Prematurity Stereotyping and Mothers’ Interactions With Their Premature and Full-Term Infants During the First Year

Marilyn Stern, 1 Ph.D, Katherine Karraker, 2 Ph.D, Bonnie McIntosh, 3 Ph.D, Sara Moritzen, 4 Ph.D, and Michelle Olexa, 5 Ph.D

1Virginia Commonwealth University, 2West Virginia University, 3Cincinnati Children’s Hospital, 4University at Albany, and 5Shippensburg University Counseling Center

Objective To longitudinally assess stability and correlates of prematurity stereotyping and perceptions of infant vulnerability in mothers of premature (N = 56) and full-term (N = 59) infants. Method At 5, 9, and 12 months, mothers rated videotapes of unfamiliar infants with a full-term label (FTI) or a preterm label (PL), interacted with their own infant, and completed other questionnaires. A subgroup of infants were administered a developmental assessment at 32 months. Results Mothers rated PL infants more negatively than FTI infants at each age. Individual differences in stereotyping were not stable. Mothers who negatively rated infants labeled with the same birth status of their own infant exhibited more negative interactive behaviors with their infants. Mothers who viewed their own infant as more vulnerable and who showed more prematurity stereotyping at 5 months had infants with lower 32-month mental scores. Conclusions The results suggest an association between early maternal cognitions and both contemporaneous maternal behavior and later child developmental outcomes.

Key words mother–infant interactions; perceived vulnerability; premature infants; prematurity stereotyping.

Expectations play an important role in guiding all social interactions (Snyder & Stukas, 1999), and mothers’ perceptions of and expectations for their infants may have a critical impact on early mother–infant interactions (Olexa & Stern, 1999). These perceptions and expectations can affect both caregiving behavior and infant outcomes through a self-fulfilling prophesy process whereby mothers treat infants in accord with their perceptions and expectations, eventually producing the infant outcomes they expect. An illustration of this process is provided by a series of studies by Stern and colleagues (e.g., Stern & Karraker, 1988, 1989, 1992; Stern, Karraker, Meldrum-Sopko, & Norman, 2000) that demonstrated that mothers hold more negative perceptions about and behave in less desirable ways toward infants they believe were born prematurely than toward infants they believe were born at full term, independent of the infants’ actual characteristics or behavior. Stern and colleagues labeled this process “prematurity stereotyping” and have suggested that it may function to compromise the development of some prematurely born infants through the self-fulfilling prophesy process.

Most previous research on prematurity stereotyping has examined mothers’ and other adults’ prematurity stereotyping in the context of responding to unfamiliar infants in experimental labeling studies. These experiments (e.g., Stern & Karraker, 1988, 1989, 1992; Stern et al., 2000) assessed adults’ perceptions of and reactions to infants who were labeled premature or full-term independent of their actual birth status. This procedure allows the researchers to determine the extent to which adults’ perceptions and behaviors are influenced by beliefs about prematurity separately from any effects of the actual characteristics or behavior of the infants.
Collectively, these studies have shown that mothers of full-term infants, mothers of premature infants, mothers of ill full-term infants, and other adults hold negative beliefs about infants who are described as having been born prematurely, rating them as less physically mature in appearance, less physically potent, less sociable, less cognitively competent, less behaviorally mature, and less liked than when the same infants are described as having been born at full-term (Stern & Karraker, 1988, 1989, 1992). Furthermore, during interactions with labeled unfamiliar infants, mothers have been observed to treat infants who were labeled premature more negatively than infants who were labeled full-term. In turn, infants labeled premature responded to mothers differently from their full-term–labeled counterparts (e.g., Stern et al., 2000). These studies provide experimental evidence for the presence and functioning of a prematurity stereotype, but do not directly address whether prematurity stereotyping influences individual mothers’ interactions with their own premature infants. Support for this possibility comes from studies showing that some mothers treat their premature infants differently from full-term infants even when evidence is available that the premature infants are capable and competent and that the infants’ behavior is no different from that of their full-term counterparts (e.g., Barnard, Bee, & Hammond, 1984; Barratt, Roach, & Leavitt, 1996; Harrison & Magill-Evans, 1996).

Studies of mothers’ perceptions of infant vulnerability also provide parallel evidence for the impact of mothers’ perceptions on mother–infant interactions and infant outcomes. For example, Perrin, West, and Culley (1989) found that some mothers of prematurely born infants viewed their child as especially fragile or susceptible to damage in the absence of any objective data, even in comparison with mothers who had infants who were in fact objectively more at risk. Mothers who viewed their prematurely born infants as vulnerable have been shown to continue to treat their children differently from those mothers who did not view their children as vulnerable and, in turn, those infants perceived as highly vulnerable were more likely to manifest behavior problems and lowered developmental competence as compared with other infants (e.g., Estroff, Yando, Burke, & Snyder, 1994).

