., 2000).

A transactional perspective frees temperament research from the debate about whether temperament is a parental construct or inherent in the child: temperament must lie neither within the child nor within the adult perceptions of the child, but somewhere between the two, in the interaction of adult and child, mutually defined and refined. By moving the discussion of temperament into a relationship-centered focus, as urged by Seifer and Sameroff (1986), the important question for researchers to ask about temperament is not "How does the child behave?" but "What is the nature of the behavior-style relationship with the parent?"

Because a mother's perception of her child's temperament forms the basis of her response in interacting with the child, a parent-report measure of temperament is crucial to understanding goodness of fit. Following the same argument, a self-report measure of maternal temperament is also crucial. How objective these measures are thought to be loses relevance when it is maternal perception of her own and her infant's temperaments that are the focus of the research. What is important is the extent to which the measures

have been used in the literature, so that comparisons to other studies may be made and the dialogue about goodness of fit extended.

Defining a Difficult Temperament

Thomas, Chess, and colleagues (1963) originally defined a difficult child as one who is predominantly negative in mood, intense in mood expression, withdrawing from new stimuli, low in adaptability, and unpredictable in rhythms of bodily function; through their extensive interviews with parents, they had determined that these dimensions of temperament were most troublesome to parents. Thomas and Chess (1982) suggested that temperament scores be used as a continuous variable; this continuous variable would mirror the range of temperaments, from very easy to very difficult, and would provide a more realistic picture of temperament than would categorizing children into easy and difficult groups.

In creating an operational definition of a difficult temperament, many researchers have followed this advice and constructed a 'difficultness composite' score by summing individual scores for different dimensions of temperament; most scores from temperament questionnaires run from least difficult to most difficult so that scores may be summed in such a way. For instance, Mednick, Hocevar, Schulsinger, and Baker (1996) summed standardized scores for mood, approach/withdrawal, and flexibility/rigidity (from the Dimensions of Temperament Survey-Revised; Windle & Lerner, 1986a) to get an overall difficulty rating in their study of the relationship of maternal characteristics to child temperament. Similarly, Sheeber and Johnson (1992), using a temperament questionnaire developed by Thomas and Chess (1977), summed the scores for mood, intensity, rhythmicity, approach/withdrawal, and adaptability to create a

global temperament score in their study of the relationship of child temperament and maternal adjustment. Following Thomas and Chess, Carey and McDevitt (1978), in designing their much-used Revised Infant Temperament Questionnaire, defined a child of difficult temperament as one who is arrhythmic, withdrawing, low in adaptability, intense in reaction, and predominantly negative in mood.

As new approaches to temperament arose, the Thomas and Chess definition of a difficult temperament was criticized, leading to the formation of new definitions of a difficult temperament; now several such definitions exist, including an equating of difficult temperament with simple irritability (e.g., van den Boom, 1994). One of the main criticisms leveled at the Thomas and Chess definition came out of their own work, which revealed a cultural bias in the way rhythmicity functioned, with White, middle-class parents reporting arhythmicity in sleep patterns as problematic, while working-class Puerto Rican immigrant parents did not find it to be a problem (Thomas, Chess, Sillen, & Mendez, 1974). Because of this finding, researchers who choose instruments for assessing temperament which include the classic Thomas and Chess dimensions of temperament often omit rhythmicity from their operational definition of a difficult temperament.

Mediators and Correlates of Child Temperament

Maternal perception of child temperament has been found to be related to many other factors, including other child characteristics, maternal characteristics, and maternal interactive behavior. Evidence of a correlation between maternal interactive behavior and infant and child temperament is substantial. This literature as it pertains to the proposed research is reviewed beginning on page 27.

Birth order and gender are the child characteristics that have been shown to be most robustly related to ratings of temperament (Bates, Freedland, & Lounsbury, 1979; Klein, 1984; Gordon, 1983; Sanson & Rothbart, 1995). Mothers with more than one child tend to rate their youngest child as less difficult than do first time mothers (Bates, Freeland, & Lounsbury, 1979). While boys and girls have been shown by observation to differ only in ratings of activity level, maternal behavior toward intense infant girls (those with strong responses, either positive or negative, to various situations) has been shown to differ from that toward intense infant boys in an Israeli sample (Klein, 1984). It has been found that more negativity is accepted from boys than girls, and more shyness is accepted from girls than boys (Sanson & Rothbart, 1995).

Maternal characteristics have also been found to be related to ratings of child temperament, particularly maternal personality and mental health factors. Maternal extroversion has been found to be negatively correlated with child's difficult temperament rating; however, correlations between maternal extroversion scores and observer ratings of infant crying and soothability were very low, indicating that extroverted mothers perceive their babies as less difficult although their babies do not appear to behave differently on these dimensions (Bates, Freeland, & Lounsbury, 1979).

Negative affectivity in the mother is positively related to proneness-to-distress in infants (Mangelsdorf, Gunnar, Kestenbaum, Lang, & Andreas, 1990). Maternal positive affectivity was negatively related to infants' proneness-to-distress (Mangelsdorf, et al., 1990). Additionally, mothers who report themselves to be more emotionally stable and socially dominant also report their children to be temperamentally easier (Matheny, Wilson, & Thoben, 1987). Mothers of temperamentally difficult children reported more

anxiety and depression, as well as lower self-rated parenting competence (Sheeber & Johnson, 1992); the link between maternal depression and child difficult temperament has been found elsewhere, as well (Gross, Conrad, Fogg, & Wothke, 1994), as has the link between maternal anxiety and child difficult temperament (Mednick, Hocevar, Schulsinger, & Baker, 1996), though no direction of effect has been demonstrated. Of course, all of these concurrences in maternal personality and child temperament could be due to common underlying, heritable personalities. Most research thus far has not been able to address this issue.

Maternal Sensitivity

Maternal sensitivity in infancy, defined as a mother's ability to read, interpret, and respond to her infant's cues, has been found to be a key component of later child cognitive and social development. This was first found for maternal behavioral sensitivity, or her ability to respond to her infant's cues (e.g., van den Boom, 1995) and then explored in the psychological and cognitive process underlying this behavioral sensitivity (e.g., Slade, Grienenberger, Bernbach, Levy, & Locker, 2001). This underlying process, maternal ability to interpret and make sense of others' states of mind and reasons for action, is called reflective functioning (e.g., Steele, Steele, & Fonagy, 1996). Maternal capacity for reflective functioning has been strongly linked to mother-infant attachment (Slade, et al., 2001) and has been found to be negatively related to maternal negative interactive behavior (Grienenberger, Kelly, & Slade, 2001). Maternal sensitivity implies the ability to read, interpret, and act on infant cues, a combination of psychological functions (reading and interpreting) and behavioral functions (acting on). A study of maternal sensitivity in infancy ought to include measures of both capacities.

What researchers can observe in a mother's behavior with her infant and what that mother can tell researchers about her beliefs or her internal working model of her infant are two complementary objects of study, both of which serve to elucidate maternal sensitivity toward her infant.

The design of the current study draws on the literature on maternal behavioral sensitivity, or interactive behavior, and on maternal reflective functioning, thought to be a key component underlying the maternal behavioral sensitivity. The literature on the consequences of maternal interactive behavior for child outcome is reviewed below, illustrating the importance of studying maternal interactive behavior. This is followed by a review of the literature linking maternal interactive behavior and child temperament; the literature linking maternal interactive behavior and maternal characteristics; and the literature describing maternal infant play interactions. Together, these literatures inform the design of the current study, which explores the relationships among maternal interactive behavior (as assessed in mother-infant play interactions), child temperament, and maternal temperament. Finally, the literature on maternal reflective functioning and its relationship to maternal interactive behavior is reviewed, completing the review of the literature on which the current study is based.

Maternal Interactive Behavior

The consequences of maternal interactive behavior for child outcome. Both discrete maternal behaviors, over time, and more global maternal interactive styles have been shown to play a determining role in children's social and cognitive developmental outcomes, illustrating the critical importance of maternal behavior for child development. A general finding, repeated in study after study, links maternal behavioral sensitivity or

responsivity during the child's infancy to positive outcomes later in childhood in child language, cognitive, and social development (van den Boom, 1995; Elardo, Bradley, & Caldwell, 1975; Plomin & DeFries, 1983; Donovan & Leavitt, 1978; De Wolff & van Ijzendoorn, 1997). Behavioral responsivity can be defined as sensitive maternal reaction to child cues, for example comforting a crying baby, responding to infant vocalizations, or toning down the stimulation level of a play episode when the infant averts its gaze or shows signs of agitation. In an extremely far-reaching finding, maternal responsivity at 6 months has been shown to be related to teacher evaluation of children at 10 years of age (Bradley, Caldwell, & Rock, 1988).

Many studies have considered the influence of maternal behavior on developmentally disabled or at-risk children because, as Sameroff and Chandler (1975) argue, the impact of the environment, and particularly of maternal behavior, is amplified in this population. Consistently responsive maternal behavior throughout infancy and toddlerhood was related to leaps to near normal levels in cognitive development at two years for a subset of perinatally brain-injured children who had previously lagged behind their non-injured peers in a longitudinal study. This difference was maintained when the children were tested at five and eight years of age (Beckwith & Parmelee, 1986). The quality of maternal interactive behavior, as scored by a global behavioral rating scale, accounted for up to 25% of the variance in child Stanford-Binet scores at 24 and 36 months for children in a child-focused early intervention program which did not directly target maternal interactive behavior as part of the intervention (Wheeden, Mahoney, Fewell, & Spiker, 1997). These studies clearly indicate a large role for maternal interactive behavior in determining child outcome.

Mother-infant attachment is considered to have important and far-reaching social and cognitive development implications for children. Attachment is generally assessed at one year through the Strange Situation Procedure (Ainsworth, Blehar, Waters, & Wall, 1978), which involves a series of separations and reunions between mother, stranger, and child. The child behavior upon reunion with the mother is considered paramount. Maternal sensitive responsiveness to infant signals is seen by many as the key determinant of infant attachment status (Ainsworth, et al., 1978). Mother-infant dyads observed as early as 3 months were seen to exhibit behavior that was later found to be predictive of secure attachments (Isabella & Belsky, 1991). Dyads who were developing secure attachments were observed to have highly synchronous interactions, which appeared to be mutually rewarding; mothers of infants developing insecure attachments were observed to be minimally involved, unresponsive, or intrusive (Isabella & Belsky, 1991).

Maternal interactive behavior and child temperament. Past literature indicates a relationship between maternal interactive behavior and child temperament that the current study seeks to build upon. Infant temperament has generally been shown to be strongly related to maternal interactive behavior and to particular maternal behaviors (e.g., Milliones, 1978; Campbell, 1979; van den Boom, 1995), though some researchers have found no relationship between the two (e.g., Park, 2001). That there is a main effect of child temperament on maternal interactive behavior does not negate the possibility that there is also an interaction effect among maternal and child temperament and maternal interactive behavior, as per the goodness of fit hypothesis.

The focus of the original forays into temperament research, children with difficult temperament continue to be the focus of many studies. For instance, Milliones (1978) and Campbell (1979) found a significant negative relationship between maternal ratings of infant difficulty and observer-rated maternal responsiveness. Because of the correlational nature of the data, however, no conclusion can be drawn about causal direction.

A longitudinal study which addresses this problem of determining causation was conducted on a sample of Finnish mothers and infants at home in a free-play situation at 3 and 12 months (Kivijaervi, Voeten, Niemelae, Raeihae, Lertola, & Piha, 2001). Although neither maternal nor infant behavior were strongly related over time, there were strong correlations between mothers' sensitive behavior at 3 months and infant behavior at 12 months and between infant behavior at 3 months and maternal sensitive behavior at 12 months. This indicates that infant behavior influenced later maternal behavior and vice versa. Further analyses revealed that the strongest association was between infant happiness and positive mood at 3 months and maternal sensitive behavior at 12 months. The researchers conclude that positive infant aspects of mood, play, and social behavior seem to develop in parallel with maternal sensitivity in the first year (Kivijaervi, et al., 2001). This longitudinal study gives support to a transactional view of development.

An intervention study (van den Boom, 1995) found that toddlers rated as difficult in infancy show lower levels of cooperation with their mothers; however, this did not hold true for toddlers rated as difficult in infancy whose mothers were trained in responsiveness during the second half of the first year, indicating that the less-optimal interaction patterns of difficult infants and their mothers can be changed, and that this

change in interaction patterns will be reflected in child behavior. Indeed, intervention research can be a powerful tool in uncovering causal mechanisms and disentangling multiple determinants in child outcome.

For this intervention study, van den Boom (1994; 1995) selected infants high on irritability in the neonatal period, then provided the mothers with three, two-hour, in-home visits when their babies were between six and nine months of age. The visits focused on raising mothers' awareness of positive interactions with the child (because mothers of difficult infants tend to focus on the negative interactions) and on teaching the mothers to read and respond appropriately to the infants' cues. The infants included in the study scored higher on measures of sociability, self-soothing, and exploration than a control group, and spent less time crying. Their mothers were significantly more responsive, stimulating, visually attentive, and controlling of the infants' behavior at nine months than were those in the control group. Finally, the infants in the intervention group were significantly more likely to be securely attached at 12 months (van den Boom, 1994). Positive effects of the intervention continued to be seen through age 3 ½ years in follow-up studies (van den Boom, 1995). These studies suggest that although neonatal irritability (often considered a harbinger of later infant difficult temperament) can be related to poorer maternal interactive behavior and poorer child outcome, mothers of these children can be taught to change their interactive behavior, thereby changing the outcomes for the children.

However, not all studies have found a negative relationship between difficult child temperament and maternal behavior. For instance, mothers of difficult six-month-olds were found to be significantly more likely to exhibit affectionate contact, teaching

behavior, and object stimulation than mothers of easy babies (Bates, Olson, Petit, & Bayles, 1982). In pondering such paradoxical findings, Crockenberg (1986) posits that differences in temperament do influence caregiving, but that this influence is also dependent on characteristics of the caregiver and the caregiving environment. In explaining how infant negative emotionality, or difficulty, has been shown to both negatively and positively influence maternal behavior, Crockenberg (1986) notes that it could be that age differences are underlying the findings, with mothers gradually withdrawing from their difficult babies over time; or that there is a curvilinear relationship between infant difficulty and caregiving, in which moderately difficult infants elicit better caregiving but very difficult infants elicit maternal withdrawal; or there could be an interaction between temperament and a third variable, possibly maternal temperament, underlying the relationship. This third explanation parallels the goodness of fit hypothesis.

It is clear from the available research that maternal interactive behavior and child temperament are linked. While it is not possible to say which way the direction of influence flows—whether infant temperament influences parental behavior or vice versa—a transactional perspective negates the necessity of asking this question, because it posits that the two develop in tandem, mutually influencing one another. Some traits appear to have a more organic basis, such as smiling (Crockenberg & Acredolo, 1983) or activity level (Fish & Crockenberg, 1981), while others appear to be more influenced by maternal behavior in the first year, such as sociability (Fish & Crockenberg, 1981). Determination of infant temperament and maternal behavior is likely to be a fluid process, with influence flowing both ways, each partner in interaction offering both

stimulation and feedback to the other and slowly shaping each others' characteristic behavior over time. While the current study seeks to investigate the relationships among child temperament, maternal interactive behavior, and a third variable, maternal temperament, it draws on this earlier research into the direct relationship between maternal interactive behavior and child temperament.

Maternal interactive behavior and maternal characteristics. The role of maternal personality, or temperament, is addressed in the current study as part of the assessment of goodness of fit. The relationship of maternal personality and parenting behavior has been noted as a neglected area of research (Belsky, 1984), but some studies do exist; those that do often point to the powerful predictive utility of maternal personality on parenting behavior (e.g., Clark, Kochanska, & Ready, 2000). In a review of empirical studies on this relationship, Belsky and Barends (2002) report that, overall, child development outcomes appear to be best when parents are mature and psychologically healthy, including being high in self-esteem and having an internal locus of control.

Maternal self-efficacy has proven to be a mediating variable between infant proneness-to-distress and sensitive maternal behavior: when self-efficacy was moderately low or extremely high, high infant proneness-to-distress occurred with less sensitive maternal behavior, whereas when self-efficacy was moderately high, high infant proneness-to-distress was associated with more sensitive maternal behavior (Leerkes & Crockenberg, 2002). Such a study points to the complexity of relationships of the variables involved in determining parenting behavior.

Additionally, socio-economic status has long been considered as a determinant of parenting (e.g., Luster & Okagaki, 1993). Recent research indicates that maternal

education level accounts for most of the effect of socio-economic status on child behavioral outcomes during infancy (Bornstein, Hahn, Suwalsky, & Haynes, 2003). The current study promises to further the research linking maternal characteristics and maternal interactive behavior.

Mother-infant play interactions. For the design of its investigation, the current study draws upon the literature on mother-infant play. Observation of mother-infant play is considered to be an excellent way of capturing maternal interactive behavior. Mother-infant play is defined as those moments in which the attention of one or both partners is on the social behaviors of the other, and response to these social behaviors involves social behavior from the other partner (Stern, 1977). It is a concept bounded by culture, for the idea of a parent as a play partner arises out of Western conceptions of the parent-child relationship as friendly and sociable (Uzgiris & Raeff, 1995). Many infancy researchers have described the structure of mother-infant interaction during play periods, notably Daniel Stern (1977, 1983, 1985, 1990, 1995). From earliest infancy, play periods consist of two distinct parts: episodes of engagement and time-out episodes. These two types of episodes alternate, some lasting seconds, some minutes, but both serving to regulate the interaction so that it stays within the boundaries of the optimal ranges of stimulation. The time-out episodes occur in response to the breaching of a stimulation limit and provide a chance to adjust the level and direction of attention, excitement, and affect so that the play period may go on.

Play changes as infants develop, with simple face-to-face social interaction evolving into play with objects around four to six months (Uzgiris & Raeff, 1995). There is an inverted U-shaped curve of optimal stimulus level/attention for all play interactions:

low level stimulus does not get attention, high level stimulus gets an aversive reaction, and attention is best at midrange (Stern, 1977). Bell (1974) speaks of upper limit parental controls, intended to redirect or reduce infant behavior, and lower limit controls, which occur in response to insufficient or delayed infant behavior.

Stern (1977, 1983, 1985, 1990, 1995) describes mother-infant play episodes as becoming smoother and more complex over time as the mother-infant relationship develops. He likens their play to a structured dance, with a set tempo for each interval of interaction, and often synchronous movements. However, there are misregulations in the dance which can be the result of either partner. Maternal behavior that is controlling or intrusive is most often the cause of over-stimulation, because these maternal behaviors interfere with the infant's self-regulation. Similarly, under-stimulation can be caused by too little maternal behavior which is not of the intensity required to elicit a response from the baby. Stern notes that depressed mothers appear to understimulate their infants when interacting with them. An intensive study of five mother-infant pairs interacting over the first 20 weeks of life showed that one of the most important rules of interaction was the mother's sensitivity to the infant's signals for capacity of attention and nonattention (Brazelton, Koslowski, & Main, 1974). Although mothers are primarily the cause of misregulations in play interactions, infants can be the cause as well. For example, an infant who is hypoactive due to brain injury or Down Syndrome would be difficult to stimulate, while a cocaine-exposed infant who is extremely sensitive would be easy to over-stimulate.

In navigating their play interactions, mothers and infants must read, interpret, and respond to each other's cues in an elaborate dance. Evidence described above points to

the influence of infant and maternal characteristics on maternal sensitivity, a marker of the 'smoothness' of the interaction of the pair. The current research seeks to address whether the fit of maternal and infant temperaments influences this interaction.

Maternal Reflective Functioning

The term reflective functioning refers to the capacity to attribute mental states to others, thereby making others' behavior meaningful and predictable (Fonagy & Target, 1997). As such, it is a crucial component of any social interaction. The current study seeks to investigate the relationship of maternal reflective functioning to maternal behavior, as well as its relationship to the fit of maternal and infant temperaments.

Individuals differ in their capacity for reflectiveness, specifically in the extent to which they consider their own and others' beliefs, desires, plans, and the like in making sense of behavior; this is their reflective functioning (Fonagy & Target, 1997). The term reflective functioning was coined by Peter Fonagy and his colleagues (Fonagy, Steele, Steele, Leigh, Kennedy, Mattoon & Target, 1995) to describe these capacities; reflective functioning was first assessed in an empirical way on the basis of adults' responses to the Adult Attachment Interview (George, Kaplan, & Main, 1985), an interview with adults about their own childhood attachment experiences. Following up on Main's idea that a mother's metacognitive capacity predicts to her infant's attachment status, Fonagy and colleagues began to explore the relationship between parental reflective functioning and adult and child attachment security. They found that women who, when interviewed in the third trimester of pregnancy, could reflectively describe their own childhood attachment relationships to each parent were far more likely to subsequently have an infant who was securely attached at one year (Fonagy, Steele, & Steele, 1991). As the

researchers explained it, "predictive power resides, it seems, not in the quality of past experience but in the overall organization of mental structures underlying relationships and attachment-related issues" (Fonagy, Steele, & Steele, 1991, p. 901).

Mothers high in reflective functioning were also more likely to describe their childhood attachment experiences in coherent and collaborative ways, and were thus judged securely attached (Fonagy, et al., 1995). Yet a large part of the relationship between parental adult attachment classification and infant attachment classification remained unexplained; the mechanisms transmitting attachment status from mother to child were not understood.

Slade and colleagues sought to address this attachment "transmission gap" by looking directly at mothers' talk about their children, rather than about the mothers' own childhood experiences. The researchers believed that the assessment of maternal reflective functioning in the context of a narrative about the "live and ongoing relationship with the child" (Slade, et al., 2001, p. 5) would reveal more about the factors contributing to child attachment status. They used the Parent Development Interview (Aber, Slade, Berger, Bresgi, & Kaplan, 1985) to code maternal reflective functioning; this interview is a 45-item, semi-structured clinical interview which examines parents' representations of 1) themselves as parents, 2) their children, and 3) their relationships with their children. Slade and colleagues developed an adaptation of Fonagy's reflective function coding system for use with these parental development interviews (Slade, Bernbach, Grienenberger, Levy, & Locker, 2001).

Slade and colleagues assessed a sample of 40 mothers in their third trimester of pregnancy with the Adult Attachment Interview, and, when the infants were 10 months

old, with the Parent Development Interview; the infants were also assessed at 14 months with the Strange Situation Procedure (Ainsworth, Blehar, Waters, & Wall, 1978). It was found that maternal reflective functioning (measured when the infant was 10 months old) was highly predicted by maternal adult attachment status (measured in the third trimester of pregnancy) and, in turn, maternal reflective functioning predicted the quality of the infant's attachment (measured when the infant was 14 months old) (Slade, Grienenberger, Bernbach, Levy, & Locker, 2001).

In a related study, Grienenberger, Kelly, and Slade (2001) examined the relationship of maternal reflective functioning to maternal interactive behavior within the Strange Situation Procedure. Maternal interactive behavior was scored for disrupted affective communication. It was found that maternal reflective functioning was inversely related to disrupted maternal behavior. In addition, this disrupted maternal behavior was related to both adult attachment status and infant attachment status (Greinenberger, Kelly, & Slade, 2001). Finally, maternal reflective functioning and maternal disrupted behavior both appeared to act as mediators in the relationship between maternal adult attachment and infant attachment; when considered together, they explain about 25% of the variance in infant attachment security (Greinenberger, Kelly, & Slade, 2001). The relationship between maternal reflective functioning and maternal interactive behavior has important implications for the current study.

Fonagy and colleagues posit that "security of attachment is an indicator of that quality of infant-caregiver relationship which generates psychological understanding" (Fonagy & Target, 1997, p. 687). The ability of a mother high in reflective functioning to treat her child as a mentalizing being leads her to reflect back to the infant his emotions,

which in turn helps the child to come to understand both his own emotions and, on a larger scale, himself as a mentalizing being. Indeed, attachment security has been found to be related to later symbolic abilities and precocious mentalizing; this points to an underlying psychological understanding which is fostered by secure attachment and is passed along, in a sense, from mother to child (Fonagy & Target, 1997).

While the focus of the research described here has been mainly on the link between maternal reflective functioning and adult and child attachment status, some research may be seen as implicating a role for child temperament in that relationship. In an assessment of mothers' and fathers' adult attachment status, based on the Adult Attachment Interview, as well as child attachment security with both mother and father, based on the Strange Situation procedure, it was found that while adult attachment status of each parent predicted well to child attachment status with that parent, there was also a relationship between the child's attachment to the mother and the child's attachment to the father (Steele, Steele, & Fonagy, 1996). The authors suggest that this relationship could be explained by the influence of child temperament, in other words something within the child, that was constant in both relationships thereby accounting for the similarities in the child's attachment to the mother and the child's attachment to the father.

