Child Development, August 1997, Volume 68, Number 4, Pages 571–591

Sensitivity and Attachment: A Meta-Analysis on Parental Antecedents of Infant Attachment

Marianne S. De Wolff and Marinus H. van IJzendoorn

This meta-analysis included 66 studies (N = 4,176) on parental antecedents of attachment security. The question addressed was whether maternal sensitivity is associated with infant attachment security, and what the strength of this relation is. It was hypothesized that studies more similar to Ainsworth's Baltimore study (Ainsworth, Blehar, Waters, & Wall, 1978) would show stronger associations than studies diverging from this pioneering study. To create conceptually homogeneous sets of studies, experts divided the studies into 9 groups with similar constructs and measures of parenting. For each domain, a meta-analysis was performed to describe the central tendency, variability, and relevant moderators. After correction for attenuation, the 21 studies (N = 1,099) in which the Strange Situation procedure in nonclinical samples was used, as well as preceding or concurrent observational sensitivity measures, showed a combined effect size of $r(1{,}097) = .24$. According to Cohen's (1988) conventional criteria, the association is moderately strong. It is concluded that in normal settings sensitivity is an important but not exclusive condition of attachment security. Several other dimensions of parenting are identified as playing an equally important role. In attachment theory, a move to the contextual level is required to interpret the complex transactions between context and sensitivity in less stable and more stressful settings, and to pay more attention to nonshared environmental influences.

INTRODUCTION

In the first volume of his trilogy, *Attachment and Loss*, John Bowlby (1969) signaled an urgent need to determine the antecedent conditions that influence the development of attachment. Bowlby (1969) suggested that one of the conditions contributing to the development of a secure attachment relationship may be the attachment figure's sensitivity in responding to the baby's signals: When infants experience that their social initiatives are successful in establishing a reciprocal interchange with the mother, it is likely that an active and happy interaction between the couple will ensue and that a secure attachment relationship will develop.

After more than 25 years of research on the antecedents of attachment security, we may now be in a position to answer Bowlby's question. Is parental sensitivity indeed an important condition for the development of a secure attachment relationship between infant and parent?

Ainsworth and her colleagues were the first to examine the relation between parental behavior in the home and security of attachment (Ainsworth, Blehar, Waters, & Wall, 1978). They observed 26 middle-class mother-infant dyads from Baltimore throughout the first year of life; more than 70 hr of observation were spent in each home. At the time of the infant's first birthday, mother and infant came to the laboratory for assessment in the Strange Situation. This standardized procedure for assessing the infant-parent attachment relationship was developed in Ainsworth's pioneering study (Ainsworth & Wittig, 1969), and in the past few decades it has been used worldwide (Thompson, in press; van IJzendoorn & Kroonenberg, 1988). Ainsworth and her colleagues assessed a great variety of dimensions of maternal behavior at home. In particular, four rating scales (sensitivity, acceptance, cooperation, and accessibility) were found to be strongly related to attachment security, and the authors concluded that "the most important aspect of maternal behavior commonly associated with the security-anxiety dimension of infant attachment is manifested in different specific ways in different situations, but in each it emerges as sensitive responsiveness to infant signals and communications" (Ainsworth et al., 1978, p. 152).

Two decades after the Baltimore study, however, there is still great controversy over the parental antecedents of the Strange Situation attachment classifications. Gewirtz and Boyd (1977) and Lamb, Thompson, Gardner, and Charnov (1985) sparked this controversy by arguing that in her exploratory study Ainsworth had overgeneralized from the findings of her small sample. Lamb et al. contended that Ainsworth et al.'s (1978) "exciting hypotheses about the specific antecedents of Strange Situation behavior remain unproven except in their most general form" (p. 97). Although many studies had indicated that the infant's prior experiences at home were indeed related to their Strange Situation behavior, according to

Lamb et al. it is still unclear which specific maternal behaviors are of formative importance for attachment security. In the first meta-analysis on attachment and sensitivity, Goldsmith and Alansky (1987, pp. 811, 813) also concluded that "many of the studies . . . replicate Ainsworth et al.'s (1978) original findings of the predictive power of maternal sensitivity when replication is evaluated in terms of statistical significance." However, they cautioned that the actual size of the predictive effect of maternal sensitivity is much smaller than once was believed, suggesting only a weak relation between attachment security and parental sensitivity.