Little is known about the relation between prematurity stereotyping and perceptions of infant vulnerability. Prematurity stereotyping reflects a general belief that premature infants are more fragile and less capable than full-term infants, whereas a mother’s perception of her own infant’s vulnerability (to illness, injury, delayed development, etc.) may or may not relate to her infant’s birth status. Mothers of premature infants who engage in prematurity stereotyping may be more inclined to perceive their own premature infants as vulnerable, and either or both of these maternal cognitions may influence how these mothers interact with their infants.

The general goal of this study was to extend existing knowledge about prematurity stereotyping and perceptions of vulnerability by examining how individual differences in mothers’ tendencies to engage in prematurity stereotyping and to perceive their infants as vulnerable relate to mothers’ interactions with their own premature or full-term infants during the infants’ first year. An ancillary goal of this study was to examine the utility of the stereotyping assessment procedure used by Stern and colleagues as a tool for identifying mothers who show prematurity stereotyping, with the aim of ultimately using this procedure to identify these mothers so that any negative effects associated with stereotyping might be thwarted through targeted intervention. The following specific research questions were addressed:

1. Do mothers of premature and full-term infants show similar prematurity stereotyping of unfamiliar infants and perceptions of their own infants’ vulnerability?
2. Are individual differences in mothers’ prematurity stereotyping and perceptions of infant vulnerability stable across time?
3. Are individual differences in mothers’ prematurity stereotyping and perceptions of infant vulnerability related?
4. Are individual differences in mothers’ prematurity stereotyping and perceptions of infant vulnerability related to mother–infant interactions?
5. Are individual differences in mothers’ prematurity stereotyping and perceptions of infant vulnerability related to later infant outcomes?

Method
Participants

Mothers who had given birth to a premature infant who met inclusion criteria at three hospitals in northeastern New York State and mothers of full-term infants from the same area who were identified through newspaper birth listings were sent letters by hospital personnel or research staff inviting them to participate in the study. Appropriate Institutional Review Board (IRB) approvals were obtained before beginning recruitment. Approximately 25% of all mothers who received a letter contacted...
the researchers about participating in the project, and 22% subsequently agreed to participate. The letter sent to the mothers indicated that the study would include assessments at 5, 9, and 12 months following the infant's expected date of birth.

Of 115 mothers, 59 with their full-term infants (29 male and 30 female) and 56 with their premature infants (32 male and 24 female) participated in the 5-month assessment. From this group, 98 mothers (55 full term and 43 premature) participated in the 9-month assessment and, of these, 92 (52 full term and 40 prema-
ture) participated in the 12-month assessment. Mothers who completed these assessments were later contacted and invited to participate in a previously unplanned follow-up assessment (for which IRB approval was first obtained) when their child was between 25 and 35 months (M age = 31.7 months). Only 28 mothers of full-term and 28 mothers of premature infants could be contacted and were available and willing to participate in this assessment.

The sample of mothers participating in the study can generally be characterized as middle class. Nearly all of the mothers and infants enrolled in the study were Caucasian and lived in suburban areas (98%), and all of the mothers in the full-term infant group and all but four of the mothers in the premature infant group were married or cohabiting with the father of their infant. Most mothers of full-term infants (72%) and premature infants (60%) reported at least some college education. The mean age for mothers of full-term infants was 31.64 years, and the mean age for mothers of premature infants was 30.24 years. Nearly one half of both mothers of full-term infants (44.1%) and mothers of premature infants (46.3%) were primiparous.

Only those prematurely born infants who had spent at least 2 days in the Neonatal Intensive Care Unit (NICU), but who had not been seriously ill following birth, were selected for the study. Relatively healthy infants were targeted for the study, because prematurity stereotyping effects are more easily discerned when infant characteristics contradict mothers’ stereotyped perceptions and expectations and because prematurity stereotyping effects may be most problematic for premature infants who have few difficulties resulting directly from their prematurity. Excluded from the study were infants with obvious neurological damage, infants born to mothers with drug addictions, infants with serious handicapping conditions (such as, cerebral palsy, genetic disorders, and other conditions known to lead to delayed development), and infants who were born at less than 28 weeks of gestation. Most of the prematurely born infants in the study (70%) were in the NICU for observation or for a variety of minor medical problems, including minor bowel difficulties, infections, jaundice, and abnormal blood sugar. The remaining prematurely born infants (30%) were diagnosed with varying types of respiratory disorders (fluid in the lungs, trouble breathing, respiratory distress syndrome, etc.). The mothers and infants in the full-term and premature groups differed only in infant gestational age at birth, t(114) = 7.70, p = .001 (full-term M = 39.25, range from 38 to 42 weeks; premature M = 33.67 weeks, range from 28 to 36 weeks), birth weight, t(114) = 13.68, p = .001 (full-term M = 3401.94; premature M = 1950.45 g), and length of hospital stay, t(114) = 7.07, p = .001 (full-term M = .08; premature M = 16.05 days, range from 2 to 56 days). The premature infants in the study were not in any special follow-up or intervention programs.