Maternal reflective functioning is an important component of maternal sensitivity distinct from her observable behavioral repertoire. The inclusion of a measure of reflective functioning in a study of maternal sensitivity provides a complement to observable phenomena and adds a naturalistic complexity to the picture of that maternal

behavior. This complexity is somewhat analogous to the relationship between beliefs and behavior.

Summary

The concept of goodness of fit is an inherently transactional concept, taking into account a recursive process in which child temperament affects parental behavior, which in turn affects child temperament, in an emergent process of development. Temperament is most often defined as the stylistic component of behavior, one of the major determinants of individual differences, an "active mechanism by which individuals adapt to their environment" (Van Heck, 1991, p. 173). It was originally conceived in a study of the antecedents of problematic behavior in childhood. The difficult temperament was defined as consisting of negative mood, intense mood expression, tendency to withdraw from new stimuli, low adaptability, and unpredictable rhythms of bodily function. However, the predictive value of a difficult temperament depends on the child's environment, particularly the family environment, and how the child's temperament fits within that environment. According to the over-arching hypothesis of this study, it is this goodness of fit which helps shape development.

Maternal perception of child temperament has been shown to be related to certain child characteristics, such as gender and birth order; to maternal characteristics, such as extroversion, negative and positive emotionality, and maternal depression; and to maternal interactive behavior, particularly responsiveness to child cues. As noted earlier, some argue that the biological purpose of infant temperament is to make infant behavior more coherent and therefore somewhat predictable for caregivers, so that they may tailor the stimulation they present to the infant. In this way, ideally, the infant receives an

appropriate level of stimulation. Understanding this as the function of temperament, it becomes clear that infant temperament is social-communicative; the dimensions of temperament should reflect this. Because the function of infant temperament is social-communicative, we can assume that temperament is expressed as behavior in the infant because it is through behavior that a social partner can read an infant's temperament.

Although researchers do not agree on the extent to which child temperament is a within-child construct versus a within-interaction construct, or a parental perception construct, a transactional view of temperament obviates the need to determine this. Maternal perception of child temperament will determine how she responds to the child, makes sense of the child's behavior, and predicts the child's behavior. This will have an important determining effect on child outcome and therefore merits study. Maternal questionnaire measures of rates of specific child behaviors provide the best way to measure maternal perception of child temperament.

Goodness of fit refers to the match of the child's temperament with the environmental demands. In infancy, the bulk of these environmental demands can be considered to lie within the mother-infant relationship because of the primacy of that relationship in the first year. The primary 'fit' in the mother-infant relationship is the fit of the mother's and infant's temperaments. A good fit could consist of the relationship between maternal and infant temperamental characteristics.

Maternal sensitivity has been shown to have broad and far-reaching influences on child outcome, both social and cognitive. Maternal sensitivity has both behavioral and psychological aspects. Behaviorally, maternal sensitivity can be defined as maternal interactive behavior. Underlying this behavioral sensitivity is maternal reflective

functioning, or a mother's ability to understand the link between her infant's behavior and his or her underlying mental states.

Mothers who have a good fit with their infants should enjoy high quality interactions with them; this is the implication of a good fit. Thus, maternal sensitivity can be used as a marker of a good fit. In this way, patterns of maternal and infant temperament which constitute a good fit can be discerned by determining the relationships of the various temperamental dimensions to maternal sensitivity in mother-infant pairs.

Research Aims

The aim of this research is to begin to define patterns of infant and maternal temperament which, using maternal sensitivity as a marker of good fit, can be shown to constitute a good fit between mother and infant. In this study, relations were examined between 1) maternal perception of her own and her infant's temperaments and 2) maternal sensitivity. Infant and mother temperament were measured using established questionnaires completed by the mother. Maternal sensitivity was assessed through behavior as an observer-rated, time-sampled observation and through reflective functioning as an analysis of maternal use of mental state terms in maternal reports of highpoints and lowpoints in interaction with their infants.

The central research question of this study is: Can we begin to define a good fit by discerning patterns of maternal and child temperaments that are related to maternal sensitivity, measured as both behavior and reflective functioning?

Hypotheses and Predictions

I. When mother and infant temperaments are both 'easy' (*i.e.*, approaching, adaptable, positive mood, low intensity), maternal sensitivity will be more optimal as measured through behavior and reflective functioning.

When infant temperament scores for approach, adaptability, and mood are relatively high and scores for intensity are relatively low in conjunction with relatively high maternal temperament scores for approach, flexibility, and mood,

Ia) the scores for the quality of interaction will be relatively high, and

Ib) the scores for maternal reflective functioning will be relatively high.

II. When mother and infant temperaments are both 'difficult' (*i.e.*, withdrawing, non-adaptable, negative mood, and high intensity), maternal sensitivity will be less optimal, as measured through behavior and reflective functioning.

When the infant temperament scores for approach, adaptability, and mood are relatively low and scores for intensity are relatively high in conjunction with relatively low maternal scores for approach, flexibility, and mood,

IIa) the scores for the quality of interaction will be relatively low, and

IIb) the scores for maternal reflective functioning will be relatively low.

III. When a mother-infant pair consists of a mother who is more rigid and less adaptable and an infant who is generally negative in mood, maternal sensitivity is expected to be less optimal than other mother-infant pairs as measured through behavior and reflective functioning.

When the infant score for mood is low and the maternal score for rigidity is high,

IIIa) the score for quality of interaction will be low, and

IIIb) the score for maternal reflective functioning will be low.

IV. Maternal reflective functioning and the quality of interaction are expected to be positively related, such that a mother who is high in reflective functioning will also be observed to engage in high-quality interactions, and vice-versa.

IVa) Scores for maternal reflective functioning will be positively correlated with scores for quality of interaction.

II. Methods

Participants

Thirty-four first-time mothers participated in this study. They were recruited through visits of the investigator to mothers' support groups, through fliers advertising the study, and through word-of-mouth. When mothers were first contacted, they were screened to see if they fit the selection criteria. These included: that this was the first baby born to this mother; that English was the first language between mother and baby; that the baby was born full-term and with no perinatal complications; that the baby was healthy and normally developing; and that the baby was seven months old or less. For mothers who met the criteria, a date was set for the investigator to visit the participant in her home while the baby was between four and seven months old.

The majority of mothers were white, old for primiparous mothers, and well educated; 12% (or four) of the mothers classified themselves as other than white, non-Hispanic. Of these mothers, two classified themselves as Hispanic and two classified themselves as Asian. Mothers ranged in age from 29 to 42 years, with a mean of 34.5 years and a median of 34 years. All mothers had at least a college degree; 56% of the mothers also had a graduate degree. All but one of the mothers was married; that one mother had been artificially inseminated and was parenting without a partner. Sixty-two percent of the mothers were stay-at-home mothers at the time of the interview; 21% were working part time (less than 20 hours a week) and 17% were working full time (20 hours or more per week). There was no significant correlation between education and employment, so that mothers with graduate degrees were no more or less likely to be employed than were mothers with only college degrees ($r = .125$, $p = .481$). Fifty-six

percent of the mothers had participated in a parenting support group of some kind, although only six of them had been recruited from support groups. Half of the mothers had had experience with infants before becoming mothers (e.g., had cared for much younger siblings, nieces and/or nephews, or had babysitting experience).

Twenty of the infants were breast fed only (including those infants sometimes bottle-fed expressed breast milk); three were in transition and received both formula and breast milk; three were weaned from the breast after three months of age and were exclusively formula fed; and eight were weaned from the breast before three months of age and were exclusively formula fed; these groups were each made up of approximately equal numbers of boy and girl infants. Breastfeeding history was unrelated to maternal education ($r = -.170$, $p = .337$) and maternal employment ($r = -.196$, $p = .267$). In the total sample, half of the infants were boys and half were girls and they ranged in age on the day of visit from 122 to 240 days, or just over 4 months to 7.5 months. Two of the infants had begun to crawl before the day of the visit; none of the others were crawling yet. The infants' birth weights ranged from 6 pounds, 6 ounces to 9 pounds, 14 ounces; all had been born between 38 and 42 gestational weeks.

The data were collected in and around two major metropolitan areas. Fifty-three percent of the participants lived on the East coast of the United States and 47% of them lived on the West coast. However, both the East coast and West coast sample seemed to be itinerant and not representative of either area. Of the 14 West coast mothers who mentioned where they themselves had grown up, only six had grown up in the region in which they were now parenting; of the 12 East coast mothers who mentioned where they themselves had grown up, only five had grown up in the region in which they were now

parenting. Most mothers had moved to the region within the past year to five years and could not be considered representative of their region; in fact, two of the West coast mothers had recently moved from the East coast region included in this study. Therefore, it was determined that East/West comparisons of the data would be meaningless and this variable was left out of the analyses. Because of this, the two samples were treated as one. Additionally, there were no correlations between infant gender and any of the sample characteristics, independent variables, or dependent variables, so mother-infant pairs with boys and pairs with girls were treated as one sample. Although previous studies have noted differences in mother-infant interaction based on infant gender and based on an interaction effect of temperament by gender (e.g., Klein, 1984; Gordon, 1983), this was not found to be true in the current research, perhaps because the sample of 17 girls and 17 boys was too small to detect significant differences.

Measures

Demographics/Background Information

Family demographics were reported by the mother, including marital status, ethnic background, household occupants and ages, parental education levels, parental employment and occupations. Background information included mothers' previous experience with infants, participation in parenting support groups, pregnancy history, babies' birth weight, gestational age at birth, occurrence of perinatal complications, maternity leave arrangements, and infant feeding method (see Appendix A). Finally, the interviewer kept a record of the number of infant's cry bouts and total cry time during the home visit.

Revised Infant Temperament Questionnaire (RITQ)

The RITQ (Carey & McDevitt, 1978) consists of 95 questions about infant behavior. Mothers indicate the frequency with which their infant displays each behavior on a scale of one to six. Nine categories of temperament, are scored: his/her activity level, predictability of the child's schedule (rhythmicity), tendency to approach/withdraw from new situations, how adaptable the child is, how intense his/her reactions are, his/her sensory threshold, his/her mood, how distractible s/he is, and how persistent s/he is. These scores can range from a possible low of one to a possible high of six, with one always representing the least difficult expression of that dimension of temperament and six the most difficult.

This scale is the most widely used of all mother-report infant temperament ratings and therefore allows for maximum comparability with existing literature. Additionally, the RITQ has been suggested as an appropriate measure of temperament for studies focusing on the relationship of child temperament to parent-child interaction, because of its "demonstrated sensitivity to interactional parameters" (Hubert, Wachs, Peters-Martin, & Gandour, 1982, p. 581). The scale was revised in order to improve its psychometric properties; the scale was lengthened, the rating points were increased from 3 to 6 on each question, the order of questions was randomized, and more questions were reverse scored. The test-retest reliability for the whole scale is .86, while the internal consistency is .83 (Carey & McDevitt, 1978). The revised scale was standardized on 203 mothers of four- to eight-month-olds recruited from private pediatric practices, predominantly White, middle-class, but representing all social strata.

A maternal-report measure of temperament was chosen here rather than an observer rating of temperament because it is the maternal perception of infant temperament upon which the mother is basing her conception of the child and, hence, her interaction with the child. Therefore, a measure of maternal perception of infant temperament was necessary for this study.

Dimensions of Temperament Survey-Revised (DOTS-R)

The DOTS-R (Windle & Lerner, 1986a) is a 54-item questionnaire which asks respondents to rate each item from one to four to describe themselves. The items consist of basic behavior, such as feeling sleepy at the same time each night or feeling comfortable immediately in a new environment, and together the 54 items provide a profile of temperament, including activity level—general; activity level—sleep; approach-withdrawal; flexibility-rigidity; quality of mood; rhythmicity—sleep; rhythmicity—eating; rhythmicity—daily habits; task orientation—distractibility; task orientation—persistence. The scores range from a possible low of one to a possible high of four, with one representing the most difficult expression of that dimension of temperament and four representing the easiest.

The revised scale enjoys good internal consistency of the temperamental factors, with a Cronbach's alpha ranging from .78 to .89, except for Rhythmicity—Daily Habits which has an alpha of .62. Separate DOTS-R scales have also been constructed for preschoolers and elementary school children, and the three scales also show relatively high congruence for most of the temperament attributes across samples in pairwise comparisons; this speaks to its good construct validity. The scale offers maximum

congruence with the RITQ in terms of the dimensions of and conceptualization of temperament represented.

Parent/Caregiver Involvement Scale (PCIS)

The videotapes of maternal child interaction were coded using the Parent/Caregiver Involvement Scale (Farran, Kasari, Comfort, & Jay, 1986). This scale consists of 11 items of maternal interactive behavior rated for amount, quality, and appropriateness, and five ratings of the coder's general impression of the overall atmosphere of the interaction. For each of the 11 maternal interactive behavior dimensions, the amount rating indicates the frequency with which each type of behavior occurred; the quality rating has to do with whether the behavior was generally more harsh and/or cold in being carried out by the mother, or more warm and gentle; the appropriateness rating has to do with how closely matched the behavior was to the infant's developmental level, or how closely synchronized the behavior was to the infant's. The 11 maternal interactive behavior dimensions measured by the scale are: physical involvement, verbal involvement, responsiveness, play interaction, teaching behavior, control over child's activities, directives, relationship among activities begun by the adult, display of positive emotions, display of negative emotions, and goal setting or expectations. The five general impression rating areas are: the availability of the adult to the child, the acceptance and approval of the child as manifested by the adult, the atmosphere of caregiver-child interaction, the enjoyment of the child demonstrated by the adult, and the adult's provision of a learning environment. The coding system can be found in Appendix B. A training videotape and workbook complement the anchored, five-point scale scoring manual and serve to standardize procedures better than some

other maternal-child rating scales available to date. It has been used extensively (e.g., Gorman, Lourie, & Choudhury, 2001; Seifer, Schiller, Sameroff, Resnick, & Riordan, 1996; Simeonsson, Bailey, Huntington, & Comfort, 1986) and was designed for use in research addressing the effects of certain child characteristics on parental behavior.

Another advantage of this particular scale is that 'amount,' 'quality,' and 'appropriateness' are rated separately. Thus, the frequency of behavior (amount) is separate from the quality and appropriateness, which are the dimensions more closely related to maternal sensitivity. Maternal sensitivity is therefore not confounded with how active the mother is during the interaction.

The scale shows good reliability (as reviewed by Towle, Farran, & Comfort, 1998; and Munson & Odom, 1996). In a study including 71 subjects, the subscales showed inter-rater reliability of .53 to .92, based on generalizability (Munson & Odom, 1996). Another study involving 49 subjects reported intraclass correlations of .80 or above using all items but Teaching Behavior (Seifer, Schiller, Sameroff, Resnick, & Riordan, 1996). The same researchers report a test-retest reliability of .4 to .5 for single-time assessments and .77 to .81 across a three-month span (Seifer, et al., 1996). As for validity, Munson and Odom (1996) report that correlations suggest that the scale is related to actual behavior in the ways expected, and that the scale is based on a revision of a previously developed scale. The current investigator's previous experience with the scale has shown it to be sensitive to subtle differences in maternal behavior, revealing a significant difference in one aspect of behavior (use of directives) in very small samples of mothers of normally developing and developmentally disabled toddlers. This scale has

also been used in other studies of goodness of fit between mothers and children (e.g., Simeonsson, et al., 1986).

Maternal Reflective Functioning

Maternal reflective functioning (Slade, Bernbach, Grienenberger, Levy, & Locker, 2000) was coded according to the coding scheme developed for the Parent Development Interview (PDI; Aber, Slade, Berger, Bresgi, & Kaplan, 1985). It was coded based on transcriptions of mothers' audiotaped answers in the Highpoints/Lowpoints Interview that was developed for this study. For the Highpoints/Lowpoints Interview, mothers were asked to describe three highpoints and three lowpoints in their interactions with their babies in the previous weeks (see Appendix C). Probes requested that mothers demonstrate their reflective functioning capacity in discussing these highpoints and lowpoints (i.e., "What were you feeling?" and "What do you think the baby was feeling?").

The Highpoints/Lowpoints Interview was developed as a method for measuring maternal sensitivity from a different perspective than behavioral interaction. It was intended to complement and expand on the observational data obtained from the videotaped interaction. The decision to ask about highpoints and lowpoints was based on the investigator's experience with mothers of infants; it was thought that it would elicit important information about the mothers' conceptions of their relationships with their babies.

In reviewing the initial conceptualization of this study, Arietta Slade noted the similarity between the Highpoints/Lowpoints Interview and two questions on the PDI: "Describe a time in the last week when you and your baby really clicked" and "Describe

a time in the last week when you and your baby really weren't clicking." These questions from the PDI are similar but are intended to elicit a description of mutuality, the presence of which is key to a higher score for reflective functioning. It was expected that in asking for a more general description of highpoints and lowpoints a broader range of scores would be elicited; it should be noted that the PDI contains 45 questions and takes from 1 ½ to 2 hours to administer, giving parents ample chance to demonstrate a wide range of answers. The Highpoints/Lowpoints Interview has only two questions and takes only 10 to 15 minutes to administer.

The Highpoints/Lowpoints Interview was piloted on three mothers of infants in order to assess whether or not the interview would elicit responses which could be coded using the PDI reflective function coding scheme. The three interviews were transcribed and then reviewed by an author of the coding scheme. It was concluded that the Highpoints/Lowpoints Interview data could be coded using the PDI reflective function coding scheme.

Each description of a highpoint or lowpoint was coded, receiving a score ranging from –1 (repudiation of reflective function) through 5 (definite or ordinary reflective function) to 9 (full or exceptional reflective function). The overall transcript of the interview was then coded with the same scores. Please see Appendix D for examples of the coding system.

Reflective functioning has been reliably coded on the Parent Development Interview in the past. Four judges coding reflective functioning in Parent Development Interviews achieved solid inter-rater reliability as evaluated using intraclass correlation coefficients. The Intraclass Correlation Coefficient (2, k) ranged from .78 to .95 with a

mean of .88 for individual passage scores and an Intraclass Correlation Coefficient (2, k) of .87 for ratings of overall reflective functioning in these transcripts (A. Slade, personal communication, June 2001).

Procedure

A few days before the date for the home visit, a packet was mailed to each participant containing two questionnaires and a letter confirming the date and time of the home visit and asking the participant to complete the questionnaires before the visit. The questionnaires, which are described above, were the *Dimensions of Temperament Survey-Revised* (DOTS-R) and the *Revised Infant Temperament Questionnaire* (RITQ).

At the start of the visit, the mother was given two consent forms to read, one for herself and one for her infant (see Appendixes E and F). Any questions she had were addressed and she and the investigator signed the forms. She was then interviewed about demographic and background variables and asked to complete any incomplete portions of the questionnaires that had been mailed to the house; only two mothers needed time to complete questionnaires. Upon completion of the questionnaires, the mother was asked to recount three highpoints and three lowpoints she had experienced in interacting with her baby in the preceding two weeks. This interview was audiotaped. If she did not spontaneously mention the mental states of both herself and the baby in the recounting of each episode, she was prompted to do so by the follow-up question, "What were you feeling?" and/or, "What do you think [baby] was feeling?" (Please see Appendix C for the interview questions).

Next, the mother was asked to choose a room in which she usually played with the baby and the video camera was set up in that room. This was a relatively small

camcorder that recorded onto compact VHS-C videotapes. The recorder was set up on a tripod positioned to capture as wide an area as possible; this area was indicated to the mother before the recording began. The mother was asked to play with and interact with her baby in as natural a manner as possible while being videotaped for 30 minutes. The videotaped interaction was later scored using the Parent/Caregiver Involvement Scale (PCIS; Farran, Kasari, Comfort, & Jay, 1986). Twenty minutes is considered adequate observation time for being able to score the interaction according to the manual for the Parent/Caregiver Involvement Scale; therefore, videotaping was ended at the end of 20 minutes or between 20 and 30 minutes for the few infants who became very fussy before 30 minutes had elapsed. Three infants needed to be fed after the videotaping had begun; for these infants, the videotaping was restarted after they had eaten. When needed, the order of the visit was rearranged to fit the babies' needs, so that a baby who was alert but soon to need a nap was videotaped immediately, a mother needing to feed the baby was interviewed at that time, etc. Generally, mothers were interviewed first and then videotaped. During the videotaping, the investigator either left the room or, when that wasn't possible, turned her back to complete the scoring of the infant temperament questionnaire during the videotaping, so as to remain as inconspicuous as possible. The infant temperament scores were shared with and explained to the mothers at the very end of the visit.

Mothers were asked not to answer the phone or the doorbell during this 30 minutes and any other occupants of the house were asked to go to another room. Mothers who asked how they should interact were encouraged to do "what you usually do" and to "follow the baby's cues." Some mothers who were hesitant or expressed confusion or

nervousness about what to do were given some examples of what other mothers had done: "I've had mothers read, sing, play with toys, I've had babies on the couch, the floor, the bed, in an *Exersaucer* [common baby toy, in which baby is held upright and can bounce], under a *Gymini* [another common baby toy, in which baby lies prone with toys dangling from two crossed overhead bars], all kinds of things." Infant toys were not provided by the researcher, but were used by all mothers. Finally, the starting and ending times of the visit were noted, or the starting and ending times of the baby being awake, if they differed. The number of cry bouts was tallied and the total cry time was recorded with a stopwatch.

Coding

Maternal Interactive Behavior

The interaction videotapes were coded primarily by the investigator, with 24% of the tapes coded by a second coder trained to reliability with the first, as measured by Cohen's *Kappa.* Both were blind to the temperament scores of the mothers and infants at the time of coding (although the investigator had coded the infant temperament scores at the time of the home visits, from one to fourteen months had elapsed between the home visits and the interaction coding and she was not capable of remembering individual scores). The coding for one dimension of behavior proved to be unreliable; *Teaching Behavior* had an inter-rater reliability score of .42 using Cohen's *Kappa.* Though this score was significant at the .01 level, it fell so far below the scores for the other dimensions of behavior that it was dropped from the analyses. One dimension of behavior, *Negative Statements/Regard,* was so rarely used by the mothers in this study that it, too, was dropped from the analyses. The remaining nine dimensions of behavior

and five general impression ratings had an inter-rater reliability score of .77 using Cohen's *Kappa*, with individual scores ranging from .60 to .94; the data are presented in Table 1. Consensus was reached on all disagreements. After reaching reliability with the second coder, the investigator coded all remaining videotapes within two weeks.

Table 1.

Kappa *Statistics of Interrater Reliability for Behavioral Interaction Measure*

Behavioral Subscore	*n* of scores[a]	*Kappa* value (Standard Error)	Approximate Significance
Overall Score	257	.77 (0.031)	< .001
Physical Interaction	25	.82 (0.094)	< .001
Verbal Interaction	24	.77 (0.107)	< .001
Responsiveness	25	.83 (0.089)	< .001
Play Interaction	25	.84 (0.086)	< .001
Control Over Interaction	25	.73 (0.104)	< .001
Use of Directives	22	.60 (0.125)	< .001
Relationship Among Activities begun by Adult	24	.63 (0.113)	< .001
Positive Regard	24	.72 (0.116)	< .001
Goal Setting	23	.94 (0.056)	< .001

General Impression[b]	40	.80 (0.075)	< .001

[a] Each interaction was scored for amount in each of these subscores; if amount was greater than 1, each interaction was also scored for quality and appropriateness in each of these subscores.

[b] General Impression consists of 5 individual subscores for each interaction, not scored for amount, quality, or appropriateness; the 5 subscores that make up General Impression are Availability, Acceptance, Atmosphere, Enjoyment, and Learning Environment.

Maternal Reflective Functioning

The Highpoints/Lowpoints interviews were transcribed from the audiotapes by the investigator and coded by a co-author of the *Addendum to the Reflective Functioning Scoring Manual for Use with the Parent Development Interview* (Slade, Bernbach, Grienenberger, Levy, & Locker, 2001). The coder was blind to the conditions and purposes of the study. Because the coder had helped to develop the scale and had been trained to extremely high reliability using this scale, it was determined that no specific reliability testing was necessary for these interviews because the coding process was exactly the same (A. Slade, personal communication, April, 2003). The coder assigned a single score to each recounting of a highpoint or lowpoint and an overall score for the entire transcript.