Despite all of the skepticism, many authors continue to embrace Bowlby's (1969) original proposition that maternal sensitivity is a crucial antecedent of attachment security (Bretherton, 1985; Main, 1990; Sroufe, 1988). Recently, Isabella (1993) stated that attachment theory and research highlight maternal sensitivity as an all-important characteristic of interaction that has been consistently linked to attachment security. The fact that some studies have yielded much weaker associations than Ainsworth's study can, in this view, be attributed to methodological weaknesses of the replication studies. Some researchers have restricted their observations to a single home visit, whereas others have used brief laboratory assessments of sensitivity instead of extensive home observations (e.g., Frodi, Grolnick, & Bridges, 1985). Some studies included interviews to assess positive parental attitude toward the infant (Benn, 1986), whereas others focused on the frequency of physical contact (Kerns & Barth, 1995). Although these approaches may tap into some dimension of a broad concept of parenting that is pertinent to the development of attachment, they may not capture the original concept of sensitivity. Studies designed according to the basic features and conceptualizations of Ainsworth's Baltimore study (e.g., Belsky, Rovine, & Taylor, 1984; Grossmann, Grossmann, Spangler, Suess, & Unzner, 1985; Isabella, 1993) have yielded results that are closer to the original findings (Pederson et al., 1990). This latter view, which we will refer to as the "orthodox" position, can be summarized as follows: (1) The mother's interactive behavior, in particular, her sensitivity, is considered to be the primary determinant of attachment quality; (2) however, if this association is to be detected, the observations of maternal behavior must be sufficiently intensive and reliable, and the observed dimensions of maternal behavior must be conceptually close to sensitivity.

In this meta-analysis, we integrate the available studies on parenting and attachment in a quantitative manner. In a controversial field like attachment theory, a narrative review may not contribute to resolution of the debate because of its "subjective" and less systematic nature (Bretherton, 1985; Lamb et al., 1985). Meta-analysis offers a way of bringing some degree of order to a large and inconsistent body of findings (Rosenthal, 1991) at the same time that it allows for the testing of specific hypotheses statistically. For example, one might test the idea that studies which show a greater similarity to the Baltimore study show stronger associations between attachment and sensitivity. Almost 10 years after Goldsmith and Alansky's (1987) meta-analysis of 13 studies, it is time to take stock of the growing literature on attachment and sensitivity and to focus on the core issue of variability among the pertinent studies. More than 60 studies have investigated maternal behavior in relation to attachment security. In the current meta-analysis, we address three issues (Mullen, 1989): (1) Central tendency: What is the typical strength of association between maternal behavior and attachment security? (2) Variability: Is the set of study results heterogeneous, in the sense that outcomes are relatively variable across the studies? (3) Prediction: Can the variation between studies be explained by study features that are relevant to the controversy between the skeptical and orthodox positions?

In particular, we test the "orthodox" hypothesis that studies which more closely resemble the original Baltimore study show the strongest associations between parenting and attachment, whereas studies deviating from this intensive, naturalistic longitudinal study yield less impressive results. More specifically, the following hypotheses are tested: (1) Stronger associations between maternal behavior and attachment security are found in studies defining maternal behavior as sensitivity as compared with studies defining maternal behavior differently. (2) Home-based studies show stronger associations between attachment and maternal behavior than do laboratory studies. (3) Long-term home observation studies show stronger relations between maternal behavior and infant attachment than do short-term, home-based studies. (4) Assessments of maternal behavior during the first year of life show stronger associations with attachment security than do assessments after age 1. (5) The longer the time interval between the assessment of maternal behavior and the Strange Situation procedure, the weaker the association between sensitivity and attachment (see Goldsmith & Alansky, 1987). Moreover, we will test whether the association between sensitivity and attachment is dependent on contextual factors such as the socioeconomic or (non-)clinical status of the families involved. The

orthodox hypothesis would emphasize a relatively context-free interpretation of this association. It may also be argued, however, that contextual factors partly override the influence of sensitivity on the development of attachment, and that in lower-class or clinical samples the association between sensitivity and attachment would be weaker.