Procedure

Mothers were engaged in a series of standardized interaction tasks with their own infant and then viewed and rated videotapes of two other infants and completed a series of questionnaires during the 5-, 9-, and 12-month assessments. Toddlers were administered a development-tal assessment, and mothers completed several questionnaires at the 32-month assessment.

Mother–Infant Interactions

Mothers were instructed to play with their infants as they did at home, with and without the use of any toys for 10 min. Interactions were videotaped with one camera focused on the infant (who was seated in an infant seat on a table at 5 months and in a high chair when 9 and 12 months) and another focused on the mother while also capturing the interaction with the infant from a mirror positioned behind the mother. Following the interaction, mothers were asked to rate their infants’ behavior and mood during the session on a five-point scale ranging from “unusually irritable” to “unusually quiet.”

Stereotyping Assessment

Mothers were shown two 5-min videotapes of neutrally clothed healthy male full-term infants of the same age as the mothers’ own infants, one labeled full-term and one labeled premature, playing with a set of age-appropriate, non–sex-typed toys while seated in an infant seat (at 5 months) or a high chair (at 9 and 12 months). The infants’ mothers were not visible. Because of the general absence of infant gender effects in previous studies on the prematurity stereotype (e.g., Stern & Karraker, 1992; Stern et al., 2000), gender was held constant
rather than varied by using only male infant stimuli. Mothers read a brief description of each infant, which included the appropriate birth status label, before viewing each videotape. Mothers were told that the age of the infant labeled premature was corrected (i.e., the age was based on the infant’s expected date of birth). To control for any potential effects due to order of presentation, order of labels, or specific videotape, the presentation of the two infant stimulus videotapes was counterbalanced so that an approximately equal number of mothers viewed each of the infant stimuli presented first or second and labeled as either full term or premature. After viewing each infant, mothers completed a set of ratings of the infant, indicated which of three toys they thought would be most appropriate for the infant (at 9 and 12 months only).

Measures

Mothers’ Ratings of Videotaped Infants
Mothers completed the Stern–Karraker Infant Stereotyping Scale (S-KISS) following their observation of each of the videotapes. The S-KISS\(^1\) consists of 18 seven-point adjective scales that have been used in all previous studies examining prematurity stereotyping. The 18 rating scale items were grouped according to conceptual relatedness as in previous studies (see Stern & Karraker, 1992; Stern et al., 2000; for items and groupings), resulting in four composite scores: sociability (e.g., happy/sad), cognitive competence (e.g., attentive/inattentive), appealing behavior (e.g., fun/not fun to play with), and physical potency (e.g., strong/weak). The internal consistency of each scale (coefficient alpha calculated within each age by birth status label group) ranged from .60 to .81. To assess individual differences in stereotyping, S-KISS scores were summed across all 18 items for the full-term label (FTL) and preterm label (PL) conditions separately (details about these analyses are reported below). The internal consistencies for the set of items included in these summary scores were \(\alpha = .82\) for both the FTL and PL conditions at 5 months, \(\alpha = .85\) for the FTL condition and .89 for the PL condition at 9 months, and \(\alpha = .89\) for both the FTL and PL conditions at 12 months.

Mothers’ Toy Choice
As in previous studies (e.g., Stern et al., 2000), mothers were asked to select which of three non–sex-typed, similarly colored toy rattles they thought would be most appropriate for each videotaped infant they viewed. A different set of toys was used at each age. A group of 48 middle-class mothers of full-term infants from another study (Sullum, 2000) were shown each set of toys and asked to rank order the toys for developmental appropriateness. Based on these rankings, within each of the three age levels, the least mature toy was assigned a value of 1, the intermediate toy was assigned a value of 2, and the most mature toy was assigned a value of 3.

Play Program Choice
After viewing each of the stimulus infants at the 9- and 12-month sessions, mothers were instructed to choose the most appropriate play program for the infant viewed. The program descriptions included identical infant activities, but one program was described as being appropriate for infants up to 9 months (or 12 months) and the other program was described as being appropriate for infants 9 months (or 12 months) and older.