In consultation with Arietta Slade, who developed the *Addendum to the Reflective Functioning Scoring Manual for Use with the Parent Development Interview* (Slade, et al., 2001), which was used in the coding of the Highpoints/Lowpoints Interviews in this study, it was decided to use the sum of the two highest individual answers (called 'summary RF') for the analyses. This is distinct from earlier use of the overall RF score because the Highpoints/Lowpoints Interview was composed of only six potential answers, as compared with the 21-question Parent Development Interview for which the RF coding system was developed. The greater number of RF individual answer scores

necessitated the use of an overall score in previous studies; this is not the case in the current study. Mothers' optimal capacity for RF was what was of interest in the current study and this is reflected in the two highest individual scores. Mean score would include the lowest score, which was often the result of either a mother not being warmed up in the interview or of a mother running out of significant incidents to cite as either highpoints or lowpoints; it was felt that mean score would therefore not reflect mothers' optimal capacity for RF. Single highest score was not used because it only ranged from five to seven and did not seem to distinguish between mothers who were struggling with reflective functioning in the mother-infant context from other mothers. The sum of the two highest scores ranged from nine (top individual scores of four and five) to 13 (top individual scores of six and seven) with a much larger variance than the overall score; this made it a better candidate for including in parametric analyses. This summary RF score was closely linked to the overall RF score, but added depth to it by further distinguishing among mothers by virtue of its larger range.

Data Analysis

Mother and Infant Temperament Data

Infant temperaments were calculated using the Revised Infant Temperament Questionnaire (RITQ). These scores were converted to *z*-scores. Similarly, mothers' scores on the Dimensions of Temperament Scale-Revised (DOTS-R) were calculated. Because the DOTS-R indicates more difficult behavior with a lower score, while the RITQ indicates more difficult behavior with a higher score, the scores on the DOTS-R were reversed and converted to *z*-scores for comparison with the infant scores. In this way, mothers' and infants' scores indicated the difficulty of their temperament along a

continuum, from an easy temperament in any one dimension indicated with a negative *z*-score to a more difficult one indicated with a positive *z*-score.

In an effort to reduce the statistical 'noise' of including variables extraneous to the analyses, the nine infant temperament dimensions and seven maternal temperament dimensions measured by the RITQ and the DOTS-R were reduced to four and three dimensions, respectively. Because the hypotheses refer to the difficulty or easiness of temperament, the dimensions of temperament that define these states were the ones of interest. As noted in the introduction, a difficult temperament was originally defined in the NYLS study as consisting of non-adaptability, withdrawal, negative mood, intensity of reaction, and arhythmicity (Thomas, et al., 1963). Rhythmicity, however, was found to have strong cultural parameters, as discussed earlier (Thomas, et al., 1974). For the current study, therefore, adaptability, approach/withdrawal, mood, and intensity of reaction are included as the infant temperament variables. Similarly, the maternal temperament dimensions that were included in the analyses are Mood, Approach/Withdrawal, and Flexibility/Rigidity; the DOTS-R, used to assess maternal temperament, does not include a measure of intensity of reaction. A previous study used these same dimensions of the DOTS-R in considering temperament along a difficultness continuum (Mednick, et al., 1996).

The decision to reduce the number of variables was made a priori. Additionally, for some analyses, the temperament *z*-scores were aggregated and for some analyses, each of the seven individual dimensions of temperament *z*-scores was included. This is indicated in the description of each analysis. However, for all of the analyses, a more difficult temperament is shown by a higher score, whether composite or individual.

Maternal Interactive Behavior Data

The videotapes of interaction were scored by an observer using the Parent/Caregiver Involvement Scale (PCIS) and total scores for amount, quality, and appropriateness of the interaction were calculated by averaging the nine subscores for each dimension (omitting *Teaching Behavior* and *Negative Statements,* as previously mentioned, because the former had low inter-rater reliability and the latter was rarely seen in the videotapes). This yielded three summary scores (amount, quality, and appropriateness) between one and five for each mother, with one representing the least amount, lowest quality, and least appropriate level of interaction. For this research, seven subscales from the "quality" dimension and two subscales from the "appropriateness" dimension, along with the five general impression scores, were averaged to create a score for maternal interactive behavior because they were believed to be most closely related to the meaning of maternal behavioral sensitivity as it is used in this study. The variable representing this aggregate score is referred to as the 'quality of interaction.'

While 'quality' subscores were used for the aggregate score for most behavior dimensions, the 'appropriateness' score for *Verbal Involvement* and *Responsiveness* were used in place of the quality score because they come closer to the working definition of maternal sensitivity for the age of the infants used in this research. The quality score for *Verbal Involvement* is for the extent to which the adult adjusts the level of her language to ensure child's comprehension, which is difficult to judge with four- to seven-month-old infants. The appropriateness score is for the extent to which the adult uses language to comment on her own and the infant's activities during their interaction; this is very appropriate with four- to seven-month-old infants. The quality score for *Responsiveness*

is for the intensity of the response that the adult provides to infant cues. The appropriateness score is for the contingency or synchrony of the timing of the responses, an idea very close to the working definition of maternal sensitivity in this study. The five General Impression scores reflect the observer's overall impression of the mother-infant interaction and include scores for Availability (of mother to infant); Acceptance (of infant by mother); Atmosphere (harmoniousness of the interaction); Enjoyment (of the infant by the mother); and Learning Environment (extent to which the mother facilitated the infant's concentration on any one activity). Please see Table 2 for an overview of the variables used in this study.

Table 2.

Variables Used in the Analyses

Variable Type	Variable Name	Components	Source
Independent	Maternal Temperament	1. Approach/Withdrawal 2. Flexibility/Rigidity 3. Mood	Maternal report from the Dimension of Temperament-Revised (DOTS-R) Questionnaire
	Infant Temperament	1. Approach/Withdrawal 2. Adaptability 3. Intensity 4. Mood	Maternal report from the Revised Infant Temperament Questionnaire (RITQ)
Dependent	Quality of Interaction	1. Quality of physical interaction 2. Appropriateness of verbal interaction 3. Appropriateness of responsiveness 4. Quality of play interaction 5. Quality of control over interaction 6. Quality of directives 7. Quality of the relationship among activities begun by adult 8. Quality of positive regard 9. Quality of goal setting 10. Availability 11. Acceptance 12. Atmosphere 13. Enjoyment 14. Learning environment	Coded from the videotaped interaction using the Parent/Caregiver Involvement Scale (PCIS)
	Maternal Reflective Functioning	1. Scores on individual highpoints/lowpoints 2. Overall score	Coded from the Highpoints/Lowpoints interview

Statistical Analyses

Distributions were run for all independent and dependent variables. The dependent variables were analyzed for correlations, as were the maternal and infant temperament variables. The data were also analyzed for correlations between maternal and infant temperaments. A large correlation matrix was created to consider relationships among the various subscores. Demographic variables found to be correlated with the dependent variables were included as covariates in further analyses.

The first and second hypotheses, that when mother and infant temperaments are both 'easy,' quality of interaction and maternal reflective functioning will be more optimal and that when mother and infant temperaments are both 'difficult,' quality of interaction and maternal reflective functioning will be less optimal, were tested using a simple linear regression of each of the dependent variables on each of the seven dimensions of temperament *z*-scores. Covariates and hypothesized predictors were entered together and the Backward procedure was used to remove variables one at a time which were not contributing significantly to the equation's prediction value. Additional models not generated by the Backward procedure were tested as well, using the Bonferroni correction to adjust the *p* value according to the number of models considered. Finally, the ten highest and ten lowest difficult temperament mother-infant pairs' means on the dependent variables were contrasted using student *t*-tests. This comparison of the top third, or most difficult, and bottom third, or easiest, temperament participants based on questionnaire responses has been used in previous research (Gordon, 1983), albeit for children's scores alone, not a combination of mother and infant scores as is the case here.

The third hypothesis, that when a mother is rigid and an infant is negative in mood, quality of interaction and maternal reflective functioning will be less optimal, was tested using a regression analysis procedure like that described for the first and second hypotheses. The fourth and final hypothesis, that quality of interaction and maternal reflective functioning will be positively related, was tested using Pearson's product moment correlation to investigate a relationship between reflective functioning and quality of interaction.

Additional exploratory analyses were undertaken to better understand the data. A difficulty index was created for mothers and infants by summing their respective *z*-scores for each of the temperament dimensions included in the analyses. The absolute difference in these difficulty indexes was calculated, and its relationship to the overall quality of interaction and overall reflective functioning was calculated using regression analyses. In this absolute difference score, called the *Mismatch Score*, mothers and infants who differed greatly in their difficulty indexes had the highest mismatch scores while mothers and infants who were more similar had the lowest mismatch scores. Subsequently, the relative, not absolute, difference was analyzed the same way; this was calculated by subtracting the infant difficulty index from the maternal difficulty index[1]. The resulting scores were highest (i.e., positive) for very difficult mothers of substantially less difficult, in some cases easy, infants. The resulting scores were lowest (i.e., negative) for very difficult infants of substantially less difficult, in some cases easy, mothers. Mothers and infants who were closely matched in difficulty scored nearer the middle (i.e., close to zero) using this directional difference approach.

Because results indicated a counter-intuitive finding that mothers who rated themselves as more difficult than their infants had significantly higher quality interactions, further analyses were undertaken to understand this finding. One possible explanation was that these difficult mothers of easier infants were less socially dominant, engaged in more social monitoring, and had infants whose cues were easier to read. A set of seven behavioral interaction subscores was hypothesized as providing evidence for this interpretation; these seven behavioral interaction subscores were: 1) quality of control (how flexible the mother was in following through on her suggestions for play), 2) amount of responsiveness, 3) appropriateness of responsiveness (synchronicity with infant behavior), 4) amount of positive regard, 5) appropriateness of positive regard (timing in relation to infant behavior), 6) amount of control used in the interaction, and 7) amount of directives. Only one of these seven subscores had been included in the aggregate *quality of interaction* score; they were of interest here because of their hypothesized relationship to the above interpretation. These seven scores were factor analyzed using a principal components analysis and the difficulty index difference score was regressed on the resulting factors to better understand the relationship between these scores and the mother-infant pairs with high difficulty index difference scores[2].

---

[1] It is the difference in difficulty indexes between mother and infant that is of interest here. Subtracting maternal difficulty index from infant difficulty index would have been equally valid and would have yielded the same results, although they would have been inverted.

[2] Cattell (1978) recommends a ratio of 3 to 6 for number of variables in a factor analysis to number of participants in the research; the current study meets this criterium. MacCullum, Widaman, Zhang, and Hong (1999) argue that when the communalities of the variables comprising a factor are high (above 0.6), the impact of sample size is reduced; the communalities of the variables in the first factor here, used in further analyses, have an average communality of 0.68.

III. Results

Preliminary Analyses

The distributions for each variable were analyzed, including skewness and kurtosis; all were found to have acceptable distributions. Because no data were missing, all of the analyses discussed in this section have an *n* of 34, with the exception of any paternal variables; since one mother was parenting without a partner, all analyses involving a paternal variable had an *n* of 33. The independent variables were checked for correlations using the Pearson product moment correlation; it was found that the maternal temperament scores used in the analyses (flexibility/rigidity, approach/withdrawal, and mood) were highly intercorrelated, as were the infant temperament scores used in the analyses (approach/withdrawal, adaptability, intensity, and mood) such that those with high scores in one temperament dimension were likely to have high scores in the other temperament dimensions. Questionnaires such as the RITQ and the DOTS-R are designed to have dimensions that are correlated because that contributes to the assessment of their validity by indicating that they are measuring a single construct. Notably, however, the mother and infant scores were not significantly correlated with each other, indicating that mothers were rating infants' temperaments independently of their own temperaments (see Table 3 for the *r* scores and significance levels). This does not rule out that other maternal qualities or characteristics did not influence the maternal ratings of infant temperament; for this reason, 'infant temperament' in this study is understood to mean maternal perception of infant temperament.

Table 3.

Intercorrelations Among Maternal and Infant Temperament Variables

	Maternal Temperament Variables *r* values			Infant Temperament Variables *r* values			
Temperament	Approach	Flexibility	Mood	Approach	Adapt.	Intensity	Mood
Maternal Approach	-	.580**	.373*	ns	ns	ns	ns
Maternal Flexibility	.580**	-	.345*	ns	ns	ns	ns
Maternal Mood	.373*	.345*	-	ns	ns	ns	ns
Infant Approach	ns	ns	ns	-	.599**	.339*	.464**
Infant Adapt.	ns	ns	ns	.599**	-	. ns	.530**
Infant Intensity	ns	ns	ns	.339*	ns	-	ns
Infant Mood	ns	ns	ns	.464**	.530**	ns	-

* $p < .05$ ** $p < .01$

Demographic variables, including infant gender, geographic location (East coast versus West coast), infant age (in days), percentage of the home visit that the infant spent crying, maternal age, paternal age, birth weight, maternal education, maternal employment (unemployed, employed less than 20 hours a week, employed 20 hours a week or more), maternal ethnicity (White-non-Hispanic, or Other), maternal prior experience with infants, and infants' breastfeeding history (weaned [in part or completely] before three months, weaned after three months, and not weaned) were all examined for correlation with the two outcome variables (reflective functioning and quality of interaction) and the seven independent variables (maternal and infant temperament dimensions) using the Pearson product-moment correlation (see Appendixes G and H for correlation tables).

Maternal education and maternal age were each found to be slightly negatively correlated with the quality of interaction score, indicating that older and better educated mothers had lower-quality interactions than did their younger and less-educated counterparts, though maternal education and maternal age were not significantly correlated with one another. The mothers in this sample were all highly educated, with all mothers having at least a college degree and over half the sample having a graduate degree as well. So in this sample, college educated mothers tended to have higher quality interactions with their infants than did those mothers who held graduate degrees (r = -.360, p = .044). Mothers in this sample were also quite old for a primiparous sample, ranging in age from 29 to 42 years. In this sample, relatively younger mothers tended to have higher quality interactions with their infants than did the older mothers (r = -.346, p = .045). Of interest among the relationships among the background variables, maternal age and education were unrelated (r =.284, p = .103), as were maternal education and employment (r = .125, p = .481). While Zaslow, Pedersen, Suwalsky, and Rabinovich (1989) report differences in homemaker versus working mother interactions with their infants, this was not replicated in the current study (r = .011, p = .952 for the correlation of maternal employment with quality of interaction). However, Zaslow and colleagues found the differences in maternal interactive behavior only in evening visits with the father present and not in the daytime visits with only the mother and infant present; perhaps evening visits, because evening tends to be a more stressful time of day for both mothers and infants, would have elicited significant differences between the groups, although given the extremely low correlation found here it is not probable. Finally, summary reflective functioning, the sum of the two highest reflective functioning answer

scores, though unrelated to either maternal age (r = -.057, p= .750) or education (r= .001, p = .997), was found to be correlated with breastfeeding history. The longer mothers went without introducing formula, the more likely they were to have a higher summary reflective function score than were other mothers less dedicated to breast feeding (r = .377, p = .028). Quality of interaction was unrelated to breastfeeding history (r = .042, p = .813). Geographic location was slightly correlated with two dimensions of infant temperament, approach (r = -.348, p = .044) and intensity (r = -.350, p = .044), such that mothers living on the West coast rated their infants as less withdrawing and intense. However, due to the itinerant nature of the samples, this variable was not included as a covariate in further analyses because it was thought that it had no explanatory value.

Means and standard deviations on the temperament variables were compared to normative means. However, the only normative means for the Dimensions of Temperament Survey-Revised (DOTS-R), used here for measuring maternal temperament, are for a group of older adolescent girls, not adult women. The age discrepancy may account for some of the differences in means. Overall, the women in this sample are slightly more approaching, less flexible, and report an overall better mood than the adolescents on whom the questionnaire was standardized. The scores on the Revised Infant Temperament Questionnaire (RITQ) were very similar to normative means, though the infants in the sample tend to be slightly more adaptable, less intense, and express overall better mood. All of these slight differences may reflect a self-selection tendency among the mothers who decided to participate in the study, such that more approaching, more positive mood women with slightly easier babies may have been more likely to invite a developmental psychologist to observe their interactions with their

infants. For both mothers and infants, the differences from normative means are so slight that they appear to come from the same population (Table 4).

Table 4.

Means and Standard Deviations for RITQ[a] *and DOTS-R*[b]

Variable	Mean (SD) for this sample	Normative Mean (SD)
RITQ Approach[c]	2.27 (.72)[d]	2.27 (.78)[d]
RITQ Adaptability[c]	1.93 (.52)[d]	2.02 (.59)[d]
RITQ Intensity[c]	3.25 (.73)[d]	3.42 (.71)[d]
RITQ Mood[c]	2.76 (.63)[d]	2.81 (.68)[d]
DOTS-R Approach[e]	2.99 (.54)[f]	2.87 (.51)[f]
DOTS-R Flexibility[e]	2.78 (.78)[f]	2.97 (.54)[f]
DOTS-R Mood[e]	3.62 (.47)[f]	3.49 (.58)[f]

[a] Revised Infant Temperament Questionnaire
[b] Dimensions of Temperament Survey-Revised (for measuring maternal temperament)
[c] Possible range from 1 to 6
[d] Higher score indicates more difficult temperament (less approaching, less adaptable, more intense, more negative mood)
[e] Possible range from 1 to 4
[f] Higher score indicates less difficult temperament (more approaching, more flexible, more positive mood)

The aggregate quality of interaction scores derived from the Parent/Caregiver Involvement Scale (PCIS) ranged from 2.43 to 4.93 (a range of 2.50) out of a possible range of 1 to 5. It should be noted that a score of 1 would indicate a seriously adverse mother-infant interaction; none of this self-selected group of mothers' interaction fell into that category. While norms for the PCIS are not available, one of the authors of the scale suggested that means be compared to those in a previous study with the closest match for gender of parent and age and developmental status of child (D. Farran, personal

communication, June, 2003). Means and standard deviations for mean 'Quality' and 'Appropriateness' summary scores were compared to those reported in a study including a similar group of 49 White, married, mid- to high-socioeconomic status mothers with their normally developing six-month-old infants (Seifer, Schiller, Sameroff, Resnick, & Riordan, 1996). The mothers in the Seifer study were probably somewhat less well-educated (authors report only that all but one had completed high school) and were somewhat younger (mean age 29 years, as compared to 34.5 in the current study) than the mothers in the current study, which most likely accounts for their slightly lower mean scores; these two background variables, maternal age and education, were each found to be significantly related to mothers' interaction scores in the current study. While Seifer and colleagues (1996) report a mean score of 3.50 for quality and a standard deviation of 0.40, the current study found a mean score of 3.88 and a standard deviation of 0.78 for quality; for appropriateness, Seifer and colleagues (1996) report a mean score of 3.49 and a standard deviation of 0.40, while the current study found a mean score of 3.93 and a standard deviation of 0.77. In sum, the mothers of four- to seven-month-olds in the current study scored slightly higher on average, and more disparately overall, than a somewhat similar group of mothers with their normally developing six-month-olds (see Table 5 for distribution information for the aggregate score for the current study and see Appendix I for distribution information for each subscore).

Reflective functioning (RF) had a typical range for normal parents of three to six, out of a possible negative one to nine. Participants were prompted during the Highpoints/Lowpoints Interview to talk about their own and their infants' mental states, making a score below a three unlikely. An interview receiving a score of three contains

some suggestion of mental states, but does not make reflective functioning explicit; a score below that would be for an interview in which mental states could be inferred, at best, or were missing or repudiated, all highly unlikely in the context of the Highpoints/Lowpoints Interview. Individual answers on the Highpoints/Lowpoints Interview ranged in score from two to seven; four mothers gave only two examples of highpoints and five mothers gave only two examples of lowpoints in the interview for a total of nine missing answers. As discussed in the Methods section, each interview received an *Overall Reflective Functioning* score which was the single score for the entire transcript assigned by the coder. In addition, each interview received a *Summary Reflective Functioning* score which was the sum of the two highest individual answers scores (see Table 5 for distributions of both the Overall RF and Summary RF scores and see Appendix J for distributions of each Highpoint and Lowpoint score).

Table 5.

Distributions for the Dependent Variable Scores

Variable	Minimum	Maximum	Mean (SD)	Median	Variance
Quality of Interaction Score	2.43	4.93	3.84 (0.77)	4.00	0.59
Overall Reflective Functioning Score	3	6	4.97 (0.91)	5	0.82
Summary Reflective Functioning[a]	9	13	10.79 (1.32)	10	1.74

[a] Summary RF is derived from the sum of the two highest individual RF answers

Direct Relationships between the Variables

Before the more complex hypothesized interactions among all the variables in the goodness of fit analyses were considered, the more simple, direct relationships between

the predictor and outcome variables were considered. To reduce the large number of potential correlations for these analyses, and thus the potential for Type I errors, aggregate *z*-scores were used for maternal temperament and infant temperament.

Direct Relationships between Infant Temperament and Quality of Interaction

Simple relationships between the infant temperament scores and quality of interaction were considered first. Summing the infant temperament scores to indicate overall level of 'difficult' temperament, an Infant Difficulty Index was created by adding the z-scores for approach, adaptability, intensity, and mood. This Infant Difficulty Index showed significant correlations with several dimensions of behavior measured during the interaction.

Overall, mothers who reported their infants as having difficult temperaments showed lower-quality interactions with them. The more difficult the infant's temperament (i.e., the more withdrawing, less adaptable, more intense, and more negative in mood), the less verbal interaction was seen with the mother ('amount verbal;' $r = -.421, p = .01$), the less the pair engaged in mutual play ('amount play;' $r = -.359, p < .05$), the fewer statements or demonstrations of positive regard the mother gave him or her ('amount positive regard;' $r = -.445, p < .01$), the less likely these statements or demonstrations of positive regard were contingent to the child's behavior ('appropriateness positive regard;' $r = -.352, p < .05$), and the less acceptance and approval the mother manifested of the child ('acceptance;' $r = -.417, p = .01$). Finally, the more difficult the infant's temperament, the less the mother interacted with the child overall ('mean amount;' $r = -.414, p = .015$). Table 6 provides an overview of these correlations.

Direct Relationships between Maternal Temperament and Quality of Interaction

Similarly, a Maternal Difficulty Index was created by summing the individual temperament *z*-scores for flexibility/rigidity, approach/withdrawal, and mood. Again, this index had several significant correlations with the interaction subscores. All of these correlations were counter-intuitive: overall, the more difficult the mother's temperament, the higher the quality of interaction with her infant. The more difficult the mother's temperament (i.e., the more rigid, more withdrawing, and more negative in mood), the more she responded to the child's initiations, verbalizations and/or distress ('amount responsiveness;' $r = .35, p < .05$); the better synchronized these responses were to the child's behavior ('appropriateness responsiveness;' $r = .368, p < .05$); the less the mother organized and controlled the child's activities ('amount control;' $r = -.377, p < .05$); and the more acceptance and approval of the child was manifested by the mother ('acceptance;' $r = .352, p < .05$). Again, Table 6 provides an overview of these correlations.

Table 6.

Correlations between Maternal / Infant Difficult Temperament and Maternal Interactive Behavior

Behavioral Subscore	Maternal Difficulty Index	Infant Difficulty Index
	r value	*r* value
Amount Responsiveness	.350*	ns
Approp. Responsiveness	.368*	ns
Amount Control	-.377*	ns
Acceptance	.352*	-.417*
Amount Verbal	ns	-.421**
Amount Play	ns	-.359*
Amount Positive Regard	ns	-.445**
Approp. Positive Regard	ns	-.352*
Mean Amount[a]	ns	-.414*

[a] Mean Amount is the average of the nine subscores for Amount
*p < .05. **p < .01

These correlational analyses indicate that more difficult mothers enjoyed higher quality interactions than did easier mothers, while infants reported to be easier by their mothers enjoyed higher quality interactions than did infants reported to be difficult by their mothers.

Direct Relationships between Maternal and Infant Temperament and Maternal Reflective Functioning

The relationship of maternal and infant difficulty indexes was tested independently for correlations with summary maternal reflective functioning (the sum of the two highest individual RF answers). No significant correlations were found between either the infant

difficulty index and maternal reflective functioning or between the maternal difficulty index and maternal reflective functioning.

These analyses of direct relationships between predictor and outcome variables do not tell us about any potential interaction effects of mother and infant temperaments on maternal sensitivity. To investigate this goodness of fit, the hypotheses were tested.

Hypothesis Testing

Relationships among Maternal and Infant Temperaments and Measures of Maternal Sensitivity

Hypothesis one, that when maternal and infant temperaments are easier, maternal sensitivity is higher, and hypothesis two, that when maternal and infant temperament are more difficult, maternal sensitivity is lower, are the inverse of one another. Because of this, they could be tested with the same regression analyses. The first prediction of each of these two hypotheses is quality of interaction. This was tested first. The second prediction of each of these two hypotheses is maternal reflective functioning. This was tested second.

As outlined in the Methods section, the first step in analyzing these hypotheses was to adjust the temperament scores so that maternal and infant scores were comparable, for ease of interpretation. The maternal temperament scores were reverse coded so that a high score in both maternal and infant temperaments would indicate a more 'difficult' temperament (i.e., withdrawing, non-adaptable, intense, and predominantly negative mood for infants; withdrawing, rigid, and predominantly negative mood for mothers). *Z*-scores were calculated for both maternal and infant temperament scores to adjust for the differences in scale between the two scores (maternal temperament scores could range

from one to four, while infant temperament scores could range from one to six). Once this was accomplished, the maternal and infant temperament scores were comparable.