In the first study, we report on the construction of homogeneous sets of concepts related to maternal interactive behavior, through a new method of expert sorting and rating procedures. Meta-analysis has often been criticized for mixing apples and oranges. Careful, systematic *conceptual* analysis is required to establish sets of similar studies that can be included in separate meta-analyses (Cooper & Hedges, 1994). This study paves the way for the second study in which the methods and results of a series of meta-analyses on the association between various aspects of maternal interactive behavior and attachment security are described. In the second study, the crucial issue of homogeneity will be addressed *statistically*.

STUDY 1

Method

Because the grouping of predictor variables may have an important effect on the outcome of a meta-analysis, experts were asked to categorize aspects of maternal interactive behaviors. In sorting the constructs, the experts were blind to the effect sizes associated with the specific constructs. In the total set of pertinent research papers (see Study 2), 55 different constructs were identified, all referring to various aspects of maternal behavior, for example, maternal body contact, maternal involvement, support, stimulation, sensitivity, verbal responsiveness, delight in the interaction, mutuality between mother and child, and frequency of positive responses to child. For each construct, we also identified the formal definition that was presented in the introduction or method sections of the research report, as well as the method of assessing that particular construct. On the basis of these descriptions, the experts were asked to sort the constructs. However, because the task of sorting 55 constructs would have been extremely time consuming, a smaller set was created. The first author identified four groups of constructs that were relatively self-evident. To reduce the total set, 15 constructs were preliminarily assigned to one of the four straightforward categories. These four groups encompassed the following aspects of maternal behavior: (1) Sensitivity, (2) Contiguity of Response, (3) Physical Contact, and (4) Cooperation.

The first group, Sensitivity, included all constructs that conformed to the original definition of Mary Ainsworth and her colleagues (Ainsworth, Bell, & Stayton, 1974): the mother's ability to perceive the infant's signals accurately, and the ability to respond to these signals promptly and appropriately. Constructs that had been assessed with Ainsworth's rating scale (Ainsworth et al., 1974) or a rating scale that was explicitly based on Ainsworth's original rating scale were also assigned to this first group. The second group of constructs was maternal Contiguity of Response, which we defined as promptness or, more generally, frequency of response to the infant's signals. The most important factor distinguishing contiguity of response from sensitivity is the absence of any qualitative assessment of the mother's behavior. Only promptness or frequency of the mother's responses contribute to contiguity of response; appropriateness is irrelevant. In the literature, contiguity of response is often referred to as "responsiveness." Because the term responsiveness is sometimes also used to indicate sensitivity (e.g., "sensitive responsiveness"), the concept "contiguity" was preferred to avoid confusion (see Bornstein, 1989). The third group consisted of all constructs that referred to quality or quantity of Physical Contact. The fourth group, Cooperation, included constructs that bore on the presence or absence of intrusive or interfering maternal behavior. This concept was originally defined and operationalized by Ainsworth (Ainsworth et al., 1974). The second author independently sorted 15 studies in the four categories. The percentage agreement was 93%, which documented the self-evident nature of this preliminary step in the sorting procedure.

Fifteen constructs fitted into one of the four groups; 40 constructs referring to other aspects of maternal interactive behavior remained to be classified. Two methods were used to create conceptually homogeneous groups of concepts: A sorting task in which the experts sorted the concepts into a few subgroups, and a rating task in which the experts rated each concept in terms of its similarity to Ainsworth's sensitivity construct and in terms of its importance for the development of attachment relationships.

Participants

A total of 27 persons who were actively involved in attachment research were asked to participate in the study. Experts were defined as persons who had been actively involved in attachment research for several years and who were at least participating in a graduate program in the behavioral sciences. Com-

plete data were obtained for 19 respondents (12 women). Because the sorting task was rather time consuming, four persons were unable to comply with our request. Four nonrespondents refused for other personal reasons. Almost half of the sample had obtained a doctoral degree ($n = 9$); the remaining participants were graduate students. The respondents had been actively involved in the attachment research for 5 years on average ($SD = 3.89$).