Observers’ Ratings of Mother and Infant Interaction Behaviors
Trained observers rated the behavior of each mother and infant at 5, 9, and 12 months from videotapes using Field’s Interaction Rating Scales (FIRS; Field, 1980). Ten items focus on the mother’s behavior, and seven items focus on the infant’s behavior, with each item rated on a three-point Likert scale with higher ratings being optimal. Total scores for mothers and infants also were calculated. Because the FIRS coding focuses on specific behaviors rather than more general patterns of behavior, coders also were trained to assess three components of maternal behavior in a relational context derived from the Emotional Availability Scales (EAS; Biringen & Robinson, 1991; Biringen, Robinson, & Emde, 1988). These nine-point rating scales included maternal sensitivity, maternal intrusiveness, and maternal hostility. Finally, because the FIRS has been primarily used only with young infants (Field, 1980; Hart, Field, Stern, \& Jones, 1997), an additional coding scheme was applied to the data to more fully capture individual differences in mother–infant interactions at 9 and 12 months. Items from a behavior rating scale developed by Vaughn, Taraldson, Crichton, and Egeland (1980) were rated from videotapes of mother–infant interaction at 9 and 12 months only. Detailed criteria for each item were obtained from the authors. Total scores were calculated separately for both maternal and infant behaviors. Individual items for each of these rating scales are summarized in Tables I and II.

Two coders were kept blind to all research hypotheses and were trained to rate the videotaped interactions until they reached an inter-rater agreement rate of .85 for each variable (using Cohen’s Kappa). These assistants

\(^1\)Copies of the entire questionnaire can be obtained from the first author.
### Table I. Means for Maternal Behaviors and Vulnerable Child Scale (VCS) Scores at 5, 9, and 12 Months by Type of Mother and Mothers’ Stereotyping Group

<table>
<thead>
<tr>
<th>Behavior</th>
<th>5 Months</th>
<th>9 Months</th>
<th>12 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mothers of premature infants</td>
<td>Mothers of full-term infants</td>
<td>Mothers of premature infants</td>
</tr>
<tr>
<td></td>
<td>FTL+</td>
<td>PL+</td>
<td>FTL+</td>
</tr>
<tr>
<td><strong>FIRS ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State</td>
<td>2.83</td>
<td>2.96</td>
<td>3.00</td>
</tr>
<tr>
<td>Head orientation</td>
<td>2.86</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Facial expression</td>
<td>2.78</td>
<td>2.89</td>
<td>3.00</td>
</tr>
<tr>
<td>Imitative behavior</td>
<td>1.00</td>
<td>1.18</td>
<td>1.10</td>
</tr>
<tr>
<td>Game playing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>21.78</td>
<td>23.00</td>
<td>22.90</td>
</tr>
<tr>
<td><strong>EAS ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrusiveness</td>
<td>5.32</td>
<td>4.61</td>
<td>5.10</td>
</tr>
<tr>
<td>Hostility</td>
<td>1.13</td>
<td>1.04</td>
<td>1.03</td>
</tr>
<tr>
<td><strong>Vaughn et al. ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patience</td>
<td>5.68</td>
<td>6.26</td>
<td>6.08</td>
</tr>
<tr>
<td>Attitude toward play</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression of positive regard</td>
<td>5.75</td>
<td>6.25</td>
<td>5.96</td>
</tr>
<tr>
<td>Expression of negative regard</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression of delight</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>66.13</td>
<td>69.71</td>
<td>70.23</td>
</tr>
<tr>
<td><strong>Vulnerable Child Scale scores</strong></td>
<td>51.45</td>
<td>52.35</td>
<td>53.06</td>
</tr>
</tbody>
</table>

FTL+, mothers who rated full-term label (FTL) infants more positively; PL+, mothers who rated preterm label (PL) infants more positively. Means are provided only when a t test comparison of the FTL+ and PL+ groups was significant at p < .05. The score indicating more positive behavior in each pair of means is underlined.

*There were no effects on the Field's Interaction Rating Scales (FIRS) for physical activity, gaze behavior, silence, frequency of vocalization, and contingency with infant cues.

*There were no effects in the Vaughn et al. (1980) ratings for reaction to being observed, inventiveness in play, support or involvement, reciprocal play, or quality of verbalizations.

### Table II. Means for Infant Behaviors at 5, 9, and 12 Months by Type of Mother and Mothers’ Stereotyping Group

<table>
<thead>
<tr>
<th>Behavior</th>
<th>5 Months</th>
<th>9 Months</th>
<th>12 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mothers of premature infants</td>
<td>Mothers of full-term infants</td>
<td>Mothers of premature infants</td>
</tr>
<tr>
<td></td>
<td>FTL+</td>
<td>PL+</td>
<td>FTL+</td>
</tr>
<tr>
<td><strong>FIRS ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical activity</td>
<td>2.68</td>
<td>2.95</td>
<td>3.00</td>
</tr>
<tr>
<td>Head orientation</td>
<td>1.60</td>
<td>1.90</td>
<td>2.20</td>
</tr>
<tr>
<td>Gaze behavior</td>
<td>2.00</td>
<td>2.20</td>
<td>2.45</td>
</tr>
<tr>
<td><strong>Vaughn et al. ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General activity level</td>
<td>3.17</td>
<td>3.31</td>
<td>3.29</td>
</tr>
<tr>
<td>Satisfaction in play</td>
<td>4.75</td>
<td>5.06</td>
<td>5.13</td>
</tr>
</tbody>
</table>

FTL+, mothers who rated full-term label (FTL) infants more positively; PL+, mothers who rated preterm label (PL) infants more positively. Means are provided only when a t test comparison of the FTL+ and PL+ groups was significant at p < .05. The score indicating more positive behavior in each pair of means is underlined.