In order to test the first prediction about the relationship of maternal and infant temperament to quality of interaction, the quality of interaction scores were regressed on all of the maternal and infant temperament *z*-scores. Because earlier correlational analyses suggested that maternal age and education either together or singly might serve as demographic covariates for the quality of interaction, both were entered into the equation. If either or both were not significantly contributing to the prediction, they were removed. Whereas in the earlier reported correlational analyses, the temperament scores were aggregated into single difficulty indexes to help reduce the number of correlations tested, here in the regression analyses each individual score could be included as a potential parameter in a single model. The Backward command was used, resulting in the step-by-step elimination of variables which were not contributing to the prediction value of the equation. The resulting regression equation indicated that quality of interaction can be predicted using one of the covariates, specifically maternal age, and one temperament score, specifically maternal flexibility. Thus, the younger and more rigid the mother, the higher the quality of the interaction (F [2, 31] = 4.695, p = .017; R^2 = .232). This model did not include an infant temperament variable.

Maternal education had been the last variable eliminated and infant adaptability had been the second-to-last variable eliminated in the Backward procedure described above. In an attempt to create a model that included both a maternal and an infant variable, two additional models not considered in the Backward procedure were tested. Each included the infant and maternal temperament variable that contributed the most to the prediction

value of the above model and one of the two covariates. The first model included maternal age, maternal flexibility, and infant adaptability, while the second model included maternal education, maternal flexibility, and infant adaptability.

The second model emerged as significant (F [3, 30] = 4.081, p = .015; R^2 = .290), though the parameter estimates were not; however, this lack of significance may be due to the small sample size and the relationship merits discussion and further investigation in future research. This model suggests that, when controlling for maternal education, mothers who are more rigid and who report their infants to be more adaptable have higher quality interactions with their infants than do other mothers. The parameter estimates for both the above-described models are in Table 7, with the narrowly rejected model for predicting quality of interaction from temperament scores first and the accepted model second. Because three models were tested here (the model generated by the Backward procedure and the two additional models), the Bonferroni correction was used to calculate that each of the F tests should have a p value less than or equal to 0.017; this is true for both of the models reported here.

Table 7.

Parameter Estimates for Regression Equations for Predicting Quality of Interaction

Model	Parameter	Parameter Estimate (standard error)	*t*-value	Significance Level
Rejected Model	(Constant)	5.674 (0.749)	7.574	.000
	Maternal education	-0.299 (0.120)	-2.480	.019
	Infant adaptability	-0.242 (0.119)	-2.030	.051
	Maternal flexibility/rigidity	0.207 (0.119)	1.748	.091
Accepted Model	(Constant)	6.959 (1.216)	5.721	.000
	Maternal age	-0.091 (0.035)	-2.578	.015
	Maternal flexibility/rigidity	0.264 (0.124)	2.134	.041

In order to test the second prediction about the relationship of temperament to reflective functioning, the summary reflective functioning scores were regressed on all of the maternal and infant temperament *z*-scores. Because earlier correlational analyses suggested that breastfeeding history might serve as a covariate for reflective functioning, it was entered into the equation. Again, a Backward procedure was used, eliminating those variables which were not contributing significantly to the prediction value of the equation. The resulting regression equation indicated that summary reflective functioning cannot be predicted from the maternal and infant temperament dimensions

used in this study. Breastfeeding history alone is the best predictor of summary reflective functioning (F [1, 32] = 5.301, p = .028; R^2 = .142). Thus, mothers who introduce formula earlier tend to have lower capacity for reflective functioning than do those mothers who continue to nurse and express breast milk.

As above, however, a second model of the relationships between temperament and reflective functioning, when controlling for breastfeeding history, emerged which was significant. Again, as above, this significant model had a not quite significant t-value for one parameter estimate. That model, the second-to-last one to emerge in the backward procedure, included parameter estimates for infant mood and breastfeeding history (F [2, 31] = 4.379, p = .021; R^2 = .220), indicating that perhaps with a larger sample one would find that, when controlling for breastfeeding history, mothers who are high in reflective functioning tend to experience their infants as more positive in mood. The parameter estimates for both of these models for predicting reflective functioning from temperament scores can be found in Table 8.

Table 8.

Parameter Estimates for Regression Equations for Predicting Summary Reflective Functioning

Model	Parameter	Parameter Estimate (standard error)	*t*-value	Significance Level
Rejected Model	(Constant)	9.526 (0.536)	17.768	.000
	Infant mood	-0.372 (0.211)	-1.763	.088
	Breastfeeding history	0.419 (0.163)	2.562	.015
Accepted Model	(Constant)	9.625 (0.550)	17.486	.000
	Breastfeeding history	0.386 (0.168)	2.302	.028

In a final analysis of the first two hypotheses, an alternative way of looking at the data was framed, in which the easiest mother-infant pairs and the most difficult mother-infant pairs could be compared. A difficulty index score was constructed by summing the four infant and three maternal temperament *z*-scores, with a high score indicating more difficult overall temperament of the mother-infant pair. This way, the ten most difficult and ten easiest temperament mother-infant pairs were identified and compared for differences in quality of interaction and reflective functioning. Student *t*-tests of the differences in the means of the high difficult (HD) versus low difficult (LD) groups revealed no significant differences on quality of interaction ($t = .711$, $p = .486$; mean for HD = 3.814, mean for LD = 4.057) or summary reflective functioning ($t = 1.255$, $p = .226$; mean for HD = 10.500, mean for LD = 11.300). Earlier analyses of the relationship

of infant difficult temperament to behavioral interaction had revealed a significant negative correlation between infant difficult temperament and amount of interaction. The amount of interaction, coded as part of the Parent/Caregiver Involvement Scale, was not captured by the aggregate quality of interaction score used in prior analyses of hypotheses one and two. Because of the earlier finding of the significant role of amount, versus quality, of interaction, the high and low difficulty groups described above were analyzed for differences in their mean score for amount of interaction. This analysis revealed that the low difficult pairs interacted more than did the high difficult pairs (t = 3.099, p = .006; mean for HD = 3.322, mean for LD = 3.756), indicating that easier mothers of easier infants interacted more with their infants during a free play interaction than did more difficult mothers of more difficult infants.

Thus, hypothesis one, that when mother and infant temperaments are both easy, maternal sensitivity will be more optimal as measured both through quality of interaction and maternal reflective functioning, was not supported. In addition, hypothesis two, the converse of hypothesis one, that when mother and infant temperaments are both difficult, maternal sensitivity will be less optimal as measured both through quality of interaction and maternal reflective functioning, was similarly not supported. However, evidence of relationships among the variables was found. Counter to the hypothesized relationship between easier temperaments and higher quality of interaction, mothers who were rigid (more difficult on the adaptability dimension of temperament indicating that they are slow to adapt to change) had higher quality interactions with their infants, when controlling for maternal age. Some evidence also pointed to a role for infant adaptability, such that more adaptable infants with more rigid mothers had higher quality interactions,

when controlling for maternal education level. Similarly, some evidence pointed to a relationship between reflective functioning and infant mood, such that mothers who were high in reflective functioning had infants whom they experienced as more positive in mood, when controlling for breastfeeding history. So while maternal rigidity and infant adaptability were related to high quality interaction, infant positive mood was related to maternal reflective functioning.

While a relationship between temperamental difficulty and *quality* of interaction was hypothesized, it was *amount* of interaction that was found to differentiate between high difficulty and low difficulty mother-infant pairs, indicating that though mothers who report themselves and their infants to be of easier temperaments do not differ significantly in the quality of their interaction from mothers who report themselves and their infants to be of more difficult temperaments, the easy pairs interact more than do the difficult pairs.

The third hypothesis, that when a mother-infant pair consists of a mother who is more rigid and less flexible and an infant who is generally negative in mood, maternal sensitivity will be less optimal than in other mother-infant pairs as measured both through quality of interaction and maternal reflective functioning, was not supported, either. Like hypothesis one and two, hypothesis three was tested with a regression model, analyzing the relationship of the temperament variables to the quality of interaction first and to maternal reflective functioning second.

The quality of interaction scores were regressed first on the z-scores for maternal adaptability and infant mood, using the Backward procedure and controlling for maternal age and education as discussed above. As before, the resulting regression equation

indicated that quality of interaction can be predicted using maternal age and adaptability, such that the younger and more rigid the mother, the higher the quality of the interaction (F [2, 31] = 4.695, p = .017; R^2 = .232); the t-value for the parameter estimate for infant mood was not significant in the model which included it. Maternal education appeared to play no role in the regression equation, either (see Table 7, "accepted model," above, for parameter estimates). Thus, once again, the data show that, when controlling for maternal age, more rigid mothers have higher quality interactions with their infants no matter what the infant's general mood.

Next, the relationship of infant mood and maternal adaptability to maternal reflective functioning were assessed. The summary reflective functioning scores were regressed on the two temperament variables, maternal adaptability and infant mood, controlling for breastfeeding history. No model predicting reflective functioning from infant mood or maternal adaptability emerged. Therefore, hypothesis three was not supported; indeed, evidence suggested that maternal rigidity, rather than flexibility, is related to higher quality interactions; no such relationship was found for reflective functioning.

Relationships between Quality of Interaction and Maternal Reflective Functioning

Hypothesis four, that maternal reflective functioning and the quality of interaction are positively related such that a mother who is high in reflective functioning will also be observed to engage in high-quality interactions, was not supported. A correlational analysis using the Pearson product-moment correlation revealed no relationship between the quality of interaction and reflective functioning scores (r = .021, p = .91). However, a significant correlation between overall reflective functioning and the interaction subscore for amount of physical contact emerged, indicating that mothers who showed a more

complex understanding of their own and their infants' mental states touched their infants more during interaction ($r = .455$, $p = .007$). This finding must be looked upon conditionally, for it is open to Type I error; because of the earlier relationship between amount of interaction and temperamental difficulty, the relationships between the overall and summary reflective functioning scores and the interaction subscores for *amount* were investigated. While overall (a single score assigned to the interview as a whole) rather than summary (the sum of the two highest individual answer scores) reflective functioning scores were significantly correlated with amount of physical contact, it is highlighted it here because it aids in the discussion of the relationship between reflective functioning and maternal interactive behavior which was hypothesized in hypothesis four. Additionally, the overall and summary reflective functioning scores were highly intercorrelated ($r = .858$, $p < .001$).

Further Analyses

Additional analyses were undertaken to further explore the relationship between more difficult maternal temperament and higher quality interaction. Because it had been different than expected, the relationship of mother and infant difficulty was considered. If a mother-infant match on easy temperament was not the best predictor of maternal sensitivity, what relationship between mother and infant temperaments was? In order to explore the possibilities, the overall temperament of each infant and its relationship to overall maternal temperament was considered.

As delineated above and in the Methods section, the Infant Difficulty Index is a sum of the *z*-scores for the four infant temperament dimensions; it gives an overall indication of how easy or difficult each infant's temperament is. Because *z*-scores indicate distance

and direction from the mean, difficult temperament dimension scores are positive and easy temperament scores are negative. So if, based on maternal report, an infant was scored as very difficult in all four dimensions, the highly positive *z*-scores for each dimension would create an extremely positive Infant Difficulty Index when added together; an infant who was unadaptable (difficult in the adaptability dimension) but otherwise easy would receive a much lower, probably negative, Infant Difficulty Index reflecting the sum of three negative and one positive *z*-scores. Similarly, a Maternal Difficulty Index was created by summing the individual temperament *z*-scores for flexibility/rigidity, approach/withdrawal, and mood. This Maternal Difficulty Index was, like the Infant Difficulty Index, negative for overall easy temperament mothers and positive for overall difficult temperament mothers. These Difficulty Index scores are not categorical but rather fall on a continuum, maintaining the full power of the scores in statistical analyses. A classic definition of a difficult temperament is a child who is high in difficulty in all four of the dimensions considered here; while some of the mothers and infants who participated in the current study fit that definition of difficult, the following analyses are not limited to considering only those mothers and infants because each participant has a Difficulty Index score which indicates how difficult (or easy) he or she is along a continuum. All temperament profiles are of interest in goodness of fit research, not just difficult temperaments.

The relationship of these two scores, the Infant Difficulty Index and the Maternal Difficulty Index, were used in subsequent analyses. Earlier analyses had indicated that reflective functioning and quality of interaction did not vary on the basis of easy-easy mother-infant pairs versus difficult-difficult pairs; it was not similarity in the difficulty

indexes of the pair that seemed to be related to maternal sensitivity. Furthermore, earlier analyses showed that while difficult maternal temperament was related to some specific elements of higher quality interaction, difficult infant temperament was the opposite; that is, difficult infant temperament was related to some elements of lower quality interaction. Neither infant difficulty index nor maternal difficulty index by itself was predictive of quality of interaction. Therefore, the difference in the difficulty indexes of the pair was of interest here.

As discussed in the Methods section, a Difficulty Index Difference score was created by subtracting the infant difficulty index from the maternal difficulty index, marking the difference in the Difficulty Index scores for the pair and showing how unalike the mother had rated them. For the first analyses, the absolute value of this score was considered. The absolute Difficulty Index Difference score is a measure of overall temperament match or mismatch between mother and infant, and for the purpose of clarity will be referred to as the Mismatch score. Mother-infant pairs who were very different in overall temperament, such as easy mother-difficult infant pairs and the reverse, had higher Mismatch scores while mother-infant pairs who were more similar in overall temperament, such as both difficult or both easy, had lower Mismatch scores. These Mismatch scores ranged from a low of 0.08 (for an almost equally difficult mother-infant pair) to a high of 7.57 (for a very difficult mother with a very easy baby).

In order to test whether or not mothers' and infants' differences in temperamental difficulty were related to maternal sensitivity, this Mismatch score was assessed for its relationship to quality of interaction and reflective functioning. The quality of interaction scores were first regressed on the Mismatch scores, using maternal age and education as

covariates. No significant models emerged from the analysis, indicating that mismatch versus match in maternal and infant temperamental difficulty was not related to the quality of interaction. So too, no significant models emerged when summary reflective functioning scores were regressed on the Mismatch score, using breastfeeding history as a covariate. This indicates that mismatch versus match in maternal and infant temperamental difficulty was not related to maternal reflective functioning. Therefore, it was not how similar or how different mothers and infants were in difficulty that was related to maternal sensitivity.

The next step entailed considering the differences between pairs in which mothers were easy and infants were difficult and pairs in which mothers were difficult and infants were easy. This was done by considering the directional Difficulty Index Difference score rather than the absolute Difficulty Index Difference, or Mismatch, score considered above. Here, scores ranged from –6.64 to +7.57, with the most difficult infants who were rated much more difficult than their mothers having the lowest scores and the most difficult mothers who rated themselves as much more difficult than their infants having the highest scores. In this way, the difficult-difficult and easy-easy pairs tended to fall in the middle; the easy mothers with difficult infants fell at the low end; and the difficult mothers with easy infants fell at the high end. However, the real measure here is how much more difficult than the infant the mother reports herself to be; to that end, the two most difficult mothers in the group had moderately difficult infants (two difficult-difficult pairs) and yet their Difficulty Indexes were so far above their infants' that they scored fifth and sixth highest on Difficulty Index Difference scores. For the most part, however, the higher scores reflect difficult mothers with easy infants. The lower scores reflect

difficult infants with easy mothers. The sample in the current study had more difficult mothers of easy infants than easy mothers of difficult infants, so that the relationships are clearer at the high end of the Difficulty Index Difference scores.

Given that difficult mothers and easy infants had each independently been related to higher quality interactions, the question of how these temperaments together influenced maternal sensitivity was the next reasonable one to ask. In order to assess this, the quality of interaction and maternal reflective functioning scores were regressed on the Difficulty Index Difference scores. First, a Backward procedure was used to create a model which best predicted quality of interaction from Difficulty Index Difference scores; maternal age and education were also entered as covariates. A significant model emerged including Difficulty Index Difference score and maternal age (F [2, 31] = 4.54, p = .019; R^2 = .227), indicating that, when controlling for the effect of maternal age, the mothers who reported themselves to be greatly more difficult than their infants demonstrated higher quality interactions than other mothers. Maternal education was not a significant predictor in this regression equation. Apparently then, when controlling for the effect of maternal age, difficult mothers of infants they found to be easier than themselves had higher quality interactions than did easy mothers of infants they found to be more difficult than themselves. No significant equation emerged describing the relationship of the Difficulty Index Difference score to reflective functioning. This indicates that how much more difficult a mother reports herself to be than she reports her infant to be is not related to the extent to which she takes her infant's and her own interior life into account when she is discussing their interaction.

Table 9.

Parameter Estimates for the Regression Equation for Predicting Quality of Interaction from Difficulty Index Difference

Parameter	Parameter Estimate (standard error)	*t*-value	Significance Level
(Constant)	6.632 (1.200)	5.53	< .001
Maternal Age	-0.081 (0.035)	-2.34	.026
Difficulty Index Difference	0.072 (0.035)	2.07	.047

This surprising finding, that mothers who reported themselves to be much more difficult than they reported their infants to be had the highest quality interactions, raised many questions. Further analyses were necessary in order to make sense of the finding. One hypothetical explanation for the relationship of these variables was that 'difficult' was a misnomer for these mothers who were, in fact, less approaching, less adaptable, and less exuberant, though not of negative overall mood. It was posited that these mothers, because of their more cautious temperaments, engaged in more self- and social-monitoring and as such were perhaps more alert to infants' cues, and less likely to dominate a social interaction. Their infants, whom they perceived to be easier than themselves, also likely had easier to read cues and generated more positive regard. If this interpretation is correct, several of the specific behaviors measured in the Parent/Caregiver Involvement Scale, though not necessarily included in the quality of interaction aggregate score, would be related to the Difficulty Index Difference score.

While such things as physical interaction, verbal interaction, mutual play, relationship among the activities begun by the mother, and maternal goal-setting would not likely be related to this increase in readability of infant cues, increased maternal

ability to read cues, and decreased tendency to dominate an interaction, other areas of interaction would theoretically be related. Certainly the *amount of responsiveness* of the mother to the infant would be an area to look at; this is a score which represents both the infant cues caught and the infant cues missed by the mothers. While the quality of the response, how gentle and sensitive it is, would not necessarily be related, the *appropriateness of the response*, or how well timed to infant behavior it is, would be because an observant mother who is picking up on more readable infant cues more readily would also time her responses to infant behavior.

If these introverted mothers have less of a tendency to dominate in social interaction, then the *amount of control* score should be lower, while the score for *quality* of that control, that is how flexible the mother is about implementing her own suggestions and the extent to which she takes the infant's reaction (which would be more readable) into account when controlling the interaction, would be expected to be higher. The score for the *appropriateness* of the control would not be expected to be related here, because it measures how well the amount of controlling the mother does corresponds with the infant's need for control, based on developmental level of the infant.

Similar to amount of control, the score for *amount of directives* these introverted mothers used with their readable infants would be expected to be lower, with these less dominating mothers telling or showing their infants what to do less of the time. But the intensity of these directives (quality of directives score) and reasonableness of them for the infant's abilities (appropriateness of directives scores) would not necessarily be related to maternal introversion and infant easiness.

Finally, the *amount of positive regard* the mother offers the infant would be expected to be higher, because, as described earlier, mothers showed more positive regard to infants they experience as easy. Also, the extent to which these instances of positive regard are contingent to infant behavior (*appropriateness of positive regard* score) would be expected to be related to maternal introversion and infant easiness because it is essentially a measure of a specific kind of responsivity.

These seven scores were factor analyzed using a principal component analysis in order to reduce their number so that their relationship to the Difficulty Index Difference score could be analyzed without raising the risk of a Type I error. Factor analysis revealed that the seven subscores fell into two factors. One consisted of the flexibility of maternal control (quality of control), the amount of responsiveness she showed the infant, the timing of her responses in relation to the infant's behavior (appropriateness of responsiveness), her demonstration of her positive regard for her infant (amount of positive regard), and the timing of that positive regard in relation to the infant's behavior (appropriateness of positive regard) (Factor 1, called 'Warm Responsiveness'); and the other consisted of the amount of controlling she did of the infant and the amount of directives used with the infant (Factor 2, called 'Control'). The two factors were negatively intercorrelated, as would be expected, so that mothers who were more controlling and used more directives with their infants were less flexible in their control, responded less to their infants, responded less contingently to infant behavior, demonstrated less positive regard to their infants, and these demonstrations of positive regard were less contingent to infant behavior.

Summary scores for these two factors were created next by summing the variables that made up each factor. In order to investigate whether Control and Warm Responsiveness were related to the Difficulty Index Difference score, and therefore could bolster the proposed explanation for the higher quality interactions seen in the difficult mother-easy infant pairs, two final regression analyses were performed, one for each summary score. Maternal age and education were again used as covariates. Difficulty Index Difference did not predict Factor 2, Control, by itself (F [1, 32] = 1.521, p = .226; R^2 = 0.045), or after controlling for maternal education and maternal age (F [3, 30] = 0.982, p = .415; R^2 = 0.089), indicating that difficult mothers of easy infants did not do less controlling of their interactions with their infants and did not use fewer directives than did other mothers. Overall, however, this sample of mothers were low in use of directives and amount of control and perhaps the small variance in these subscores made it difficult to detect differences. However, difficulty index difference by itself was a significant predictor of the first factor, Warm Responsiveness, (F [1, 32] = 8.074, p = .008; R^2 = 0.201), indicating that the more temperamentally difficult a mother reports herself to be than she reports her infant to be, the more flexible in control and responsive to the infant's cues she is, the more her responses are synchronized with her infant's behavior, the more she demonstrates her positive regard for her infant, and the more that positive regard is contingent to infant behavior. Table 10 shows the parameter estimates for both models.

Table 10.

Parameter Estimates for Regression Equations for Predicting Factors 1 and 2 from Difficulty Index Difference

Factor	Parameter	Parameter Estimate (standard error)	*t*-value	Significance Level
1. Warm Responsiveness	(Constant)	19.765 (0.645)	30.66	< .001
	Difficulty Index Difference	0.529 (0.186)	2.84	.008
2. Control	(Constant)	6.235 (0.244)	25.505	< .001
	Difficulty Index Difference	-0.087 (0.071)	-1.233	.226

IV. Discussion and Conclusion

In this study, the relationship of maternal ratings of infant and maternal temperament to maternal sensitivity, as measured through quality of interaction and maternal reflective functioning, were investigated in an exploration of the concept of goodness of fit. While expectations about a match of easy temperaments or difficult temperaments being correlated with high or low maternal sensitivity, respectively, were not met, several interesting relationships among temperament, quality of interaction, and maternal reflective functioning did emerge.

Two dimensions of maternal sensitivity were posited and measured: 1) maternal ability to respond to infant cues, measured here as quality of interaction and 2) maternal ability to read and interpret infants' cues, measured here as maternal reflective functioning. Quality of interaction was scored from videotaped observations of maternal-infant free play. Reflective functioning consists of a mother's ability to consider her own and her infant's mental states when talking about their interactions and was scored from interviews with mothers about highpoints and lowpoints in their interactions with their infants.

The findings for quality of interaction, and the relationship of those findings to previous research, will be discussed first. The findings for maternal reflective functioning, and the relationship of those findings to previous research, will be discussed next. Because there are many more findings about quality of interaction, this discussion is divided into several sections, covering direct relationships between infant temperament and quality of interaction, direct relationships between maternal temperament and quality of interaction, and relationships among maternal and infant temperament and quality of

interaction. Finally, some conclusions about the findings will be drawn and directions for future research will be discussed.

Quality of Interaction

Direct Relationships between Infant Temperament and Quality of Interaction

It should be noted that because of the correlational nature of the findings reported here, and the fact that the data were collected at one time point on one age group rather than longitudinally or cross-lagged, direction of causation cannot be discerned. It should be understood that each finding could indicate an influence of infant temperament on maternal sensitivity, an influence of maternal sensitivity on ratings of infant temperament, an underlying heritable trait impacting both mother and infant, an underlying maternal characteristic that influences both her behavior and her rating and/or perception of her infant's temperament, or some other complex relationship between the two variables, including a transactional relationship in which each is influencing the other. Whether maternal sensitivity impacts infant temperament or vice versa, the relationship of the two has an important bearing on child development. This will be discussed further in the Conclusion.

In addition, infant temperament scores were derived from maternal report. While we cannot know to what extent these temperament scores reflect actual infant behavior or maternal bias, what is important is that they reflect maternal perception of infant temperament; it is this perception that a mother responds to in interacting with and talking about her infant. While maternal perception of infant temperament is the focus of this study, for ease of expression in the following discussion, 'difficult infant' and the

like are often referred to; it should be understood that this represents maternal impression of difficult infant temperament.