Materials

The Similarity Sorting Task was constructed to investigate how the experts evaluated the conceptual similarities between the concepts. The sorting task consisted of 40 small cards (15 × 10.5 cm), on each of which a particular concept was presented as well as a conceptual definition of that concept and, if necessary, the way in which it was assessed. On the cards, any information about the source of the concept was absent to keep the sorters blind as to the effect sizes related to the concepts. The experts were instructed to sort the cards into maximally 10 groups of relatively similar concepts. The maximum of 10 groups was chosen to facilitate the sorting task and to guarantee the statistical power of the subsequent analyses (Verkes, Van der Kloot, & Van der Meij, 1989). The rating task consisted of the same set of 40 cards with conceptual definitions. We now asked the experts to rate each concept on a 7 point scale in terms of its similarity to Ainsworth's sensitivity construct, which was presented along with the cards. This scale ranged from "not any similarity" (1) to "exact similarity" (7). Besides the similarity to sensitivity, we also asked the experts to rate each concept in terms of its importance for the development of an attachment relationship. Experts could indicate the concept's importance on a 7 point rating scale that ranged from "not at all important" (1) to "very important" (7).

Procedure

The two tasks and a written instruction were sent to the experts. We asked the experts to complete the Sorting Task before starting to accomplish the Rating Task.

Data Analysis

The sorting data were analyzed with Homogeneity analysis using alternating least squares (HOMALS; Gifi, 1990). We used the SPSS-PC program HOMALS (SPSS Inc., 1990). HOMALS can be considered as an equivalent to principal components analysis in the case of categorical data. A graphical configuration is constructed in which variables (i.e., the experts who sorted the concepts) and categories (i.e., the groups into which the concepts were sorted) are being represented. In calculating the coordinates of variables and categories, HOMALS uses a distance model: Those concepts that are sorted more frequently into the same group are represented in the configuration relatively close to each other. HOMALS has proved to be a fruitful method for analyzing sorting data of large numbers of stimuli (see Van der Kloot & Van Herk, 1991; Verkes et al., 1989). Following Verkes et al. (1989), we also performed a centroid cluster analysis (Everitt, 1974) on the distances between the 40 concepts in the HOMALS configuration to facilitate the interpretation of its results.

Results

The Similarity Sorting Task

Concepts that were sorted by an expert into the same group received the same numerical code. Concepts that were coded into different groups were given different numerical codes. Those concepts that were not grouped with any other concept were treated as missing data. In this way a 40 (concepts) × 19 (experts) data matrix was created.

HOMALS yielded a two-dimensional configuration that is shown in Figure 1. The configuration represented the similarities and dissimilarities of the 40 concepts for maternal interactive behavior. The two dimensions had satisfactory eigenvalues of .87 and .81. This means that the 40 concepts are clearly differentiated by the two dimensions of the HOMALS solution. To support the interpretation of the HOMALS configuration, we performed a hierarchical cluster analysis on the distances between all concepts. We used the linkage coefficient to assess the appropriate number of clusters to describe the data optimally (Aldenderfer & Blashfield, 1984). Five clusters were constructed, which are also presented in Figure 1.

Synchrony can be defined as "the extent to which interaction appeared to be reciprocal and mutually rewarding" (Isabella, Belsky, & Von Eye, 1989, p. 13). Asynchronous instances of maternal and infant behavior are "those considered to reflect one-sided, unresponsive, or intrusive behavioral exchanges" (Isabella & Belsky, 1991, p. 376). "Positive mutuality," a concept that was applied by Kiser, Bates, Maslin, and Bayles (1986), exemplifies the cluster *Mutuality*. Positive mutuality is a construct that consists of the following maternal behaviors: Number of "positive exchanges where both mother and infant attend to the same thing," and "the mother's skill at modulating