*There were no effects on the Field's Interaction Rating Scales (FIRS) for state, facial expression, fussiness, vocalizations, or total score.

*There were no effects in the Vaughn et al. (1980) ratings for reaction to situation, responsiveness to mother, disposition, predominant state, attention during play, or total score.
subsequently coded the remaining videotaped interactions independently. Periodic checks of inter-rater reliability (usually after every eight tapes) were conducted to ensure against rater scoring drift. On any occasion when Kappa values fell below .85, raters were re-trained until acceptable reliability levels were again achieved.

The Vulnerable Child Scale (Perrin et al., 1989)
At the end of the 5-month session, and then again at 32 months, mothers were asked to complete the Vulnerable Child Scale (VCS), a scale consisting of 16 four-point items. The statements express various concerns about a child’s health or well-being. The VCS has been reported to have adequate internal consistency (α = .75) and excellent test–retest reliability (r = .95) (Perrin et al., 1989). Because of its purported stability, mothers were not readministered the VCS at the 9- and 12-month testing sessions.

The Mental Scale of the Bayley Scales of Infant Development (2nd Edition) (Bayley, 1993)
Infants were tested with the Bayley Scales of Infant Development (2nd edition) Mental Scale (BSID-M) at approximately 32 months of age (corrected for those born prematurely). Test administrators were not aware of the responses of mothers or infants during any of the earlier assessments, although they were aware of whether the child was born prematurely or at full-term (to properly adjust for degree of prematurity in administration and scoring of the scale).

Results
Preliminary analyses revealed that the mothers and infants who completed the study were not different from the mothers and infants who dropped out at each assessment point, and that none of the dependent variables were associated with infant gender, maternal parity, marital status, maternal age, or maternal education. At each testing session, no difference was found between the ratings of infant mood by mothers of full-term and preterm infants, and the mean rating was found to be around the midpoint of the scale.

Research Question 1: Group Differences in Prematurity Stereotyping and Perceptions of Vulnerability
Multivariate and univariate analyses of variance (MANOVAs and ANOVAs) were conducted to assess the effects of type of mother (full-term or premature), birth status label (FTL or PL), and infant age (5, 9, or 12 months) on S-KISS ratings, toy choice, and play program choice, with repeated measures on the last two factors. Follow-up univariate findings are reported only when multivariate effects first reach significance. A t test was performed to compare VCS scores of mothers of full-term and premature infants.

S-KISS Ratings
A main effect for birth status label was found, multivariate F(4, 86) = 7.30, p = .038, η² = .17; however, a higher order two-way interaction involving birth status label and infant age also reached significance, multivariate F(8, 352) = 7.30, p = .001, η² = .17. Follow-up univariate tests revealed that this interaction held for sociability, F(2, 178) = 3.53, p = .03, cognitive competence, F(2, 178) = 14.98, p = .001, and physical potency, F(2, 178) = 27.97, p = .001. Simple main effects tests were conducted to determine when the birth status label effect indicated prematurity stereotyping (i.e., when the FTL infant was rated more positively than the PL infant). At 5 months, FTL infants were rated as more sociable, F(1, 114) = 3.81, p = .05 (M FTL = 16.52; M PL = 16.10), cognitively competent, F(1, 114) = 15.92, p = .001 (M FTL = 16.06; M PL = 14.63), and physically potent, F(1, 114) = 20.27, p = .001 (M FTL = 24.51; M PL = 21.60), than PL infants. At 9 months, FTL infants were rated as more cognitively competent, F(1, 97) = 4.34, p = .04 (M FTL = 15.43; M PL = 14.83), and physically potent, F(1, 97) = 7.41, p = .008 (M FTL = 22.39; M PL = 20.67), than PL infants. A similar pattern of findings was found at 12 months: FTL infants were rated as more cognitively competent, F(1, 91) = 6.17, p = .015 (M FTL = 16.44; M PL = 15.31), and physically potent, F(1, 91) = 12.84, p = .001 (M FTL = 25.15; M PL = 22.49), than PL infants. No other multivariate main effects or interactions reached significance.