The results of this investigation reiterate earlier findings that more difficult infant temperament is related to lower quality maternal interaction (e.g., Sanson & Rothbart, 1995; Milliones, 1978). Infants in the current study rated as difficult based on maternal report of infant behavior tended to receive less verbal interaction, which echoes a finding by Campbell (1979) with infants at three and eight months of age. Using the same infant temperament scale, Campbell also found that mothers of difficult infants vocalized less to them. Since maternal vocalization to infants, whether talking, reading, or singing, has been shown to be an important contributor to later child cognitive development (e.g., Bradley & Caldwell, 1984), these difficult infants are at risk for less-than-optimal development because their mothers vocalize to them less than do mothers of easier infants.

Mothers in the current study were found to use expressions of positive regard less contingently to infant behavior when their infants were of a more difficult temperament, as captured in the *Appropriateness of Positive Regard* interaction subscore. This measure reflected praise, positive feedback, kisses, hugs, and smiles. Of all of the subscores of the Parent/Caregiver Involvement Scale, this score, *Appropriateness of Positive Regard*, is the score that best detected intrusiveness. Mothers of more difficult infants more often used kisses and praise non-contingently which interrupted either the flow of the interaction or the infants' attention to a toy or task, in manner that could best be described as intrusive. Lee and Bates (1985) similarly found mothers of difficult children to be more intrusive in interaction with them, although while temperament was assessed at 6

and 13 months, the interaction observation took place when the infants were older, at 24 months. This kind of maternal behavior interferes with the flow of mother-infant interaction; intrusive maternal behavior is most often the cause of infant overstimulation because it interferes with the infant's self-regulation (Stern, 1983). Therefore, difficult infants are more likely to have play interactions with their mothers cut short because of their own distress reactions to being overstimulated by their mothers' intrusive behavior. In a longitudinal study, maternal intrusiveness was also a hallmark of mother-infant pairs at three and nine months who were developing insecure attachments as assessed at 12 months (Isabella & Belsky, 1991). Therefore, the more difficult infants in the current study, who are more likely to experience intrusive maternal behavior, are at risk for developing insecure attachments.

Mothers of more difficult infants were also found to be involved less in mutual play with their infants, in which mother and infant are engaged with the same toy or activity, than were other mothers; this score, captured by the *Amount of Play Interaction* subscore, is related to measures of joint attention, though refers to joint attention to play only. This may be a corollary to the finding that mothers of more difficult infants are more intrusive; perhaps episodes of mutual play were cut short by maternal intrusiveness and subsequent infant distress and these pairs had to spend time re-orienting before they could engage again in mutual play. Mothers of difficult infants in a previous study were found to be less stimulating and visually attentive in interaction with their infants than were mothers of difficult infants who had participated in an intervention aimed at increasing their responsiveness to their infants (van den Boom, 1994). Stimulation and visual attentiveness are two aspects of mutual play, so these two findings may be related.

Certainly, joint attention of mother and infant has been shown to be a crucial building-block of child cognitive development; the reduction in joint attention for these mother-difficult infant pairs, as seen in their reduced mutual play, puts these difficult infants at risk for less-than-optimal development.

Mothers of more difficult infants in the current study also expressed less positive regard overall toward their infants, meaning they showed fewer positive overtures and reactions to the infants such as praise, positive feedback, smiles, hugs, and kisses; this was captured in the *Amount of Positive Regard* interaction subscore. Mothers of difficult infants have been found to focus less on positive interactions with their infants and more on negative interactions with them (van den Boom, 1994) so perhaps they miss the opportunities to offer their infants positive feedback.

Of related interest for this study is the finding that infants rated as difficult by their mothers cried more at three months than other infants, but by eight months, still rated as difficult by their mothers, this difference in amount of crying had disappeared (Campbell, 1979). Perhaps this same phenomenon accounts for the lack of correlation in the current study between percentage of the visit that the infant spent crying and maternal rating of infant difficulty; mean age for the infants in the current study was almost six months, closer to Campbell's eight-month-old group. Perhaps the more difficult infants in the current study had outgrown their tendency to cry more, but continued to be rated as more difficult. Several mothers reported in the course of the interviews that their baby had been a 'crier' earlier in infancy. One mother, in particular, was surprised to find her son rated low in mood when his temperament scores were reviewed with her at the end of the

visit; she wondered aloud whether her son's high level of crying in his early infancy biased her answers about current behavior, echoing Campbell's (1979) finding.

Campbell (1979) also reported that mothers of difficult infants tended to spend more time away from the infants during a home-observation period, much as the mothers of more difficult infants in the current study interacted less with them overall than did mothers of easier infants; this was captured in the overall *Amount* of interaction summary score, the average of the nine amount subscores for each mother-infant pair. Interestingly, almost all of the direct relationships found here between infant temperament and maternal interactive behavior had to do with *amount* of behavior, not its quality or appropriateness. This has not been reported in previous studies; perhaps because the Parent/Caregiver Involvement Scale makes the distinction among three overarching aspects of parental behavior (*amount*, *quality*, and *appropriateness*), this distinction was made clear here. If mothers of difficult infants tend to focus more on their negative interactions with their infants and to focus less on their positive interactions with their infants (van den Boom, 1994), then it makes sense that they would interact less: if a mother believes any interaction will likely be negative, it is understandable that she will be less likely to enter into an interaction. A large, prospective study of first-born sons showed that lower levels of interaction between mother and infant are a precursor to an insecure attachment at one year (Isabella & Belsky, 1991); reduced interaction of mothers with their difficult infants may be one of the links between difficult infant temperament and insecure attachment.

Direct Relationships between Maternal Temperament and Quality of Interaction

As noted previously, the relationship between maternal temperament and maternal interactive behavior is a neglected area of research, with many fewer studies with which to compare the current findings than exist for the relationship of infant temperament to maternal interactive behavior. The parenting-personality literature that does exist points to a positive relationship between 'easy' personality variables and high-quality parenting or parent-child interaction (e.g., Mangelsdorf, Gunnar, Kestenbaum, Lang, & Andreas, 1990; Belsky, Crnic, & Woodworth, 1995); the findings of the current study appear to define a positive relationship between more 'difficult' maternal personality variables and higher quality interactions. The findings are briefly reviewed below and then the meaning of 'difficult' maternal personality is considered.

More difficult mothers in the current study tended to be more responsive to their infants' initiations, vocalizations, demands, and distress than were easier mothers. Their responses tended to be better synchronized with the infants' behavior and they tended to evidence more acceptance and approval of their infants. All of this bodes well for the developing mother-infant relationship; such characteristics of interaction are generally seen to be precursors to infant secure attachment at one year (Ainsworth, et al., 1978; Isabella & Belsky, 1991). Finally, they tended to control their infants' behavior less, though it appears that this was optimal; while too little control in a play interaction does not provide enough structure for an infant (because she or he has few independent skills to use in playing), too much control over-rides the infant's cues, interests, and attention, losing the important give-and-take element of mother-infant interaction.

These counter-intuitive findings beg the question, what is maternal 'difficult' temperament? The DOTS-R questionnaire that was used to assess maternal temperament defines the dimensions used here this way: Approach/Withdrawal is the mother's tendency to either approach or withdraw from new stimuli, settings, and situations; Flexibility/Rigidity refers to how long it takes a mother to get used to a change in the environment; and Mood is how high the mother perceives her general mood to be.

Mothers who participated in the current study had high moods overall; of those who scored relatively lower on the mood dimension, all but one mother had scores which fell closer to the 'positive general mood' end of the score than to the 'negative general mood' end of the score. The half-way point on the scale is a score of 2.5 and the one mother who scored below that mark had a score of 2.43, only 0.07 below the midpoint. These more 'difficult' mothers are generally of lower mood than the 'easier' mothers, but still experience a positive general mood. They could be described as less exuberant than the higher scoring mothers.

So the mothers who scored as more difficult are less approaching, less adaptable, and less exuberant, but not of negative general mood. Perhaps these less approaching, less exuberant mothers who take longer to adapt to change engage in more self- and social-monitoring. As such, they may have less of a tendency to dominate an interaction, and may be more alert to infants' cues. This could account for the findings that more difficult mothers had higher quality interactions with their infants, all of which could be explained through their increased openness to their infants' cues, overtures, attention, and reactions. This interpretation gains strength from the correlational findings just reported: they picked up on more cues and missed fewer cues than other mothers (more

responsive); they had developed a clear rhythm of interaction through their sensitivity to their infants' cues (better synchronized responses); and they allowed for more of the infants' input in the interaction (less control of the interaction). Each of these findings about more difficult mothers corroborates the idea that these mothers have an increased openness to their infants' cues.

Perhaps a tendency to withdraw from new people, places, and things and a tendency to take a long time to get used to change mark assets in the maternal role while marking detriments in other roles. The 'difficult temperament' construct was developed according to personality characteristics that parents found most difficult to deal with in their children; the reverse may not be true. That is, these temperament dimensions may not be difficult for infants to deal with in their parents. The characteristics of withdrawal and non-adaptability are perhaps dimensions of temperament that are detrimental to interaction with a mature social partner while being an asset in interaction with an infant. A woman who withdraws from new situations, people, and places may be well suited to interaction in her own home with an intimate other, her infant. Perhaps being rigid in relation to schedules and plans does not impact a woman's ability to be flexible in a more microscopic way, interacting with her infant. Indeed, flexibility has been pointed to as a key maternal quality in interacting with infants (Brazleton, Koslowski, & Main, 1974), but this can be understood to be an ability to follow the infant's cues rather than one's own agenda. This understanding of flexibility would fit with rather than contrast with the findings reported here. Additionally, it is generally understood to be beneficial to infants to follow a set schedule, so these rigid mothers may be enhancing their infants'

development by creating a predictable environment and in turn, responding to their more settled infants with higher quality interactions.

Yet another possible explanation for the surprising finding that more difficult mothers had higher quality interactions with their infants than did their less difficult peers comes from considering the self-report nature of the questionnaire used to assess maternal temperament. Perhaps the mothers who scored as more difficult simply answered the questionnaire more realistically; it is possible that they were less defensive because of relatively higher psychological stability and therefore able to rate statements such as, "I smile often" in a more realistic manner. It is possible that their relatively higher psychological stability contributes to their higher quality interactions with their infants. It is also possible that they were less anxious about being evaluated, which led them to answer the questionnaire more realistically; this lowered anxiety about being evaluated may have carried over into the videotaping context, as well, leading them to be more relaxed while interacting with their infants which, in turn, contributed to their higher quality of interaction.

Finally, it is possible that the relationship between maternal difficult temperament and quality of interaction is not linear. That is, it may be shaped more like an inverted U curve than a straight line, such that mothers low in difficulty are low in quality of interaction, those moderate in difficulty are high in quality of interaction, and those high in difficulty are again low in quality of interaction. It could be the case that the DOTS-R, used to assess maternal temperament, does not distinguish extremely difficult temperaments from moderately difficult temperaments. It is possible, even likely given the likelihood of some degree of self-selection among the participants, that no mothers of

extremely difficult temperament volunteered to participate in this study. If this were the case, the moderately difficult mothers seen here, when compared with the mothers who are low in difficulty, would have higher quality interactions and the relationship would appear to be linear: low quality interaction with lower difficulty mothers, high quality interactions with higher difficulty mothers. But if we extended the study to include extremely difficult mothers, this linear relationship may then be transformed into more of an inverted U curve. This interpretation of the findings is evocative of one of Crockenberg's explanations for the inconsistent findings about the relationship of infant temperament to maternal sensitivity (Crockenberg, 1986) in which she posits that perhaps moderately difficult infants elicit more maternal sensitivity while extremely difficult infants elicit less maternal sensitivity. This interpretation gains strength from the evidence that mothers in this sample were somewhat self-selected, in that their scores on mood and approach were slightly higher than the normative sample.

Relationships among Maternal and Infant Temperaments and Quality of Interaction

It was hypothesized that quality of interaction and maternal reflective functioning would be highest when mother and infant temperaments were both easy (i.e., approaching, flexible, and positive mood for mothers; approaching, adaptable, low intensity, and positive mood for infants); conversely, it was hypothesized that the quality of interaction and maternal reflective functioning would be lowest when mother and infant temperaments were both difficult. Neither of these hypotheses was supported. However, a comparison of the 10 most difficult and 10 easiest mother-infant pairs (as determined by summing the four infant temperament scores with the three maternal temperament score for an overall difficulty score for the pair) revealed that mothers in the

easy pairs interacted with their infants more than did mothers in the difficult pairs. This supports Buss and Plomin's (1984) assertion that a simple match does not define a good fit; while the match of easy temperaments appears beneficial, the match of difficult temperaments does not. That the temperamentally easiest pairs interacted more than did the temperamentally most difficult pairs reflects findings discussed earlier that easier infants received more maternal interaction than did more difficult infants, though no similar relationship was found to be true for easier and more difficult mothers. Therefore, it may be that the easier infants, or the maternal perception of infant easiness, are driving this particular difference. Further research is necessary to consider the contributions of each partner.

Support was found for a model of interaction in which more rigid mothers with infants higher in adaptability had higher quality interactions than did other mother-infant pairs, when controlling for maternal education. Interestingly, it is adaptability in both mother and infant that emerges as the key dimension of temperament, albeit in opposite valences. This finding encourages further research into the role of adaptability in mothers and infants in determining a good fit. Again, we must ponder why a 'difficult' temperament dimension (rigidity) would be an asset to mothers. Recalling that the maternal temperament scores for each dimension were highly inter-correlated, we can infer that mothers who are more rigid most likely score as 'difficult' in the other two dimensions of temperament measured here. So, again, these rigid mothers may be the self-monitoring mothers discussed above, who are better able to read and respond to their infants' cues. Again, as discussed above, rigid mothers may also be providing a

predictable environment for their infants that in turn has a positive impact on the interaction of the pair.

Additionally, these more adaptable infants are most likely contributing to the higher quality interaction by reinforcing maternal overtures with their tendency for easy-going reactions to change. An adaptable infant will be much more open to a new interaction, new game, or new toy than a non-adaptable infant, making most maternal overtures for play successful. A mother who tries to engage her non-adaptable infant by introducing a toy for mutual play is more likely to be rebuffed in her attempt, particularly if the infant is already engaged in an activity. A non-adaptable infant is one who perhaps balks at new foods, clings in new situations, and/or becomes distressed when a bottle is too warm or too cold. Therefore, a mother of a non-adaptable infant, because of her experience with her infant's tendency to reject new things, may be more cautious in interacting with her infant, and therefore engage in lower quality interactions with her infant.

Finally, when controlling for maternal age, a significant relationship between quality of interaction and mother and infant temperament was found, such that mother-infant pairs in which mothers reported themselves to be much more difficult than they reported their infants to be had higher quality interactions. Again, perhaps these more 'difficult' mothers engaged in more self- and social-monitoring and had less of a tendency to dominate in social interaction and were therefore better able to read their infants' cues. Indeed, the infants in this group are much less difficult than their mothers in temperament and it has been shown that easier infants give more obvious cues (van den Boom, 1994). The combination of an easier infant, whose cues are more readable, and a self-monitoring mother, who has less of a social interaction agenda and is therefore more open to the

infant's cues, could lead to a mother-infant pair that demonstrates a higher quality interaction than other mother-infant pairs.

As discussed in the Methods section, if the above supposition is true, it is reasonable that these mother-infant pairs would score higher in certain areas of interaction scored on the Parent/Caregiver Involvement Scale, not all of which were included in the aggregate quality of interaction score used in earlier analyses. Seven such interactive behavior subscores were determined: amount of responsiveness to infant cues, appropriateness of this responsiveness (synchrony with infant behavior), amount of control exerted by the mother over the interaction, quality of that control (flexibility in carrying out suggestions for interaction, based on infant reaction), amount of directives issued by the mother to the infant, amount of positive regard demonstrated by the mother of the infant, and appropriateness of that positive regard (synchrony with infant behavior).

As discussed in the Methods and Results sections, these seven behaviors were analyzed in order to test the above interpretation of why difficult mothers of much less difficult infants had the highest quality interactions. To recap, two summary factors were created and tested for their relationship to the Difficulty Index Difference score, or how much more difficult the mother reports herself to be than she reports her infant to be. The first factor, Warm Responsiveness, consisted of the flexibility of maternal control, the amount of responsiveness she showed the infant, the timing of her responses in relation to the infant's behavior, her demonstration of her positive regard for her infant, and the timing of these demonstrations of positive regard in relation to the infant's behavior. This factor was found to be strongly related to the difficulty index difference score, indicating that mothers who rated themselves much more difficult in temperament than

they rated their infants were demonstrating these five behaviors at a higher level than were the rest of the sample.

The second factor, Control, consisted of the amount of controlling the mothers did and their use of directives; this factor was negatively intercorrelated with the first, such that mothers who did more controlling and used more directives were less responsive, their responses were less synchronous with infant behavior, their control was less flexible, they demonstrated less positive regard, and their positive regard was used more intrusively. The second factor, combining amount of control and use of directives, was not found to be related to the difficulty index score, even when controlling for maternal education and maternal age. This indicates that mothers who rated themselves as much more difficult in temperament than they rated their infants did not engage in significantly less controlling and fewer directives than did the other mothers. A caveat to this finding is that levels of control and use of directives were fairly optimal overall for the sample, with very few mothers using more than developmentally appropriate; this homogeneity may account for the lack of differentiation here.

So the interpretation as to why difficult mothers of much less difficult infants have higher quality interactions finds some support in these analyses. As hypothesized, these difficult mothers of much less difficult infants are mothers who are more responsive to their infants' cues; whose responses to their infants are in good synchrony with their infants' behavior (they do not overwhelm them by responding too quickly, nor do they wait so long that there is little connection the infant can draw between his or her initiation and the mother's response); they are more flexible in organizing the infants' activities, controlling the interaction by suggesting without being overly insistent, and by adapting

their control according to the infants' reactions; they use more positive verbal statements and more non-verbal signs of positive regard; and they time this positive emotion appropriately to the infants' behavior, not intruding with kisses or praise.

Interestingly, one of the key differences in interaction that distinguishes these more difficult mothers of much less difficult infants is in the timing of their behavior. Other researchers have discussed the importance of timing in mother infant interactions (e.g., Vietze & Anderson, 1981). Perhaps because these mothers engage in more monitoring and have less of a social agenda, they watch their infant interaction partner more closely, their infant interaction partner gives more readable cues, and the mothers are therefore able to respond to more cues and to time their responses more appropriately. Also, despite being labeled 'rigid' by their responses to the temperament questionnaire, these mothers are more flexible in their control of their interactions with their infants: they gage their infants' reactions to their suggestions and adjust the interaction accordingly. Like mother infant dyads that are developing secure attachments, these dyads, too, show extremely well-timed, reciprocal, mutually rewarding interactions (Belsky & Isabella, 1991).

An additional explanation of the finding that more difficult mothers of less difficult infants had higher quality interactions comes from considering the fact that the mother is providing the information on both her own and her infant's typical behavior and reactions. Temperament here becomes not a dimension of the infant but a measure of the mother's perception of the dyad; by using maternal report to measure both her own and her infant's temperament, the measurement of the two temperaments takes on an interpersonal meaning. Mothers who rate themselves as much more difficult than their

infants are, in a sense, saying, "My baby is better than I am." Such mothers would be more likely to hold themselves accountable for any problems that arise in interaction with the infant or even in the relationship with the infant. Conversely, a mother who sees herself as easier than her infant would be likely to hold the infant accountable for any problems. This sense of maternal accountability may have a positive effect on mother-infant interaction in that a mother then may take responsibility for the interaction and may attribute a sort of control over the interaction to herself. This would in turn contribute to her parental locus of control, or the extent to which she believes herself to have control over interactions with her child, at least the negative ones (e.g., Campis, Lyman, & Prentice-Dunn, 1986). Perhaps this self-accountable orientation reflected in reporting herself to be more difficult in temperament than her infant leads a mother to monitor her interaction with her infant more closely. It is possible that this heightened monitoring results in increased responsiveness which in turn contributes to the overall higher quality of interaction seen in pairs in which the mother reports herself to be much more difficult in temperament than she reports her infant to be. At this point, this assumption is speculation; a search of the developmental literature revealed no related studies.

Finally, a look at the individual relationships between the quality of interaction scores and the difficulty index difference scores of the mother-infant pairs shows a clear trend at the high end of the difficulty index difference scores, where the difficult mothers of much less difficult infants lie, than it does toward the middle and low end. So while we can begin to draw conclusions about the mother-infant pairs that fall at the high end versus the rest of the mother-infant pairs, we cannot begin to do so about the pairs which

fall toward the middle (the easy-easy and difficult-difficult pairs) and at the low end (the easy mother-difficult infant pairs). This could be explained in part by the fact that easy mothers tend to rate their infants as less difficult (Bates, Freeland, & Lounsbury, 1979), reducing the number of easy mother/difficult infant pairs that ought to be anchoring the bottom end of the score. A larger sample size might address this issue, as would the inclusion of a confirmatory, observational measure of infant temperament.

While there were several findings about the relationships among maternal and infant temperaments and maternal interactive behavior, the hypothesized relationships were not supported. That is, mothers who reported their own and their infants' temperaments as being easier did not show higher quality interactions, nor did mothers who reported their own and their infants' temperaments as being more difficult show lower quality interactions. Maternal difficult temperament appears to be a misleading construct when applied to mother-infant interactions; it functions as an asset, particularly when paired with maternal perception of easy infant temperament.

Reflective Functioning

Direct Relationships between Infant Temperament and Maternal Reflective Functioning

The summary score for reflective functioning captured the highest maternal capacity for attributing mental states to herself and her infant; this attribution of underlying mental states aids a mother in making sense of behavior and interaction. There was some evidence for a link between summary reflective functioning and maternal ratings of infant mood, such that mothers who rated their infants' mood as more positive were more likely to have higher summary reflective functioning scores, when controlling for breastfeeding history. Infant mood was rated based on discrete behaviors, such as fussing or crying

during face washing, diapering, or examinations by the doctor. Mothers were asked to rate each item from one (almost never) to six (almost always). If a mother has a high capacity for reflective functioning, she will most likely posit the feelings that underlie her infant's reactions. This empathy the high reflective mother has for her infant may color her impression of the infant's behavior. In other words, if a mother understands that being handled naked by an unfamiliar adult (examination by a doctor) is somewhat frightening for her child, she may discredit the fussing that accompanies this procedure. When asked to report on it, she may rate the fussing lower because she understands what motivated it and was perhaps therefore less bothered by it. A mother who is lower in capacity for reflective functioning may lack such insight into the infant's fussing and may therefore find it more aversive. This might lead her to report the fussing higher. This would lead to a higher overall rating of negative mood for the infant. Thus, mothers who are high in reflective functioning may be seen to have infants high in positive mood, while mothers low in reflective functioning may be seen to have infants high in negative mood. Because of its strong correlation with reflective functioning, breastfeeding history must be controlled for before this relationship between maternal reflective functioning and maternal report of infant mood becomes apparent.

This relationship between reflective functioning and maternal report of infant mood is similar to a finding in a study that looked at infant colic. Twenty mothers who brought their infants to the pediatrician with a complaint of colic, or prolonged periods of unexplained crying, were compared with 20 mothers who brought their infants to the pediatrician for routine examinations. The mothers were asked to keep a diary of how much their infants cried, to complete an infant temperament questionnaire, and were

interviewed. Though there were no significant differences in the amount that the two groups of infants cried, the mothers who had presented their infants as having colic reported feelings of rejection and intense negative affect in response to the infants' crying. These same mothers also rated their infants as significantly higher in negative affect. What is evident in this study of colic is that it is not the infant's behavior that is crucial, but the mother's interpretation of that behavior (Pauli-Pott, Becker, & Mertesacker, 2000). It may be that maternal capacity for reflective functioning underlies this finding about colic; it seems to offer a parallel to the finding in the current study that mothers who are high in reflective functioning rated their infants as higher in positive mood (less aversive in behavior) probably because their *interpretation* of their infants' behavior differed from that of other mothers, as discussed above.

Of the hypothesized relationships among maternal and infant temperaments and maternal reflective functioning, only a relationship between infant mood and maternal reflective functioning was found. Maternal temperament appeared to be unrelated to reflective functioning; although a null finding, this is of note because no previous research had looked at the relationship of personality and reflective functioning. Obviously, because of small sample size and the homogeneity of the sample, confirmatory studies would be required before an absolute conclusion could be drawn about the relationship of personality to reflective functioning.