Toy Choice
A three-way interaction involving type of mother, birth status label, and infant age was found, F(2, 182) = 3.88, p = .022, η² = .04. In addition, an interaction between birth status label and infant age, F(2, 182) = 9.13, p = .001, η² = .09, as well as a main effect for birth status label, F(1, 91) = 12.28, p = .001, η² = .12, were subsumed by the three-way interaction. Follow-up analyses indicated that the interaction between birth status label and infant age held for mothers of premature infants, F(2, 76) = 9.53, p = .001, but not for mothers of full-term infants, F(2, 106) = 1.88, p = .16. Follow-up simple main effects tests designed to assess prematurity stereotyping indicated that mothers of premature infants chose an older toy for FTL than for PL infants at 5 months, F(1, 55) = 4.25, p = .04 (M FTL = 2.23 and
M PL = 1.90) and 12 months, \( F(1, 39) = 3.12, p = .05 \) (M FTL = 1.95 and M PL = 1.76) but that there was no difference in toy choice at 9 months (M FTL = 1.58 and M PL = 1.49).

**Play Program Choice**

Because play program choice was only measured at 9 and 12 months, the infant age variable had only two levels in this analysis. A main effect for birth status label, \( F(1, 91) = 27.02, p = .001, \eta^2 = .23 \), indicated that regardless of infant age, mothers were more likely to choose the older play program for FTL infants (M = 1.60) than for PL infants (M = 1.50). No other main effects or interactions were found for this variable.

**VCS Scores**

Mothers of premature infants did not rate their own infants as more vulnerable than mothers of full-term infants rated their own infants, \( t(107) = 1.22, p = .23 \).

**Research Questions 2 Through 5: Individual Differences in Prematurity Stereotyping and Perceptions of Infant Vulnerability**

A stereotyping score was calculated for each mother at each time of assessment.\(^2\) Positive values of this score indicate more positive perceptions of FTL infants, and negative values indicate more positive perceptions of PL infants. To assess relations between concurrent stereotyping and mothers' interactions with their own infants, mothers who perceived FTL infants more positively than PL infants were compared with mothers who perceived PL infants more positively than FTL infants at each age within each group of mothers (mothers of full-term infants and mothers of premature infants) using \( t \) tests.\(^3\) Similar analyses were conducted to evaluate the relation between stereotyping and perceived vulnerability at 5 months.

**Descriptive Information**

About half the mothers at each age rated FTL infants more positively than PL infants (56, 53, and 64\% of mothers of full-term infants, and 45, 52, and 51\% of mothers of premature infants, at 5, 9, and 12 months, respectively). Comparisons of these distributions at each age using chi square were all nonsignificant.

**Research Question 2: Stability of Prematurity Stereotyping and Perceptions of Infant Vulnerability**

Stereotyping scores were not stable: at 5 and 9 months, \( r(97) = .14 \); 5 and 12 months, \( r(91) = .12 \); 9 and 12 months, \( r(91) = .10 \), all \( p \) values >.05. Thus, individual differences in stereotyping effects were examined separately at each age in subsequent analyses. The 5- and 32-month VCS scores were significantly correlated, \( r(44) = .36, p = .013 \).

**Research Question 3: Relations Between Prematurity Stereotyping and Perceptions of Infant Vulnerability**

A significant interaction between type of mother and whether mothers rated FTL or PL infants more positively at 5 months was found to influence 5-month VCS scores, \( F(1, 105) = 5.65, p = .019 \) (see means in Table I). Mothers of premature infants who rated PL infants more positively and mothers of full-term infants who rated FTL infants more positively perceived their own infants as less vulnerable than did other mothers. Note that the correlation between the number of days the premature infants spent in the NICU and their mothers' VCS scores was not significant, \( r = .10, p > .10 \), suggesting that mothers' perceptions of vulnerability were not based on objective indicators of vulnerability.

\(^2\)Before individual differences in prematurity stereotyping could be examined, it was necessary to adjust mothers' S-KISS scores for the effects of the particular videotapes they rated and the order in which they observed the videotapes. These factors were counterbalanced across mothers to allow unbiased assessment of group differences in prematurity stereotyping. Mothers' S-KISS scores for each videotape were first summed, and then residual scores for these sums were determined after entering the effects of videotape and order via multiple regression. These residual scores were calculated separately at each assessment point (5, 9, and 12 months) and for each labeled stimulus (FTL and PL). Then, the S-KISS residual score for the PL stimulus was subtracted from the S-KISS residual score for the FTL stimulus, resulting in a stereotyping score for which positive values indicate more positive perceptions of FTL infants and for which negative values indicate more positive perceptions of PL infants.

\(^3\)Although these effects could be assessed within the context of a series of type of mother \( \times \) type of stereotyping MANOVAs, main effects were of little relevance to the goals of these analyses and interactions would still have to be assessed using follow-up \( t \) tests. The general conclusions arrived at using the present strategy, and the MANOVA strategy is virtually identical. The MANOVA results can be obtained from the authors. Because of the number of \( t \) tests conducted using this analytic strategy, the risk of finding chance results was high. However, the results (reported below) indicate that the proportion of significant \( t \) tests was substantially higher than the 5\% expected by chance. Across all the analyses reported in Tables I and II, 46 of 196, or 23.47\%, of the \( t \) tests were significant, and all of the significant effects were in the same direction.
Research Question 4: Relations Between Prematurity Stereotyping and Perceptions of Infant Vulnerability and Mother–Infant Interaction

Tables I and II provide means for the measures of mother–infant interaction at each age, according to type of mother and whether the mothers rated FTL or PL infants more positively. Means are shown only in cases where the t test comparing mothers who rated FTL infants more positively with mothers who rated PL infants more positively was significant for a particular type of mother (mothers of full-term infants or mothers of premature infants). In general, mothers of premature infants who rated PL infants more positively and mothers of full-term infants who rated FTL infants more positively showed more positive interactions with their own infants.

Additional analyses were conducted to determine whether mothers’ perceptions of their own infant’s vulnerability at 5 months related to concurrent mother–infant interactive behaviors. For full-term infants, only infant fussing was related to VCS scores. The more mothers perceived their full-term infants as vulnerable, the more their infants fussed during the interaction, \( r_{(57)} = -0.22, p = .047 \). For premature infants, both mother and infant behaviors were related to VCS scores. Mothers who perceived their premature infants as more vulnerable showed less positive facial expressions, \( r_{(50)} = -0.30, p = .02 \), more intrusiveness, \( r_{(50)} = 0.29, p = .02 \), and more hostility, \( r_{(50)} = -0.30, p = .002 \). Premature infants who were perceived by their mothers as more vulnerable engaged in less vocalization, \( r_{(50)} = 0.42, p = .001 \), showed less involvement, \( r_{(50)} = 0.24, p = .04 \), and received lower FIRS total scores, \( r_{(50)} = 0.24, p = .04 \).

Research Question 5: Relations Between Prematurity Stereotyping and Perceptions of Infant Vulnerability and Infant Outcomes

In mothers of premature infants, relatively more negative ratings of PL infants in comparison with FTL infants at 5 months (but not 9 or 12 months) were related to lower infant BSID-M scores at 32 months, \( r_{(23)} = -0.44, p = .02 \). Stereotyping and BSID-M scores were not significantly related at any age in mothers of full-term infants. Greater perceptions of infant vulnerability at 5 months were related to lower BSID-M scores at 32 months, \( r_{(41)} = 0.31, p = .029 \), for all mothers combined, but not for each group of mothers separately.

Discussion

The results of this study reiterate the existence of prematurity stereotyping in groups of mothers of both healthy premature and full-term infants. The study also furthers our understanding of the potential role of maternal cognitions, such as prematurity stereotyping and perceptions of infant vulnerability, in concurrent mother–infant relationships and subsequent infant development. The results suggest that individual differences in mothers’ perceptions of their own and other infants are related to their concurrent behavior with their own infant and may be related to their infant’s subsequent development.

As in previous research on prematurity stereotyping, mothers as a group responded more negatively to infants labeled as premature than to infants labeled as full-term, particularly in their perceptions of cognitive competence, physical potency, and appropriate play programs. In general, this stereotyping effect was consistent across age periods and did not differ between mothers of premature and full-term infants. Only when mothers were asked to select an appropriate toy for the labeled infants was there a difference between mothers of premature and full-term infants. Only mothers of premature infants selected a more immature toy for the prematurely-labeled infants than for the full-term–labeled infants. This bias toward providing less mature toys for premature infants may have resulted from the premature mothers’ experiences in providing toys for their own infants and may have reflected the mothers’ accurate knowledge about the need to make age adjustments when determining the age appropriateness of toys for young premature infants.

The most notable findings from this study concern individual differences in mothers’ prematurity stereotyping. First, the findings highlight the importance of recognizing that not all mothers rate premature-labeled infants more negatively than full-term–labeled infants (i.e., show prematurity stereotyping). Some mothers, including mothers of both premature and full-term infants, actually rated premature-labeled infants more positively than full-term–labeled infants. This finding emphasizes that assumptions about a mother’s cognitions about her own infant or about other infants should not be made based on group findings.