The hypothesized relationships among maternal and infant temperaments and maternal reflective functioning were not found. Because the complete Parent Development Interview was not used, the findings cannot be considered conclusive. However, the idea that mothers of easier temperaments who reported their infants to be of

easier temperaments would have higher reflective functioning found no support here, nor did the idea that mothers of more difficult temperaments who reported their infants to be of more difficult temperaments would have lower reflective functioning; in fact, no combination of maternal and infant temperaments was found to be related to maternal reflective functioning. Though the idea is appealing that it may be easier to attribute the thoughts and feelings underlying the behavior of a person you perceive to be similar to you than of a person you perceive to be dissimilar to you, this idea found no support in the current study. It would seem that aspects other than maternal perception of mother and infant temperament are related to maternal reflective functioning, with the possible exception of maternal perception of infant mood.

Relationships between Maternal Reflective Functioning and Quality of Interaction

Summary reflective functioning was found to be related to breastfeeding history, such that the higher a mother's capacity for reflective functioning, the longer she went without introducing formula. What does this relationship mean?

Breastfeeding is a selfless act, one in which mothers put their infants' needs before their own, particularly in the first three months. In the investigator's experience running support groups for new mothers, it has been seen that breastfeeding does not generally become a pleasurable act for the mother until the infant is about three months old. Before that time, it is a demanding, often uncomfortable, 24-hour-a-day job. Clearly, mothers who continue to breastfeed exclusively through the three month mark are those who weigh the unseen benefits to their child more heavily than their own inconvenience. This may be the clue that links breastfeeding history to summary reflective functioning: mothers who are dedicated to breastfeeding are able to take their infants' point of view.

Perhaps this ability to take their infants' point of view extends to both the decision to keep breastfeeding and the capacity for attributing mental states to the infant, thereby accounting for the relationship between the two that was found in the current sample.

Another pathway between maternal reflective functioning and breastfeeding history becomes apparent when considering a finding about the relationship of reflective functioning to observed interaction. Overall reflective functioning was positively related to amount of physical contact seen in the mother-infant free-play interaction, indicating that mothers who used more complex references to the thoughts and feelings underlying their own and their infants' behavior also touched their infants more during the free-play interaction. Here, a closer look at the individual cases elucidates a relationship that might be underlying the link between reflective functioning and amount of touch.

Fourteen mothers of the sample of 34 scored a two or a three for amount of physical interaction, indicating that they were in physical contact with their infants a moderate amount at best. In contrast, the seven mothers who received the highest score of five for amount of physical contact held or passively supported their infants for almost the entire interaction while also engaging in active touching of their infants, such as back rubbing, tickling, or kissing. It should be noted that amount of physical contact was not found to be correlated with the age of the baby. The 14 mothers who scored a two or three for amount of physical contact appeared to have an underlying goal of independence for their infants during the interaction. These mothers placed their infants away from themselves, and although they sometimes picked up their infants in order to adjust their position or to soothe them, they tended to put the infants down again, and again to place the infants away from their own bodies so that mother and infant were not in physical contact.

These fourteen mothers were giving a clear signal about expectations for independence; at the same time many of them seemed to be uncomfortable with their infants' dependence.

In the highpoints and lowpoints interviews, ten of these fourteen mothers hinted at or overtly stated that they were uncomfortable with, or at best conflicted about, their infants' dependence on them. For example, one mother described her son as having an "all-access" pass to her body, which she noted as sometimes bringing joy, but then described her son's recent tendency to pull her hair. As she described it, "I just felt like, uh, I'm just literally and figuratively being eaten up by this baby. You know, the breastfeeding, all that stuff. . . . You know, it's like the dark side of the all-access pass." This mother also refers to breastfeeding here, an act of ultimate dependency. Perhaps the link between capacity for reflective functioning, breastfeeding history, and physical contact between mother and infant is through maternal comfort with infant dependency. If a mother can take her infant's point of view through understanding the thoughts, feelings and other mentalizations that underlie his or her behavior and recognizes the mutuality of influence of her and her infant's feelings, if a mother is 'in tune' with her infant in this way, she may be more able to accept the infant's dependence on her. As a corollary to this, she may be more dedicated to breastfeeding when she is more accepting of the infant's dependence.

Two mothers received scores of two on the amount of physical interaction that went on during the video-taped free play episode, indicating that they were rarely in physical contact with their infants. These two mothers also had among the lowest scores on summary reflective functioning, overall reflective functioning, and quality of interaction;

overall, they leave an impression of relatively disturbed mother-infant relationships among a cohort of very well functioning mother-infant pairs. For this reason, they merit a closer look. Each of these two mothers clearly demonstrated an underlying goal of independence for their infants. Both were conflicted about their maternal role, as became apparent during the highpoints and lowpoints interviews. A case study of one of these mothers follows.

This mother spoke of wanting her daughter to be her 'little buddy' and distinguished that at length from her own relationship with her mother in which she felt smothered by her mother's dependence. Her goal was to have a relationship with her daughter that was like that of very good friends, hardly a realistic goal for the maternal-child relationship, and one which negates the child's dependency on the mother which is a given in the relationship. She also spoke of how drained she felt on the one day of the week that she spent alone with her daughter. Describing what she felt by the time her husband came home, she said, "I'm not responding to her because I'm tired, or she just cried for too long or she was cranky for too long, I'm not responding, I can be aware of that now, but at the beginning I don't know what that is, and I was just so, like this [demonstrates flat affect face] because I don't want to talk to her, I don't want to look at her, I just want to—it's not as if I just want to give him the baby, but I was just in a bad mood." These are the words of a mother who cannot take her infant's perspective; it appears that her lack of capacity for reflective functioning interferes with her ability to continue to respond to the infant in the face of aversive infant behavior. Perhaps because she does not so readily understand the infant's motivation for her behavior, the mother has limited tolerance for that behavior.

It is clear that maternal reflective functioning and physical contact with the infant during a free play interaction are related, and may be indicative of a deeper maternal attitude about infant dependence. Although the correlational nature of the relationship between reflective functioning and touch does not allow us to determine a direction of influence, it should be noted that maternal reflective functioning rated prenatally has been shown to be related to infant attachment at one year, a finding in which the direction of influence is clear (Fonagy & Target, 1997). Using this evidence that maternal capacity for reflective functioning is established before the birth of the child, it can be inferred that mothers in the current sample who were high in reflective functioning to begin with subsequently touched their infants more. However, this has not been tested and it remains possible that mothers who are high in physical contact with their infants are then able to talk with more understanding about their own and their infants' mental states.

Regardless of direction of influence, the fact that certain infants receive more physical contact with their mothers is significant. It has been shown that infants benefit from touch; for instance, low birth weight infants who received more positive touching showed fewer internalizing and behavioral problems at age two years (Weiss, Wilson, Seed, & Paul, 2001). Additionally, maternal contact measured in home observations at one, three, and nine months has been found to be positively related to infant sociability at nine months (Fish & Crockenberg, 1981). It would seem that the increased physical interaction of infants with their high reflective functioning mothers would be beneficial.

At least one previous study has found a specific link between maternal insightfulness about infant experience and maternal sensitivity to infant cues that was not replicated in the current study (Koren-Karie, Oppenheim, Dolev, Sher, & Etzion-Carasso, 2002).

Investigators asked mothers of 12-month-old infants to watch three videotaped segments of their own interaction with their infants and then talk about their infants' and their own thoughts and feelings during those interactions. Mothers' heightened ability to talk about the mentalizing underlying their infants' behavior was related to both laboratory-observed maternal sensitivity to infant cues and to infant attachment (Koren-Karie, et al., 2002). It is possible that the impact of maternal reflective functioning on maternal interactive behavior is more specific than that which is measured by the Parent/Caregiver Interaction Scale used here; the Koren-Karie, et al. (2002) study mentioned above found a relationship between a type of maternal insightfulness and sensitive maternal behavior, but this behavior was observed in a laboratory where nuanced infant cues and maternal response can be detected more easily.

In puzzling out why maternal reflective functioning and quality of interaction were not found to be more strongly linked in the current study, it is important to consider those mothers who scored high in one area and low in the other. The two mothers discussed above who showed relatively disturbed relationships with their infants were detected by very low scores in both reflective functioning and quality of interaction, indicating that the two scores can work in conjunction.

Two mothers who notably violated this principle of the two scores working in conjunction are the two who received the lowest scores for quality of interaction, yet received two of the highest scores for reflective functioning. Of note is the fact that one mother was herself a therapist and the other referred to having spent years in therapy; these are women who could talk eloquently about interior experience and the connection between their own feelings and their infants' feelings, yet who missed many of their

infants' cues. Interestingly, both focused mostly on their own interior experiences in their interviews, rather than on their connection with their infants which was more typical in pairs with higher quality interactions. Both mothers were probably narcissistic mothers who would have been better detected as such in the longer Parent Development Interview and subsequently would have received lower scores (A. Slade, personal communication, May, 2003).

One of these high reflective functioning, low quality of interaction mothers described as a highpoint how her son was beginning to put his arms around her neck for a sort of hug, which she interpreted as him showing his appreciation of her. When asked how she felt in that moment, she responded, "I was thinking about it, I sat down at home, I was like, 'Is that my ego? Do I, like, want to feel like totally loved and whatever and I'm just looking for that, or is this just like one of those joyous mom moments?' And I think it's just one of those mom moments. Because while I like the fact that he's independent, and he is a little bit independent, I kind of crave him just needing me, [little laugh] for a day or two, and then going back to being independent!" While she describes a high level of introspection here, it is about her own experience and not her infant's. When asked to play with her infant for the video-taped interaction, this mother set her son in a playpen by himself, and for several minutes left him alone to play, explaining that she like to encourage his independence. She did not return to him until his fussing had turned into crying. These actions contributed to her low quality of interaction score, since she missed many infant cues, and reduced the amount of mutual play and verbalizations, among other things, during this episode.

A typical highpoint for mothers high in both reflective functioning and quality of interaction focussed on connectedness of mother and infant. For instance, one mother described a moment when her son made her laugh by discovering that he could push his highchair away from the dining table with his foot. As she described it, "he smiled because he'd discovered how to do something, and I laughed also, and then I pushed him back, and he did it again, and we did it a few times and I realized we were sharing a joke! . . . That was just a very funny moment, very fun. It felt like we were connecting." This mother used rich language to describe her son's experience of this episode as well. Rather than focusing on her own interior experience, her focus is on the connection between herself and her son. While for this mother that connection is the highpoint, for the previous mother it was feeling appreciated that was the highpoint. This subtle difference was played out in their interactions, which were markedly different in score.

So while reflective functioning differentiated some mother-infant pairs who had low quality interaction, it was better able to do so when the *content* of the interviews, and not just the capacity for reflective functioning, was taken into consideration. Though not hypothesized in the current research, the *qualitative* analysis of the Highpoints and Lowpoints Interviews is the next reasonable step for future research. Additionally, future research ought to include the entire Parent Development Interview (PDI), rather than the curtailed Highpoints and Lowpoints Interview used here; the curtailed interview does not seem to have differentiated among high and low reflective mothers as effectively as the PDI.

Conclusion

The purpose of this research was to explore the concept of goodness of fit in mother-infant pairs. To this end, the relationship of maternal ratings of infant and maternal temperament to maternal sensitivity was investigated. Following a focus in the literature on difficult temperament, infant temperament dimensions of approach/withdrawal, adaptability, mood, and intensity were included in the analyses. Maternal temperament was similarly conceptualized and parallel dimensions of temperament were sought to enable potential comparisons of mothers and infants; the maternal temperament dimension of approach/withdrawal, flexibility/rigidity, and mood were included in the analyses. Both maternal and infant temperaments were rated by maternal report using established questionnaires.

Maternal sensitivity was conceptualized in two ways: as quality of interaction and as maternal reflective function. Quality of interaction consists of the mother's ability to respond to her infant's cues and was measured here in a videotaped free-play interaction between mother and infant. Reflective functioning consists of the mother's ability to read and interpret her infant's cues and was measured here in an interview with the mother about highpoints and lowpoints in interaction with the infant. The relationship of these two components of maternal sensitivity to one another was also considered.

Overall, a trend was found such that mothers of more difficult temperament and infants of easier temperament had higher quality interactions. This was true for direct relationships between infant temperament and quality of interaction, for direct relationships between maternal temperament and quality of interaction, and for the relationships among maternal and infant temperaments and quality of interaction.

Hypothesized relationships among easier mother and infant temperaments and higher quality interactions and among more difficult mother and infant temperaments and lower quality interactions were not supported.

Three possible covariates also emerged; they were maternal age, maternal education, and breastfeeding history. The older and more educated a mother, in this relatively old and extremely well-educated sample, the lower the quality of her interaction. The higher a mother's capacity for reflective functioning, the more likely she was to be more dedicated to breastfeeding.

Regardless of age and education, mothers who reported their infants to be more difficult (i.e., more withdrawing, less adaptable, more intense, and more negative in general mood) talked less to them, used fewer expressions of positive regard with them, expressed this positive regard less contingently to infant behavior, were less involved in mutual play with their infants, and interacted less overall with their infants. These findings reiterate similar findings reported in the literature and were not surprising. They indicate that these infants perceived as more difficult by their mothers are at risk for less-than-optimal development, particularly for insecure attachment.

Again, regardless of age and education, mothers who reported themselves to be of more difficult temperament (i.e., more withdrawing, more rigid, and of less positive general mood) responded more to their infants, better synchronized these responses with infant behavior, and provided more optimal control over the interaction, providing structure but allowing for the infants' input. These surprising findings were in contrast to earlier findings reported in the literature linking easier maternal characteristics with higher quality interaction. These pairs seem to be developing toward secure attachments.

When maternal and infant temperament were considered together and maternal education was controlled for, mothers who reported themselves to be more rigid and who had infants whom they reported to be more adaptable, had higher quality interactions. This leads to a conclusion that adaptability in mothers and infants, albeit in opposite valences, plays a key role in determining goodness of fit. Similarly, mothers who reported themselves to be much more difficult than their infants had higher quality interactions, when controlling for maternal age.

Several explanations for the relationship of maternal difficult temperament to higher quality interaction were offered, including that maternal difficult temperament is an asset rather than a detriment in interacting with an infant, and that the mothers who reported themselves to be more difficult were simply more realistic and less defensive and that this relative psychological stability was related to their higher quality interactions. The hypothesis that these mothers identified as difficult by the temperament questionnaire engaged in more self-monitoring, with less of a social agenda, more of tendency to social monitor, and less of a tendency to dominate in social interaction, and benefited from their easier infants' clearer signals and cues was supported by analyses of specific types of observed interactive behavior. Finally, the possibility was considered that mothers who rated themselves as more difficult than their infants were mothers who were much more likely to hold themselves accountable for any problems in interaction, which in turn may have led them to monitor the interactions more closely. This increased monitoring may have led to more responsiveness, which in turn may have led to higher quality interactions. However, this is speculative; much more research would be required to confirm this hypothesis. The reframing of temperament from a within-child construct to

an interpersonal construct by using maternal report of her own and her infant's temperament, as is the case in the current study, is a promising direction for future research.

Maternal reflective functioning was found to be related to infant mood. It was posited that mothers who could take their infants' point of view may have found infant aversive behavior less bothersome, and therefore reported less of it. More reflective mothers were also found to engage in more physical contact with their infants. This was related to the finding that mothers with a higher capacity for reflective functioning tended to be more dedicated to breastfeeding. It was proposed that these two relationships were linked by a maternal discomfort with, or conflict over, infant dependence that grew from their lesser ability to take their infants' point of view. This interpretation found support in case studies.

Finally, while reflective functioning and quality of interaction did distinguish the two most disturbed mother infant pairs in the sample, the scores sometimes diverged, with a mother scoring high in reflective functioning yet low in quality of interaction. The two most striking cases of this involved women who were very conversant about interior experience (probably due to experience with therapy and their own underlying narcissism), accounting for their high reflective functioning scores. Yet content analysis of one mother's highest score indicated that she was focusing not on her connection with her son, but rather on the complexities of her own interior experience. It was proposed that a qualitative analysis of the Highpoints and Lowpoints Interviews might reveal more of a relationship between reflective functioning and quality of interaction, rather than the quantitative analysis done here. Additionally, it was proposed that the administration of

the entire Parent Development Interview, rather than the curtailed Highpoints/Lowpoints Interview used here, would be beneficial in future research.

While no support was found for the hypothesized relationships between easy mother and infant temperaments and increased maternal sensitivity, and between difficult mother and infant temperaments and decreased maternal sensitivity, overall, strong support was found for the concept of goodness of fit. A good fit was found, albeit between difficult maternal temperament and easy infant temperament. Nevertheless, this fit indicates that both maternal temperament and maternal perception of infant temperament play roles in determining quality of interaction; this role for both maternal and infant temperaments is the hallmark of the concept of goodness of fit.

There was little indication that temperament, either maternal or infant, was closely related to maternal reflective functioning. The exception is maternal perception of infant mood, which seems to be related to maternal capacity for reflective functioning, probably because a mother's ability to understand her infant's behavior makes her perceive that behavior, and therefore report it, as less aversive. While maternal reflective functioning did not play a large role in the findings reported here, it would be important to continue to consider its relationship to quality of interaction for it seems likely that the curtailed interview used here was not an adequate assessment of maternal reflective functioning. It is also possible that a weakness of the current study arises from the administration of the interview before the videotaping of the interaction, a time at which mothers were more defended; this should be addressed in future research.

Prior evidence points to the fact that mother and infant influence one another, contributing to a style of interaction that is continually evolving. Previous research has

not considered the relative contributions of both maternal temperament and maternal perception of infant temperament to this process. The findings of the current research indicate that it is an area of inquiry rife with possibilities for exploration.

Directions for Future Research

As noted earlier, the study of the interaction of maternal personality and parenting is a neglected area of research; so, too, is the exploration of the concept of goodness of fit in infancy. This area could benefit from powerful research designs, such as longitudinal research like that of Isabella and Belsky (1991) and intervention research like that of van den Boom (1994; 1995). One possibility is to extend the current study to include maternal perception of infant temperament during toddlerhood. The second six months of the second year are a notoriously difficult time for parents, as their children become defiant and present many more problems in interaction than they did at four to seven months. It would be interesting to again videotape a free-play interaction at this time and to re-interview mothers using the entire Parent Development Interview.

Additionally, another measure of maternal temperament may be warranted. Rather than using a temperament measure standardized on adolescents and young adults, using a well-standardized personality measure may help to determine why the current findings suggest that maternal difficult temperament is positively related to higher quality interaction. Since this study has revealed that the match between mother and infant temperaments does not appear to be the determining factor in goodness of fit, there is no longer a need to have comparable measures of temperament for mother and infant.

The findings of the current study are obviously limited by the sample size and homogeneity. Even with his homogeneity, the demographic factors of maternal age and

education had an influence. The relative homogeneity of the sample served a purpose, however; it allowed findings in differences in maternal sensitivity to be attributed to the relationship among the temperaments rather than to demographic differences among the mothers.

Future research should extend this study to other populations. Studying a sample of abusive and neglectful mothers would likely provide additional insight into the relationships of maternal and infant temperament to maternal sensitivity, particularly with the addition of measures of parental locus of control. An intervention study with this population, comparing interventions aimed at increasing either maternal reflective functioning, quality of interaction, or both, would greatly increase our knowledge about the relationships of these variables. Finally, it is important to extend the research to include mother-infant pairs of other socio-economic statuses and other cultures to provide a fuller, more comprehensive understandings of the relationship of maternal and infant personality to maternal sensitivity. If such research were to be conducted, the findings of current research indicate that 1) the entire Parent Development Interview be used; 2) a measure of maternal personality should be added to elucidate the meaning of maternal temperament; and 3) the sample size be increased so that comparison groups can be formed and analyzed according to combinations of easy and difficult maternal and infant temperaments. With further research, the meaning of the good fit found here, between more difficult maternal temperament and maternal perception of easier infant temperament, can be elucidated.

Appendix A: Background Interview

Interview

1) Family composition: number of people living in house: _____
   Married Co-habitating Single Divorced Widowed
   Age of mom_____ Age of dad _____
   Baby's DOB_____
   Does your husband have any other children? _____

2) Baby info: previous pregnancies? _____
   Prenatal care? Yes No
   Pregnancy complications? _____
   Due Date _____
   Weight at birth _____
   Where delivered? Hospital Hospital birthing center Birthing Center Home
   Drugs used in delivery? _____
   Delivery complications? _____
   Newborn complications? _____
   Regular checkups and immunizations? _____
   Any diagnosed postpartum depression? _____

3) Education:

Self:	**Spouse:**
Some high school	some high school
High school diploma	high school diploma
Vocational/technical school	vocational/technical school
Some college	some college
College degree	college degree
Some graduate school	some graduate school
Graduate degree	graduate degree

4) Work history:

Self:	**Spouse:**
Do not work outside the home	Does not work outside the home
Less than 20 hours/week	Less than 20 hours/week
Full-time	Full-time
Occupation: _____	Occupation: _____
Title: _____	Title: _____

   Maternity leave arrangements: _____
   _____
   Paternity leave arrangements: _____
   Changes in work arrangements postpartum: _____
   _____

How do you feel about your work situation right now? Is this the right situation for you? Are you comfortable with it? Do you feel your work is rewarding and/or fulfilling in some ways?

5) How would you describe your ethnicity? _____
Your husband's? _____
Your child's? _____

6) Have you participated in any parenting support groups? When, where, how long?

7) Did you have experience with infants before becoming a mother?

8) Is baby: breast fed bottle fed combo eating solid food
Breast feeding history?

9) When you have a question or concern about your baby, where do you seek your information?
Pediatrician mother sister other relative baby's father friends books

other professionals support group therapist

10) How would you characterize these first few months of motherhood? How has it compared with your expectations? What has surprised you, either in a negative or a positive way?

Appendix B: Parent/Caregiver Involvement Scale Coding

The PCIS was used to code the mother-infant interactions. The entire interaction was viewed by the rater, with a checklist to note behaviors important for determining scores on the scale. Each of 11 behaviors during the interactions was rated for Amount, Quality, and Appropriateness. For this research, Quality was the scale that was used for analyses because it best captures maternal sensitive responsiveness, with two exceptions. The Appropriateness scores for Verbal Involvement and Responsiveness were used in place of the Quality scores because they come closer to the working definition of maternal behavioral sensitivity used in this research. Each behavior was rated from 1 to 5, with the odd points anchored as described below. The following is taken from Farran, Kasari, Comfort, and Jay (1986).

1. PHYSICAL INVOLVEMENT

A. Amount of bodily contact *(includes support, touching, holding)*

1	2	3	4	5
very little to none; **A and C hardly ever touch each other *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		moderate; **A** and **C** are in physical contact mostly in the service of other activities or only passive support		very much; constant, must include active touching, not just passive support

B. Quality of caregiver handling of child *(includes changing child's posture, guiding movements, carrying)*

1	2	3	4	5
never sensitive, well-executed handling; handling almost always rough, abrupt, ineffective		sometimes sensitive handling; about half-the-time *(If only passive support occurred, do not rate above a 3.)*		almost always sensitive, well-executed handling; never rough, abrupt

_____*not observed*

C. Appropriateness of caregiver positioning of child: Placement of **C** in a particular posture for the purpose of play or interaction *(e.g., sitting, standing, lying)*; placement of **A** and toys to allow easy access by **C**

1	2	3	4	5
always positioned without adequate and easy access to toys and/or adult; impeding **C**'s best approach to task)		sometimes positioned with adequate access; about half-the-time		almost always positioned adequately for **C**'s best approach

_____*not observed*

2. VERBAL INVOLVEMENT

A. Amount of verbal involvement *(includes initiating and/or responding to C's verbal or nonverbal behavior)*

1	2	3	4	5
none; **A** seldom talks to **C** *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		moderate; **A** occasionally talks to **C**; about half-the- time		very much; **A** talks to **C** throughout session/visit with practically no pauses for **C** to talk

B. Quality of verbal interaction *(adjustment for comprehension))*

1	2	3	4	5
A never adjusts speech to **C**'s level--either too high or too low		moderate adjustment for comprehension; sometimes language directed to child too "babyish" or too complicated		**A** almost always assures **C**'s comprehension of talk directed to **C**; **A** alters tone of voice to gain **C**'s attention

_____*not observed*

C. Appropriateness of verbal interaction *(How much does caregiver provide a verbal link between the child and the world?)*

1	2	3	4	5
A hardly ever comments on **C**'s activities or on **A**'s own activities		**A** occasionally directs talk to **C** about **C**'s activities, relates **A**'s activities to **C**		**A**'s talk almost always relates to **C**'s activities and explains **A**'s own activities relative to **C**. Must be both talking about **C**'s activities and **A**'s activities to receive 5

_____*not observed*

3. RESPONSIVENESS OF CAREGIVER TO CHILD

A. Amount of responsiveness to **C** *(to his initiations, verbalizations, demands, distress)*

1	2	3	4	5
never responds *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		**A** occasionally responds; responds about half-the- time		**A** almost always responds

B. Quality of caregiver responsiveness: Intensity

1	2	3	4	5
responds abruptly, forcefully, very intensely, harshly		neutral; response not intense at all		**A** responds in a gentle, sensitive, positive manner. **A** may respond enthusiastically, with delight. Spontaneity is also observed

_____*not observed*

C. Appropriateness of caregiver responsiveness: Timing

1	2	3	4	5
seldom good synchrony of response to **C**'s activities; **A** overwhelms **C** with quickness of response, is too slow in responding		moderate synchrony of response to **C**'s needs. About half-the-time **A**'s response appropriate and well-timed to **C**'s needs		response to **C** almost always appropriate to **C**'s needs. Good synchrony of response--neither too quick nor too slow

_____*not observed*

4. PLAY INTERACTION

A. Amount of play interaction: Attention/interaction of both caregiver and child to toy/activity *(May include teaching done in a play format but excludes routine child care (e.g., diapering, feeding)*

1	2	3	4	5
very little to none *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		moderate; about half-the- time		almost always

B. Quality of play between caregiver and child *(How much warmth, interest, and enthusiasm does the adult show to the child during play interactions?)*

1	2	3	4	5
A shows no warmth, interest or enthusiasm during play; **A** may seem impatient, neutral or routinized in play		**A** shows warmth, interest or enthusiasm some of the time; at other times seems routinized or detached		**A** responds in a gentle, sensitive, positive manner. **A** may respond enthusiastically, with delight. Spontaneity is also observed

_____*not observed*

C. Appropriateness of play interaction *(Adaption of toys to child's developmental level and interest*

1	2	3	4	5
A never adapts to **C**'s level of ability and interest; **A** persistently uses toys or activities conventionally although inappropriate for **C**		**A** sometimes adapts toys/activities to **C**'s level of ability and interest; about half-the-time		**A** adapts toys/activities to **C**'s level of interest; conventional use of toys and activities fits developmental needs and interests of child.

_____*not observed*

5. TEACHING BEHAVIOR

A. Amount of teaching behavior: For the purpose of teaching a particular skill *(Focus on the Amount and Time* **A** *spends teaching* **C**, *not frequency alone)*

1	2	3	4	5
very little to none *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		moderate; **A** occasionally teaches. **A** may introduce teaching activities but spends little time on each		almost always; **A** uses most of interactive time teaching **C**

B. Quality of teaching behavior

1	2	3	4	5
A subjects **C** to vigorous teaching, almost all is routinized. Non-flexible demands for learning		some teaching is spontaneous, off-the-cuff, creative; some is routinized, drill-oriented, non-flexible		teaching is almost always spontaneous, originating from and addressed to **C**'s activities. **A** creatively incorporates teaching into other activities

_____*not observed*

C. Appropriateness of teaching behavior *(related to developmental capabilities and interests)*

1	2	3	4	5
teaching tasks do not match **C**'s learning needs. **A** unmindful of **C**'s developmental capabilities		**A** sometimes teaches tasks that are appropriate to **C**'s developmental capabilities; about half the teaching		**A** encourages **C** to appropriate level of his/her developmental capabilities. **A** takes into account **C**'s capabilities in choosing what to teach and how

_____*not observed*

6. CONTROL OVER CHILD'S ACTIVITIES

A. Amount of control over child's activities exerted by caregiver

1	2	3	4	5
A never organizes **C**'s activities "Laissez-faire" - **C** on his/her own *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		**A** sometimes organizes **C**'s activities; about half- the-time		**A** almost always organizes **C**'s activities; **A** almost always tells or shows **C** what activities to stop and start

B. Quality of control: Intensity/flexibility

1	2	3	4	5
A insistent upon structure of child's activities; rigid and very firm about what **C** is to do and when		**A** sometimes insistent, demanding in organizing activities, but also somewhat flexible and will relent when **C** is not interested		**A** very flexible in organizing activities; suggests, but not overly insistent; adapts demands according to reactions of **C**

_____*not observed*

C. Appropriateness of control (Fit with child's developmental level)

1	2	3	4	5
A does much more controlling than is warranted for **C**'s developmental level, OR **A** should do a great deal more controlling because of developmental level of **C**		**A** does somewhat more controlling of **C**'s activities than is warranted, occasionally over-controls OR **A** should do somewhat more structuring for the developmental level of **C**		**A** almost always structures **C**'s activities appropriately for **C**'s developmental level. **A** anticipates needs and acts ahead of time. Expectations for amount of structure needed are appropriate to **C**'s skills

_____*not observed*

7. DIRECTIVES: NUMBER OF DEMANDS/COMMANDS MADE OF CHILD VERBALLY OR PHYSICALLY

A. Amount of directives issued by caregiver: Commands for specific behavior				
1	2	3	4	5
A never directs **C**'s specific behaviors *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		**A** issues a moderate number of directives to **C**. No more than half **A**'s verbal behavior		**A** constantly directing **C**'s behavior. Much of **A**'s verbal behavior consists of commands

B. Quality of directives: Intensity				
1	2	3	4	5
very rough; **A**'s directing statements are almost always very forceful and compelling		moderate; **A**'s directives are neutral or of mixed intensities, some forceful and some low-key		very low; **A**'s directive are almost always low-key and gentle, often phrased in the form of suggestions
		_____*not observed*		

C. Appropriateness of directives: Reasonableness of demands/commands				
1	2	3	4	5
A's demands are almost never reasonable for **C**'s abilities and interest level		**A**'s demands occasionally are reasonable, about half-the-time		**A**'s demands are almost always reasonable and appropriate to **C**'s abilities and interest level
		_____*not observed*		

8. RELATIONSHIP AMONG ACTIVITIES IN WHICH CAREGIVER WAS INVOLVED WITH CHILDREN

A. Amount of activities in which caregiver was involved				
1	2	3	4	5
almost no activities observed in which **A** was involved or which **A** helped initiate. Almost none of **A**'s time in activities with children. *(If Amount =1, Rate Not Observed on Quality and Appropriateness)*		equal balance between activities in which **A** was and was not involved with **C**. About half of **A**'s time spent in activities with children		most activities involved **A**; a large number of activities occurred whether with one toy or many. Almost all **A**'s time spent in activities with children

B. Quality of relationship among activities: Sequencing and synchrony				
1	2	3	4	5
A's sequencing of activities and tasks within activities lacks smoothness and fluidity; activities seem to begin and end rather than flow. **A** and **C** seldom ready to end activity at same time. **A** often leaves activities or shifts attention abruptly.		moderate fluidity and smoothness of sequencing between activities and tasks: about half-the-activities. **A** and **C** sometimes are synchronous on beginning and ending activities. **A** sometimes leaves activities or shifts attention abruptly. _____*not observed*		**A** almost always sequences activities and tasks so there is smooth continuity among related activities. **A** elaborates on **C**'s activities in natural order. **A** almost never leaves activities or shifts attention abruptly

C. Appropriateness of relationship among activities

1	2	3	4	5
A never sequences activities from simple to complex, or introduces change to maintain **C**'s interest; activities seem unrelated and confusing		sometimes **A** sequences activities, for example, from simple to complex, or introduces change; about half-the-activities		**A** almost always sequences activities appropriately, for example, from simple to complex, or introduces change to maintain **C**'s interest

_____*not observed*

9. POSITIVE STATEMENTS

A. Amount of expressed positive verbal statements, and non-verbal signs of positive regard *(Praise, hugs, smiles)*

1	2	3	4	5
very little to none; **A** almost never expresses positive emotion *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		moderate; **A** expresses positive emotion in moderate amounts *(about 25% of A's verbal behavior and initiations)*		very much: **A** expresses positive emotion very frequently *(more than 50% of A's verbal behavior and non-verbal initiations)*

B. Quality of expressed positive statements: Intensity

1	2	3	4	5
withdrawn, detached, positive statements made with negative voice, OR extremely overwhelming		moderate intensity; sometimes detached OR intrusive; sometimes high quality		loving, warm; variations in quality dependent on child behaviors; always high quality

_____*not observed*

C. Appropriateness of positive statements: Timing

1	2	3	4	5
A expresses positive emotion at inappropriate times, non-contingently, or inappropriate excess		sometimes inappropriate, sometimes appropriate reactions to **C**'s activities		positive emotion; almost always appropriately timed to behavior

_____*not observed*

10. NEGATIVE STATEMENTS/DISCIPLINE

A. Amount of expressed negative statements and non-verbal *(includes discipline, redirecting statements, criticism, threats, hits, impatience)*

1	2	3	4	5
very little to none; **A** almost never makes negative statements *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		moderate; **A** expresses negative statements no more than 10% of the time		very much; **A** expresses negative emotion very frequently, more than 25% of her/his verbal behaviors and non-verbal initiations

B. Quality of expressed negative statements: Intensity

1	2	3	4	5
intensely negative; **A** uses physical punishment too intensely, severely harsh tone of voice		moderate intensity; **A** occasionally uses harsh tone of voice; sometimes **A** seems impatient, sharp		**A** uses negative emotion with appropriate intensity; may frequently use reasoning to control behaviors. Redirects **C**'s attention

_____*not observed*

C. Appropriateness of negative statements: Timing

1	2	3	4	5
A expresses negative emotion not related to **C**'s activities, or with inappropriate excess		sometimes inappropriate, sometimes appropriate reactions to **C**'s activities. **A** relies on verbal control of **C**'s behavior after the fact, seldom redirects in advance		negative emotion almost always appropriately timed to **C**'s behavior

_____*not observed*

11. GOAL SETTING

A. Amount of caregiver goal setting behavior: Degree to which adult verbally or non-verbally communicates expectations for **C**'s behavior *(Goal setting implies follow through, indicating A expected certain behavior of C)*

1	2	3	4	5
none; **A** never communicates goals for **C** *(If Amount = 1, Rate Not Observed on Quality and Appropriateness)*		moderate; **A** occasionally communicates goals for **C**; half the time spent pursuing specific goals for **C**'s behavior, sometimes follows through		very frequently; **A** almost continually communicates goals for **C**, follows through in demands to get **C** to fulfill goal.

B. Quality of goal setting: Adult's overall flexibility, ability to adjust self, environment, or child so that **C** will meet success at an activity

1	2	3	4	5
A never adjusts demands, environment, toys to aid **C**'s success at attaining goals **A** has communicated		**A** sometimes is flexible, occasionally adjusts environment so **C** can be successful at achieving goals		**A** almost always adjusts to aid **C**'s success

_____*not observed*

C. Appropriateness of goal setting: Reasonableness of adult's expectations for **C**'s behavior

1	2	3	4	5
A never sets attainable, reasonable challenges for **C**; **A** unmindful of **C**'s ability level		sometimes **A**'s challenges are attainable; about half the time		**A**'s challenges are almost always moderate, attainable, and appropriate to **C**'s capabilities

_____*not observed*

12. GENERAL IMPRESSION OF CAREGIVER CHILD INTERACTION

A. Availability of **A** to **C**: Degree to which **C** has access to **A**'s attention and Involvement

1	2	3	4	5
A appears oblivious, preoccupied, inaccessible to **C**		**A** appears accessible to **C** if needed; moderately responsive to **C**; **C** receives equal attention to that given other activities.		**A** appears intensely involved, continually responsive; time seems to revolve around **C** and his/her activity

B. General acceptance and approval manifested by **A**: Extent to which **A** seems to like **C**

1	2	3	4	5
very low approval and acceptance; **A** is definitely rejecting, disapproving of **C** OR **A** is indifferent		moderate approval and acceptance; about half the time		very high, **A** exhibits much approval and acceptance

_____*not observed*

C. General atmosphere of caregiver child interaction

1	2	3	4	5
very much discord and conflict, OR indifference		sometimes the atmosphere is positive; about half the time		very harmonious, agreeable, friendly, peaceful, not one unhappy episode

_____*not observed*

D. Enjoyment

1	2	3	4	5
A never seems to take pleasure in **C**; **A** is either not involved or merely accepting		sometimes **A** seems to enjoy, take delight in and find happiness in being with **C**; about half the time OR **A** is neutral		**A** takes delight in **C**; **A**'s enjoyment is obvious and continual

E. **A**'s provision of a learning environment: That is, the provision of time, space, attention and adaption that supports **C**'s optimum concentration on single task

1	2	3	4	5
poor or non-existent toys changed interruptively or not at all, or else learning space crowded with toys. **A** unmindful of adjusting task to level of moderate challenge		moderate to good; times between **A** and **C** highlighted by occasional moments of synchronized absorption in learning; about half the time		excellent; **A** engaged in support of successful learning environment for entire session/visit

Appendix C: Highpoints/Lowpoints Interview

Directions: Every mother of a young child experiences both joys and frustrations in her interactions with her child. I am interested in what kinds of events mark the highpoints (joys) and lowpoints (frustrations) for first-time moms.

A highpoint can be anything that makes you feel good in interacting with your baby. A lowpoint can be anything that makes you feel bad in interacting with your baby. The things that make highpoints and lowpoints will vary from mother to mother. Don't worry about what other people would say, concentrate on what provides enjoyment and frustration for **you**.

Think about the past week. Could you tell me about a highpoint with your baby in the last week?

Probes: Could you describe the situation for me?
What was your reaction to the baby?
What was your baby's reaction to you?
What were you feeling?
What do you think the baby was feeling?

Tell me about another highpoint.

Same probes.

Could you tell me about one more highpoint?

Same probes.

Okay, now could you tell me about a lowpoint with your baby in the last week?

Same probes.

Tell me about another lowpoint.

Same probes

Could you tell me about one more lowpoint?

Same probes.

Appendix D: Maternal Reflective Functioning Coding

Maternal reflective functioning coding was developed for use on the Adult Attachment Interview by Fonagy, Steele, Steele, and Target (1998). An addendum to the reflective functioning scoring for use with the Parent Development Interview was developed by Slade, Bernbach, Grienenberger, Levy, and Locker (2000). The following is a brief summary, with some examples, taken from the Slade *et al.* manual for scoring of reflective functioning in parents' talk which will be used in the proposed research. The complete coding manual, which is far more complex, is available from Fonagy. The addendum is available from Slade. Please note that the examples here are based on the Parent Development Interview, not the Highpoints/Lowpoints Interview used in the current study; therefore, not all answers pertain to highpoints or lowpoints in interaction.

Each individual passage (description of a highpoint or lowpoint) will be rated, and then each mother's responses as a whole will be rated. The following describes the rating of individual passages. Both passage ratings and overall ratings range from –1 to 9.

-1 Negative Reflective Function (RF): Distinctively anti-reflective, irrational, or inappropriate.

"He's just a baby so how could we possibly be alike? You've met him. I don't see how you could possibly ask me that question."

1 Absent but not repudiated RF: A response which is passively rather than actively evasive, with no evidence of awareness of mental states, perhaps containing concrete explanations of mental states which serve to avoid references to mental states.

"(What do you like most about your child?) It's nice being around him and to watch him do things."

"I never feel guilty as a parent."

2 Rating: Vague references to RF, in which the mental state can be inferred, but is too limited and inexplicit to merit a rating of 3.

"For her, when the stroller comes out, it means we're going outside."

3 Questionable or low RF: Contains some suggestion of mentalizing efforts on the part of the subject but is devoid of any element that makes reflective functioning explicit.

"When I'm around he wants me and not her and it's a crying situation."

"Well, I must be doing something right because he's a great child."

4 Rating: Passages in which slightly more sophisticated mental state language is used than would receive a 3 rating, but not definite or elaborated enough to receive a 5 rating.

"I'm a calm parent, which is I think where she gets her disposition, and so is my husband, we're very calm, and I think that's why she is."

"She's very expressive in nonverbal ways. You can tell the difference between happy sounds and angry sounds. There are a lot of things she does physically to communicate what she needs. She can point to things or if she's tired, she'll rub her eyes."

5 Definite or ordinary RF: Passage must contain some element which makes reflection explicit, such as reference to the nature or properties of mental states—how mental states relate to behavior or mental states in relation to the interviewer—and must not be a cliché.

"I think she gets bored sometimes when she is confined, and when she decides she wants out, she gets really unhappy."

"He was very upset, and I could tell he didn't understand why I wouldn't take him back from the babysitter."

6 Rating: Passages which contain reflective statements that are more explicit and elaborated than responses rated 5, but do not meet the criteria to be rated 7.

"She's going through a phase where she wants me to hold her, but I don't want to hold her when she is whining. So we weren't clicking because she wasn't responding to me and wasn't calming down, and she was probably getting more frustrated because she was wondering why mommy wasn't holding her."

7 Marked RF: A response which contains some feature which makes reflection explicit: explicit reference to the nature or properties of mental states, how mental states relate to behavior, or mental states in relation to the interviewer. In addition, a response must meet at least one of six specific criteria (see manual for details).

"Sometimes when she does something that she knows I'll think is funny, she'll look at me and really smile or laugh. Then I laugh with her and I know that she knows she's done something to get us playing and that makes me feel good."

8 Rating: Passages which meet the criteria for a 7 rating and one of the 2 criteria for a 9 rating.

"Well, there are sometimes when, ...no, actually a better example would be last weekend. She began teething and nothing seemed to make her feel comfortable or happy. She didn't want to be held, and she didn't want to play, and she really didn't want to do much of anything. I couldn't figure out what to do to make her feel better. It really affected all of us, and we were exhausted and depressed by the end of the day."

9 Full or exceptional RF: A passage which shows the features of a 7 to an unusually high degree (in the top 10%) *or* the response must be given for a particularly charged and emotionally difficult subject in which maintaining even ordinary levels of reflective functioning could be considered exceptional. It must also have a strikingly personal character, enabling the rater to feel confident that it is experienced as personally significant and meaningful. Responses that are given a 9 frequently demonstrate full awareness of important aspects of all protagonists within an interaction. The protagonists are placed in relation to one another in terms of their feelings and beliefs and these are sufficiently complex and elaborate to convince the rater of their accuracy.

"Sometimes she gets frustrated and angry in ways that I'm not sure I understand. She points to one thing and I hand it to her, but it turns out that's not really what she wanted. It feels very confusing to me when I'm not sure how she's feeling, especially when she's upset. Sometimes she'll want to do something and I won't let her because it's dangerous so she'll get angry. I may try to pick her up and she obviously didn't want to be picked up because she's in the middle of being angry and I interrupted her. In those moments it's me who has the need to pick her up and make her feel better, so I'll put her back down."

Appendix E: Consent Form

My name is Lindsay Dunckel and I am a student in the Child Development Department at the Graduate School and University Center at the City University of New York (CUNY), and the principal investigator of this project entitled *Defining Goodness-of-Fit in Infancy.* The purpose of this study is to explore how mothers' and babies' temperaments influence how they interact. The study is expected to help us understand more about how a baby's temperament influences his or her development. I would like your permission to interview you about your experiences, to tape-record the interview, to videotape you playing with your baby for 30 minutes, and would like you to fill out 2 questionnaires.

The interview should take from 20 to 40 minutes, and the questionnaires should take less than one hour. Participation in this study is voluntary. If you choose to participate in this study, you will be identified only by code number on the questionnaires, the interview form, the cassette tape and the videotape. The audiotape of the interview will be heard only by me. The videotape of you and your baby will be seen by me and possibly one other person who will not know who you are. All information will be kept strictly confidential and will be stored in a locked filing cabinet to which only I will have access. In any report of the research, the participants will not be identified. You may refuse to answer any question or to end the videotaping or interview at any time without penalty. You may also review part or all of the videotape or audiotape made of you and withdraw any part of them without penalty.

In participating in this study, you will be contributing to our understanding of the role that babies' temperament plays in their development, particularly how it impacts the developing mother-infant relationship. You may experience some discomfort in answering questions about this emotional time of life. It should be noted that if I suspect that a mother is physically abusing or neglecting her child, I will have to report this to the authorities.

Please feel free to call me with any questions you may have about this research at (718) 855-4811. You may also contact my advisor, Dr. Herbert Saltzstein, at (212) 817-8717. If you have any questions about your rights as a participant in this study, please call Hilry Fisher, Office of Sponsored Research, Graduate School/CUNY, (212) 817-7523, hfisher@gc.cuny.edu. If you would like to have a copy of the study or a summary of the findings, please provide me with your name and address and I will send you copy when it is completed.

You will receive a copy of this consent form.

I, _____ (name), agree to participate in the research project described above. I have had the opportunity to discuss this project and my questions have been answered to my satisfaction.

_____ _____ I agree to be videotaped:
Signature of participant Date (circle one)
Yes **No**

_____ _____ I agree to be audiotaped:
Signature of investigator Date (circle one)
Yes **No**

Appendix F: Child Consent Form

My name is Lindsay Dunckel and I am a student in the Child Development Department at the Graduate School and University Center at the City University of New York (CUNY), and the principal investigator of this project entitled *Defining Goodness-of-Fit in Infancy.* The purpose of this study is to explore how mothers' and babies' temperaments influence how they interact. The study is expected to help us understand more about how a baby's temperament influences his or her development. I would like your permission to videotape your baby interacting with you for 30 minutes.

Participation in this study is voluntary. If you choose to have your baby participate in this study, he/she will be assigned a code number. This code number will be the only identification on the videotape I will make of your baby. The videotape will be viewed by me and possibly one other person who will not know who your baby is. All information will be kept strictly confidential and will be stored in a locked filing cabinet to which only I will have access. In any report of the research, the participants will not be identified. You may review and withdraw any part or all of the videotape made of your baby, or any of the information gathered about your baby. If your baby cries excessively during any part of this process, this will be respected as a withdrawal from the study. There will be no penalty for you or your baby withdrawing from the study at any time.

Please feel free to call me with any questions you may have about this research at (718) 855-4811. You may also contact my advisor, Dr. Herbert Saltzstein, at (212) 817-8717. If you have any questions about your rights as a participant in this study, please call Hilry Fisher, Office of Sponsored Research, Graduate School/CUNY, (212) 817-7523, hfisher@gc.cuny.edu. If you would like to have a copy of the study or a summary of the findings, please provide me with your name and address and I will send you copy when it is completed.

You will receive a copy of this form.

I, _____ (name), agree to have my child, _____ (child's name) participate in the research project described above. I have had the opportunity to discuss this project and my questions have been answered to my satisfaction.