A second important finding concerning individual differences was that mothers’ ratings of premature-labeled and full-term–labeled infants were associated with the mothers’ concurrent behavior with their own infants, but only when the birth status of their own infants was considered. Mothers who gave higher ratings to infants like their own interacted more positively with their own infants. In turn, some evidence for a reciprocal impact on infants was found, with infants at 9 and
12 months responding more positively to mothers who gave higher ratings to infants with the same birth status as their own. These findings lead to the conclusion that prematurity stereotyping may be associated with less than optimal mother–infant interactions and infant outcomes for mothers of premature infants, but may facilitate or reflect positive interactions and outcomes for mothers of full-term infants. Social psychological research confirms the general benefits of perceiving others who are like oneself (the in-group) more positively than others who are not like oneself (the out-group), even when one’s own group is socially stigmatized (Crocker & Major, 1989; Walsh & Banaji, 1997). Perhaps mothers differed in the extent to which they identified their infant as belonging to the group of premature or full-term infants; in general, more positive attitudes are likely toward a group with which one strongly identifies (Greenwald et al., 2002). However, some mothers may also have engaged in downward comparisons with the out-group to improve their perceptions of their own infant (Greenwald et al., 2002). This study was not designed to identify the reasons for mothers’ variations in perceptions of premature and full-term infants, but the findings suggest that future study of these social perception processes could be informative.

An alternate explanation for the association between positive behavior toward one’s own infant and positive perceptions of infants like one’s own could relate to the order in which these measures were obtained in this study. Mothers first interacted with their own infants and then completed the stereotyping task. Perhaps mothers who experienced more positive interactions with their infants were subsequently disposed to perceive other infants like their own more positively, whereas mothers who experienced more difficult interactions subsequently perceived infants like their own more negatively. Similarly, more difficult interactions were associated with mothers’ perceptions of their own infants’ vulnerability, especially in mothers of premature infants. Because premature infants and their mothers often experience more interactional difficulties than full-term infants and their mothers (e.g., Goldberg & DiVitto, 2002), the effect of interactions on perceptions could set up an ongoing circular process whereby poor interactions confirm mothers’ negative stereotypes about premature infants and perceptions of their own infants’ vulnerability, perhaps then leading to continued poor interactions and lowered expectations for their infants.

Suggestive evidence that mothers’ early cognitions about their infants can predict later developmental differences was obtained in the follow-up portion of this study. Mothers of both full-term and premature infants who perceived their infants as more vulnerable and mothers of premature infants who engaged in more prematurity stereotyping at 5 months had infants who obtained lower BSID-M scores at 32 months. The observed associations of both vulnerability perceptions and prematurity stereotyping with mother–infant interactive behaviors suggest that mothers with these negative perceptions provide less stimulating and positive social experiences for their infants, which may in turn lead to lowered developmental status in toddlerhood. Alternatively, lower developmental competence in early infancy may directly predict continued poor development and be associated with negative maternal perceptions and behaviors. Although prematurity stereotyping at 9 and 12 months was not found to relate to later BSID-M scores, further research on the relations between maternal cognitions and infant development is warranted by these findings. Questions of why individual differences in prematurity stereotyping as measured in this study were not stable and why relations with BSID-M scores were found only for measures obtained at 5 months will need to be explored.

This study has several limitations, including characteristics of the sample (self-selection, selective dropout, and homogeneity, leading to restricted generalizability), lack of consideration of other factors that may influence mothers’ perceptions (such as, infant temperament or developmental competence and maternal personality or prior experiences), and use of only one measure of developmental outcome at one point in time. Despite these limitations, the findings from this study suggest some practical implications for assessment and intervention with mothers of healthy premature infants. The stereotyping assessment procedure used here along with the VCS show promise as means for identifying the subgroup of mothers at risk for poor mother–infant interactions and subsequent compromised infant development because of negative prematurity stereotyping or unsubstantiated perceptions of their infants’ vulnerability. Another recent study (Durrett, Foster, Orrell, & Stern, 2005) has documented the feasibility of assessing perceptions of vulnerability and prematurity stereotyping in a clinical setting and has shown that doing so may yield useful clinical information for subsequent targeted intervention. These findings also point to the potential value of early cognitively focused interventions with this subgroup of mothers of premature infants. The ideal time during the reproductive and childbearing cycle to implement cognitively oriented intervention programs, as well
as precisely how to go about changing mothers’ cognitions, remain to be determined (McIntosh, Stern, & Ferguson, 2004; Stern & Bitsko, 2003).

Acknowledgment

This study was partially supported by a Social and Behavioral Sciences grant to the first author from the March of Dimes Birth Defects Foundation. We gratefully acknowledge the assistance of Jennifer Zak-Place, Puspa Das, and numerous other students who helped with the collection and coding of the data. We would also thank all the mothers who participated in the study and the staff in the Neonatal Intensive Care Units of St. Peters, Albany Medical Center, and Bellvue Hospitals who helped us identify participants for the study.

Received June 8, 2005; revisions received June 30, 2005; accepted July 5, 2005

References


Stern, M., & Karraker, K. (1989). Modifying the prematurity stereotype: The effects of information...