I agree to have my child videotaped: (circle one) **yes** **no**

_____ _____
Signature of parent Date

_____ _____
Signature of investigator Date

Appendix G: Independent and Background Variables Correlations Table

Backgrnd Variables	Maternal Temperament Variables[a] *r* values (2-tailed sig.)			Infant Temperament Variables[b] *r* values (2-tailed sig.)			
	Approach	Flexibility	Mood	Approach	Adaptabil.	Intensity	Mood
Infant Gender[c]	-.190 (.283)	.076 (.669)	-.228 (.195)	.087 (.625)	.065 (.715)	-.023 (.898)	.055 (.758)
Infant Age	-.041 (.819)	-.204 (.246)	-.145 (.414)	.011 (.951)	.172 (.331)	.304 (.080)	.047 (.792)
Infant Birth Weight	.070 (.695)	-.160 (.365)	.165 (.350)	-.108 (.543)	.134 (.449)	.000 (.999)	-.013 (.942)
Geographic Location[d]	.112 (.528)	.303 (.081)	.161 (.362)	-.348 (.044)	-.056 (.753)	-.350 (.042)	-.067 (.706)
Maternal Age	-.204 (.248)	-.198 (.262)	-.185 (.296)	.114 (.522)	.123 (.488)	-.015 (.933)	.102 (.565)
Paternal Age	-.034 (.852)	-.076 (.673)	-.338 (.055)	.119 (.510)	.137 (.447)	.253 (.156)	.137 (.449)
Maternal Education	-.078 (.662)	.028 (.875)	-.109 (.540)	.138 (.436)	-.100 (.573)	-.128 (.472)	.261 (.136)
Paternal Education	-.328 (.063)	.021 (.907)	-.427 (.013)	-.024 (.893)	.079 (.663)	-.089 (.621)	.129 (.475)
Maternal Employmt.[e]	.218 (.215)	.150 (.398)	.057 (.748)	-.134 (.451)	-.173 (.328)	-.047 (.793)	-.173 (.327)
Maternal Ethnicity[f]	.274 (.117)	.200 (.256)	.133 (.453)	.073 (.681)	.121 (.495)	.251 (.152)	-.050 (.779)
Prior Experience	.176 (.318)	.015 (.932)	.411 (.016)	-.040 (.821)	-.147 (.406)	-.011 (.949)	.131 (.459)
Breast Feeding History[g]	.046 (.795)	.204 (.247)	-.017 (.925)	-.082 (.646)	.166 (.348)	-.122 (.491)	.114 (.522)
Cry Percent[h]	.070 (.696)	-.015 (.935)	.172 (.330)	.025 (.866)	.173 (.328)	.020 (.911)	.290 (.096)

[a] Higher score indicates less difficult temperament
[b] Higher score indicates more difficult temperament
[c] Scored as 0 for male, 1 for female
[d] Scored as 0 for East coast, 1 for West coast
[e] Scored as 0 for not employed, 1 for less than 20 hours/week, 2 for 20 hours/week or more
[f] Scored as 0 for white, non-Hispanic, 1 for other
[g] Scored as 1 for weaned to formula before three months, 2 for partly weaned to formula before three months, 3 for weaned after three months, 4 for breast milk only to date of visit
[h] Percentage of the visit the infant cried

Appendix H: Dependent and Background Variables Correlations Table

Background Variable	Quality of Interaction *r* value (2-tailed sig.)	Reflective Functioning Overall[a] *r* value (2-tailed sig.)	Summary Reflective Functioning[b] *r* value (2-tailed sig.)
Infant Gender[c]	-.114 (.521)	-.099 (.577)	.068 (.703)
Infant Age	.098 (.581)	-.241 (.169)	-.219 (.213)
Infant Birth Weight	.263 (.133)	.231 (.190)	-.003 (.987)
Geographic Location[d]	-.021 (.908)	-.234 (.184)	-.168 (.343)
Maternal Age	-.346 (.045)	.004 (.980)	-.057 (.750)
Paternal Age	-.183 (.307)	-.026 (.888)	-.009 (.960)
Maternal Education	-.360 (.037)	.005 (.978)	.001 (.997)
Maternal Employmt.[e]	.011 (.952)	.024 (.894)	-.061 (.732)
Maternal Ethnicity[f]	.026 (.883)	-.090 (.611)	-.153 (.389)
Prior Experience	.233 (.184)	.231 (.188)	.113 (.524)
Breast Feeding History[g]	.042 (.813)	.312 (.072)	.377 (.028)
Cry Percent[h]	-.281 (.107)	.078 (.660)	-.028 (.874)

[a] RF Overall is a single score assigned to the whole interview by the coder
[b] Summary RF is the sum of each individual's two highest scores
[c] Scored as 0 for male, 1 for female
[d] Scored as 0 for East coast, 1 for West coast
[e] Scored as 0 for not employed, 1 for less than 20 hours/week, 2 for 20 hours/week or more
[f] Scored as 0 for White, non-Hispanic, 1 for other
[g] Scored as 1 for weaned to formula before three months, 2 for partly weaned to formula before three months, 3 for weaned after three months, 4 for breast milk only to date of visit
[h] Percentage of the visit the infant cried

Appendix I: Means and Standard Deviations for Interaction Scores

Score	Minimum	Maximum	Mean	Standard Deviation
Amount Physical	2	5	3.74	0.864
Quality Physical	2	5	3.91	0.933
Appropriateness Physical	2	5	4.32	0.912
Amount Verbal	2	5	3.53	0.662
Quality Verbal	2	5	3.88	1.066
Appropriateness Verbal	2	5	3.94	1.099
Amount Responsiveness	3	5	4.18	0.869
Quality Responsiveness	2	5	3.94	0.851
Approp. Responsiveness	1	5	3.79	1.250
Amount Play	2	5	3.50	0.992
Quality Play	2	5	3.88	0.946
Appropriateness Play	2	5	3.68	1.093
Amount Control	2	5	3.44	0.860
Quality Control	1	5	3.91	1.190
Appropriateness Control	1	5	3.59	1.209
Amount Directives	1	5	2.79	0.946
Quality Directives	3	5	4.36	0.742
Appropriateness Directives	2	5	4.48	0.795
Amount Activities	2	5	4.29	1.060
Quality Activities	1	5	3.15	1.438
Appropriateness Activities	1	5	3.18	1.029
Amount Positive	2	5	3.65	0.774

Score	Minimum	Maximum	Mean	Standard Deviation
Quality Positive	2	5	3.94	0.983
Appropriateness Positive	2	5	4.24	1.017
Amount Goal	1	4	2.21	0.687
Quality Goal	1	5	3.97	1.129
Appropriateness Goal	3	5	4.30	0.877
Availability	2	5	4.26	0.898
Acceptance	3	5	4.29	0.871
Atmosphere	2	5	3.88	0.946
Enjoyment	2	5	3.56	1.021
Learning Environment	1	5	2.94	1.205
Mean Amount Score	2.33	4.11	3.48	0.471
Mean Quality Score	2.33	5.00	3.88	0.779
Mean Appropriateness Score	2.44	5.00	3.93	0.774
Aggregate "Quality of Interaction" score	2.43	4.93	3.84	0.770

Appendix J: Means and Standard Deviations for Reflective Functioning Scores

Score	Minimum	Maximum	Mean	Standard Deviation
Highpoint 1	2	7	4.91	1.083
Highpoint 2	1	7	4.68	1.147
Highpoint 3	2	6	4.40	1.163
Lowpoint 1	2	7	4.50	1.212
Lowpoint 2	2	6	4.35	1.203
Lowpoint 3	2	6	4.31	1.137
RF Overall[a]	3	6	4.97	0.904
Summary RF[b]	9	13	10.79	1.321

[a] RF Overall is the single overall score assigned by the coder
[b] Summary RF is the sum of the two highest individual scores

References

Aber, J. L., Slade, A., Berger, B., Bresgi, I., & Kaplan, M. (1985). *The Parent Development Interview.* Unpublished manuscript.

Ainsworth, M. D., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation.* Hillsdale, NJ: Erlbaum.

Bates, J. E. (1980). The concept of the difficult child. *Merrill-Palmer Quarterly, 26(4),* 295-319.

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

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

Bates, J. E., Olson, S. L., Pettit, G. S., & Bayles, K. (1982). Dimensions of individuality in the mother-infant relationship at 6 months of age. *Child Development, 53,* 446-461.

Beckwith, L., & Parmelee, A. H. (1986). EEG patterns of preterm infants, home environment, and later IQ. *Child Development, 57,* 777-789.

Bell, R. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review, 75,* 81-95.

Belsky, J. (1984). The determinants of parenting: A process model. *Child Development, 55,* 83-96.

Belsky, J., & Barends, N. (2002). Personality and parenting. In Bornstein, M. H. (Ed.), *Handbook of parenting,* (2nd ed., Vol. 3, pp. 415–438). Mahwah, NJ: Lawrence Erlbaum.

Bornstein, M. H., Hahn, C. H., Suwalsky, J. T. D., & Haynes, O. M. (2003). Socioeconomic status, parenting, and child development: The Hollingshead Four-Factor Index of Social Status and the Socioeconomic Index of Occupations. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting, and child development*, pp. 29–82. Mahwah, NJ: Lawrence Erlbaum.

Boukydis, C.F.Z., & Burgess, R. L. (1982). Adult physiological response to infant cries: Effects of temperament of infant, parental status, and gender. *Child Development, 53,* 1291-1298.

Bradley, R. H., Caldwell, B. M., & Rock, S. L. (1988). Home environment and school performance: A ten-year follow-up and examination of three models of environmental action. *Child Development, 59,* 852-867.

Brazelton, T. B., Koslowski, B., & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver,* pp. 49-76. New York: Wiley.

Bronfenbrenner, U. (1979). *The ecology of human development.* Cambridge, MA: Harvard University Press.

Buss, A. H. (1991). The EAS theory of temperament. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement,* pp. 43-60. London: Plenum Press.

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

Campbell, S. B.G. (1979). Mother-infant interaction as a function of maternal ratings of temperament. *Child Psychiatry and Human Development, 10,* 67-76.

Campis, L. K., Lyman, R. D., & Prentice-Dunn, S. (1986). The Parental Locus of Control Scale: Development and validation. *Journal of clinical Child Psychology, 15,* 206-267.

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

Carey, W. B. (1983). Some pitfalls in infant temperament research. *Infant Behavior and Development, 6,* 247-254.

Carey, W. B., & McDevitt, S. C. (1978). Revision of the Infant Temperament Questionnaire. *Pediatrics, 61,* 735-739.

Carnicero, J. A. C., Perez-Lopez, J., Salinas, M. D. C. G., & Martinez-Fuentes, M. T. (2000). A longitudinal study of temperament in infancy: Stability and convergence of measures. *European Journal of Personality, 14,* 21 – 37.

Cattell, R. B., (1978). *The scientific use of factor analysis.* New York: Plenum Press.

Chess, S., & Thomas, A. (1991). Temperament and the concept of goodness of fit. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement,* pp. 15-28. London: Plenum Press.

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

Crockenberg, S. (1986). Are temperamental differences in babies associated with predictable differences in care giving? *New Directions for Child Development, 31,* 53-73.

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

De Wolff, M. S., & van IJzendoorn, M. H. (1997). Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment. *Child Development, 68,* 571-591.

Doelling, J. L., & Johnson, J. H. (1990). Predicting success in foster placement: The contribution of parent-child temperament characteristics. *American Journal of Orthopsychiatry, 60,* 585-593.

Donovan, W. L., & Leavitt, L. A. (1978). Early cognitive development and its relation to maternal physiologic and behavioral responsiveness. *Child Development, 49,* 1251-1254.

Elardo, R., Bradley, R., & Caldwell, B. (1975). The relation of infants' home environment to mental test performance from six to thirty-six months: A longitudinal analysis. *Child Development, 46,* 71-76.

Farran, D. C., Kasari, C., Comfort, M., & Jay, S. (1986). *Parent/Caregiver Involvement Scale.* Unpublished manuscript. Nashville, TN: Vanderbilt University.

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

Fonagy, P. & Target, M. (1997). Attachment and reflective function: Their role in self organization. *Development and Psychopathology, 9,* 679-700.

Fonagy, P., Steele, H., & Steele, M. (1991). Maternal representations of attachment during pregnancy predict the organization of infant-mother attachment at one year of age. *Child Development, 62,* 891-905.

Fonagy, P., Steele, M., Steele, H., Leigh, T. Kennedy, R., Mattoon, G., & Target, M. (1995). Attachment, the reflective self, and borderline states: The predictive specificity of the Adult Attachment Interview and pathological emotional development. In S. Goldberg, R. Muir, & J. Kerr (Eds.), *Attachment theory: Social, developmental, and clinical perspectives* (pp. 233-279). Hillsdale, NJ: Analytic Press.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-functioning manual, Version 5: For application to adult attachment interviews.* Unpublished manuscript.

Galambos, N. L., & Turner, P. K. (1999). Parent and adolescent temperaments and the quality of parent-adolescent relations. *Merrill-Palmer Quarterly, 45,* 493-511.

George, C., Kaplan, N., & Main, M. (1985). *The adult attachment interview.* Unpublished manuscript, University of California, Berkeley, Department of Psychology.

Goldsmith, H. H., & Campos, J. J. (1982). Toward a theory of infant temperament. In R. N. Emde & R. J. Harmon (Eds.), *The development of attachment and affiliative systems,* pp. 161-193. New York: Plenum.

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

Gordon, B. N. (1983). Maternal perception of child temperament and observed mother-child interaction. *Child Psychiatry and Human Development, 13,* 153-167.

Gorman, K. S., Lourie, A. E., & Choudhury, N. (2001). Differential patterns of development: The interaction of birth weight, temperament, and maternal behavior. *Journal of Developmental and Behavioral Pediatrics, 22,* 366-375.

Grienenberger, J., Kelly, K., & Slade, A. (2001, April). *Maternal reflective functioning and the caregiving relationship: The link between mental states and mother-infant affective communication.* Paper presented at the Biennial Meetings of the Society for Research in Child Development, Minneapolis, MN.

Gross, D., Conrad, B., Fogg, L., & Wothke, W. (1994). A longitudinal model of maternal self-efficacy, depression, and difficult temperament during toddlerhood. *Research in Nursing & Health, 17,* 207-215.

Grotevant, H. D., McRoy, R. G., & Jenkins, V. Y. (1988). Emotionally disturbed, adopted adolescents: Early patterns of family adaptation. *Family Process, 27,* 439-457.

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

Isabella, R. A., & Belsky, J. (1991). Interactional synchrony and the origins of infant-mother attachment: A replication study. *Child Development, 62,* 373-384.

Kagan, J. (1982). The construct of the difficult temperament: A reply to Thomas, Chess, and Korn. *Merrill-Palmer Quarterly, 28,* 21-24.

Keogh, B. K. (1986). Temperament and schooling: Meaning of "Goodness of Fit"? *New Directions for Child Development, 31,* 89-108.

Kivijaervi, M., Voeten, M. J. M., Niemelae, P., Raeihae, H., Lertola, K., & Piha, J. (2001). Maternal sensitivity behavior and infant behavior in early interaction. *Infant Mental Health Journal, 22,* 627-640.

Klein, P. S. (1984). Behavior of Israeli mothers towards infants in relation to infants' perceived temperament. *Child Development, 55,* 1212-1218.

Koren-Karie, N., Oppenhein, D., Dolev, S., Sher, E., & Etzion-Carasso, A. (2002). Mother's insightfulness regarding their infants' internal experience: Relations with maternal sensitivity and infant attachment. *Developmental Psychology, 38,* 534–542.

Lamb, M. E., & Gilbride, K. E. (1985). Compatibility in parent-infant relationships: Origins and processes. In W. Ickes (Ed.), *Compatible and incompatible relationships,* pp. 33-60. New York: Springer-Verlag.

Leerkes, E. M., & Crockenberg, S. C. (2002). The development of maternal self-efficacy and its impact on maternal behavior. *Infancy, 3,* 227 – 247.

Lerner, J. V. (1983). The role of temperament in psychosocial adaptation in early adolescents: A test of a "goodness of fit" model. *Journal of Genetic Psychology, 143,* 149-157.

Lerner, J. V., Lerner, R. M., & Zabski, S. (1985). Temperament and elementary school children's actual and rated academic performance: A test of a "goodness-of-fit" model. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 26,* 125-136.

Lerner, R. M., Castellino, D. R., Patterson, T. A., Villarruel, F. A., & McKinney, M. H. (1995). Developmental contextual perspective on parenting. In M. H. Bornstein (Ed.), *Handbook of parenting, Vol. 2: Biology and ecology of parenting*, pp. 285-309. Hillsdale, NJ: Lawrence Erlbaum.

Lerner, R. M., Palermo, M., Spiro, A. III, & Nesselroade, J. R. (1982). Assessing the dimensions of temperamental individuality across the life span: The Dimensions of Temperament Survey (DOTS). *Child Development, 53,* 149-159.

Lipetz, M. E., Cohen, I. H., Dworin, J., & Rogers, L. S. (1970). Need complementarity, marital stability, and marital satisfaction. In K. J. Gergen & D. Marlowe (Eds.), *Personality and social behavior,* pp. 201-212. Reading, MA: Addison-Wesley.

Logan, E., Kern, R., Curlette, W., & Trad, A. (1993). Couples adjustment, life-style similarity, and social interest. *Individual Psychology: Journal of Adlerian theory, research, and practice, 49,* 456-467.

Luster, T., & Okagaki, L. (1993). Multiple influences on parenting: Ecological and life-course perspectives. In T. Luster & L. Okagaki (Eds.), *Parenting: An ecological perspective*, pp. 227-250. Hillsdale, NJ: Lawrence Erlbaum.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4,* 84-99.

Mangelsdorf, S., Gunnar, M., Kestenbaum, R., Lang, S., & Andreas, D. (1990). Infant proneness-to-distress, maternal personality, and mother-infant attachment: Associations and goodness of fit. *Child Development, 61,* 820-831.

Matheny, A. P., Wilson, R. S., & Thoben, A. S. (1987). Home and mother: Relations with infant temperament. *Developmental Psychology, 23,* 323-331.

Mednick, B. R., Hocevar, D., Schulsinger, C., & Baker, R. L. (1996). Personality and demographic characteristics of mothers and their ratings of their 3- to 10-year-old children's temperament. *Merrill-Palmer Quarterly, 42,* 397-417.

Milliones, J. (1978). Relationship between perceived child temperament and maternal behaviors. *Child Development, 49,* 1255-1257.

Munson, L. J., & Odom, S. L. (1996). Review of rating scales that measure parent-infant interaction. *Topics in Early Childhood Special Education, 16,* 1-25.

Park, K. J. (2001). Attachment security of 12 month old Korean infants: Relations with maternal sensitivity and infants' temperament. *Early Child Development & Care, 167,* 27-38.

Pauli-Pott, U., Becker, K., & Mertesacker, T. (2000). Infants with "colic"—mother's perspectives on the crying problem. *Journal of Psychosomatic Research, 48,* 125-132.

Peters-Martin, P. & Wachs, T. D. (1984). A longitudinal study of temperament and its correlates in the first 12 months. *Infant Behavior and Development, 7,* 285-298.

Plomin, R. & DeFries, J. C. (1983). The Colorado Adoption Project. *Child Development, 54,* 276-289.

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

Rothbart, M. K. (1982). The concept of difficult temperament: A critical analysis of Thomas, Chess, and Korn. *Merrill-Palmer Quarterly, 28,* 35-40.

Sameroff, A. J., & Chandler, M., J. (1975). Reproductive risk and the continuum of caretaker casualty. In F. D. Horowitz (Ed.), *Review of child development research (Vol. 4).* Chicago: University of Chicago Press.

Sameroff, A. J., & Fiese, B. H. (1992). Family representations of development. In I. E. Sigel, A. V. McGillicuddy-DeLisi, & J. J. Goodnow (Eds.), *Parental belief systems: The psychological consequences for children (Second edition),* pp. 347-369. Hillsdale, NJ: Erlbaum.

Sanson, A. & Rothbart, M. K. (1995). Child temperament and parenting. In M .H. Bornstein (Ed.), *Handbook of parenting Volume 4: Applied and practical parenting,* pp. 299-321. Mahwah, NJ: Lawrence Erlbaum.

Schneirla, T. C. (1957). The concept of development in comparative psychology. In D. B. Harris (Ed.), *The concept of development.* Minneapolis: University of Minnesota Press.

Seifer, R., & Sameroff, A. J. (1986). The concept, measurement, and interpretation of temperament in young children: A survey of research issues. *Advances in Developmental and Behavioral Pediatrics, 7,* 1-43.

Seifer, R., Schiller, M., Sameroff, A., Resnick, S., & Riordan, K. (1996). Attachment, maternal sensitivity, and infant temperament during the first year of life. *Developmental Psychology, 32,* 12-25.

Sheeber, L. B., & Johnson, J. H. (1992). Child temperament, maternal adjustment, and changes in family life style. *American Journal of Orthopsychiatry, 62(2),* 178-185.

Simeonsson, R. J., Bailey, D. B. Jr., Huntington, G. S., & Comfort, M. (1986). Testing the concept of goodness of fit in early intervention. *Infant Mental Health Journal, 7,* 81-94.

Slade, A. (2002). Keeping the baby in mind: A critical factor in perinatal mental health. *Zero to Three, 22 (6),* 10-16.

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.

Slade, A., Bernbach, E., Grienenberger, J., Levy, D. W., & Locker, A. (2000). *Addendum to Reflective Functioning Scoring Manual: For use with the Parent Development Interview.* Unpublished manuscript.

Slade, A., Grienenberger, J., Bernbach, E., Levy, D., & Locker, A. (2001, April). *Maternal reflective functioning: Considering the transmission gap.* Paper presented at the Biennial Meetings of the Society for Research in Child Development, Minneapolis, MN.

Steele, H., Steele, M., & Fonagy, P. (1996). Associations among attachment classifications of mothers, fathers, and their infants. *Child Development, 67,* 541-555.

Stern, D. (1977). *The first relationship: Mother and infant.* Cambridge, MA: Harvard University Press.

Stern, D. N. (1983). Le but et la structure du jeu mere-nourrison/The goal and structure of mother-infant play. *Psychiatrie de l'Enfant, 26(1),* 193-216.

Stern, D. N. (1985). *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology.* New York: Basic Books.

Stern, D. N. (1990). *Diary of a baby.* New York: Basic Books.

Stern, D. N. (1992). The representation of relational patterns: Developmental considerations. In A. J. Sameroff & R. N. Emde (Eds.), *Relationship disturbances in early childhood: A developmental approach* (pp. 52-69). New York: Basic Books.

Stern, D. N. (1995). *The motherhood constellation: A unified view of parent-infant psychotherapy.* New York: Basic Books.

Thomas, A., & Chess, S. (1977). *Temperament and development.* New York: Bruner/Mazel.

Thomas, A., & Chess, S. (1982). Temperament and follow-up to adulthood. In *Temperamental differences in infants and young children* (Ciba Foundation Symposium No. 89) (pp. 168-175). London: Pitman Books.

Thomas, A., Chess, S., & Birch, H. (1968). *Temperament and behavioral disorders in childhood.* New York: New York University Press.

Thomas, A., Chess, S., & Korn, S. J. (1982). The reality of difficult temperament. *Merrill-Palmer Quarterly, 28,* 1-20.

Thomas, A., Chess, S., Birch, H. G., Hertzig, M., & Korn, S. (1963). *Behavioral individuality in early childhood.* New York: New York University Press.

Thomas, A., Chess, S., Sillen, J., & Mendez, O. (1974). Cross-cultural study of behavior in children with special vulnerabilities to stress. In D. F. Ricks, A. Thomas, & M. Roff (Eds.), *Life history research in psychopathology*, pp. 53-67. Minneapolis: University of Minnesota Press.

Towle, P. O., Farran, D. C., & Comfort, M. (1988). Parent-handicapped child interaction observational coding systems: A review. In K. Marfo (Ed.), *Parent-child interaction and developmental disabilities: Theory, research, and intervention,* pp.293-330.

Uzgiris, I. C., & Raeff, C. (1995). Play in parent-child interaction. In M. H. Bornstein (Ed.), *Handbook of parenting, Vol. 4: Applied and practical parenting*, pp. 353-376. Hillsdale, NJ: Lawrence Erlbaum.

van den Boom, D. C. (1994). The influence of temperament and mothering on attachment and exploration: An experimental manipulation of sensitive

responsiveness among lower-class mothers with irritable infants. *Child Development, 65,* 1457-1477.

van den Boom, D. C. (1995). Do first-year intervention effects endure? Follow-up during toddlerhood of a sample of Dutch irritable infants. *Child Development, 66,* 1798-1816.

Van Heck, G. L. (1991). Temperament and the person-situation debate. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement, pp.163-175.* London: Plenum Press.

Vietze, P. M., & Anderson, B. J. (1981). Styles of parent-child interaction. In M. Begab, H. C. Haywood, & H. Garber (Eds.), *Psycho-social influences in retarded performance,* pp. 255-283. Baltimore: University Park Press.

Wapner, S. (1993). Parental development: A holistic, developmental, systems-oriented perspective. In J. Demick, K. Bursik, & R. DiBiase (Eds.), *Parental development,* pp. 3-38. Hillsdale, NJ: Erlbaum.

Weiss, S. J., Wilson, P., Seed., M. St. J., & Paul, S. M. (2001). Early tactile experience of low birth weight children: Links to later mental health and social adaptation. *Infant and Child Development, 10,* 93-115.

Wheeden, C. A., Mahoney, G., Fewell, R., & Spiker, D. (1997, March). *The effects of a comprehensive early intervention program on mother-child interaction with low birthweight, premature infants.* Paper presented at the Society for Research in Child Development Biennial Meeting, Washington, D. C.

Wilson, R. S., & Matheny, A. P., Jr. (1986). Behavior-genetics research in infant temperament: The Louisville Twin Study. In R. Plomin & J. Dunn (Eds.), *The study of temperament: Changes, continuity, and challenges.* Hillsdale, NJ: Erlbaum.

Windle, M. & Lerner, R. M. (1984). The role of temperament in dating relationships among young adults. *Merrill-Palmer Quarterly, 30,* 163-175.

Windle, M. & Lerner, R. M. (1986a). Reassessing the dimensions of temperamental individuality across the life span: The Revised Dimensions of Temperament Survey (DOTS-R). *Journal of Adolescent Research, 1,* 213-230.

Windle, M., & Lerner, R. M. (1986b). The "goodness of fit" model of temperament-context relations: Interaction or correlation? In J. V. Lerner & R. M. Lerner (Eds.), *Temperament and social interaction in infants and children,* pp. 109-120. San Francisco: Jossey-Bass.

Windle, M., Hooker, K., Lenerz, K., East, P. L., Lerner, J. V., & Lerner, R. M. (1986). Temperament, perceived competence, and depression in early and late adolescents. *Developmental Psychology, 22,* 384-392.

Zaslow, M. J., Pedersen, F. A., Suwalsky, J. T., & Rabinovich, B. A. (1990). Maternal employment and parent infant interaction at one year. *Early Childhood Research Quarterly, 4,* 459-478.