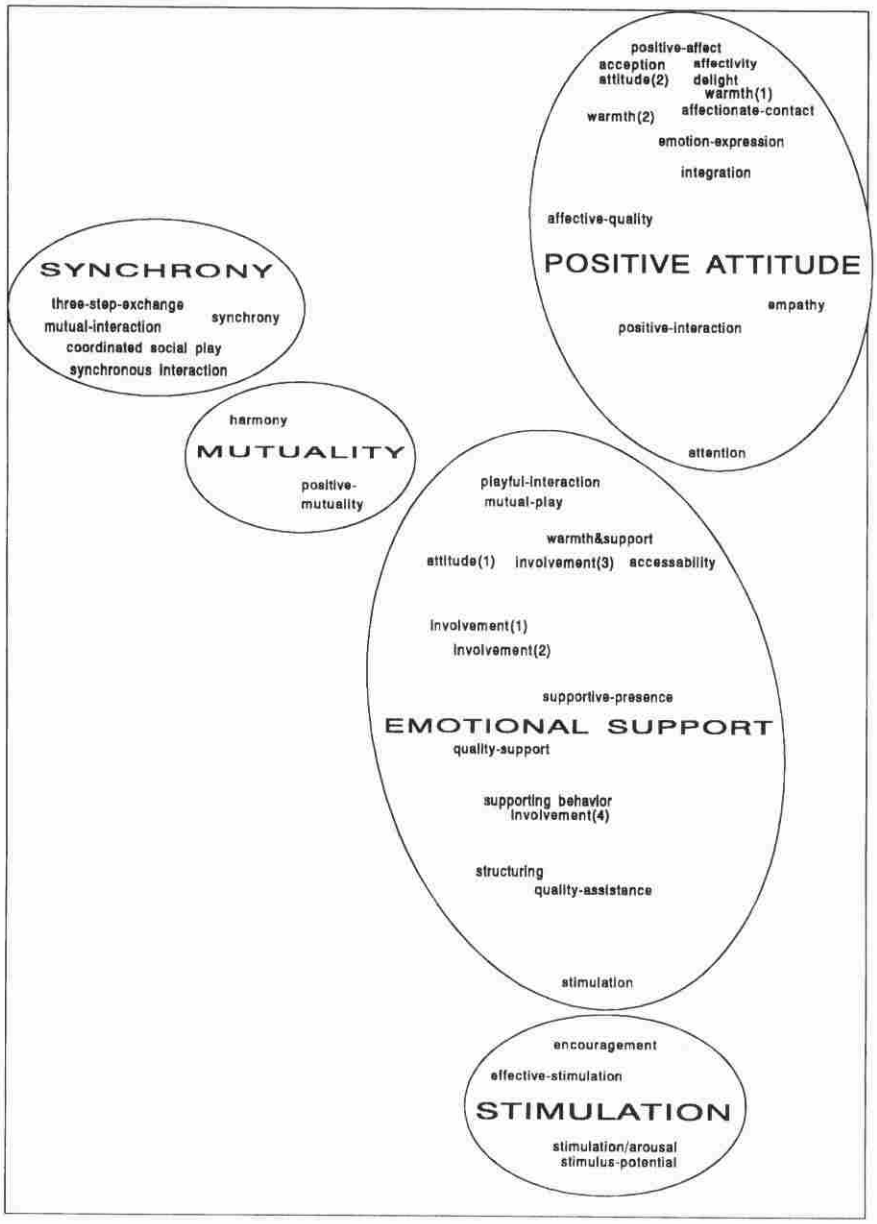

Figure 1 Two-dimensional representation of 40 concepts in the HOMALS solution

the baby's arousal, her entertainment value, and her responsiveness to the infant's cues" (p. 71); it also includes some infant behaviors: "Expression of positive affect, nonavoidance, active maintenance of the interaction, and amount of gazing at the mother" (p. 71). A central concept in the cluster *Support* is Supportive Presence, which was introduced by Erickson, Sroufe, and Egeland (1985). Matas, Arend, and Sroufe (1978, p. 350) defined this concept as follows: "The extent to which the mothers appeared attentive and available to the children and supportive to their efforts. A high score on supportive presence involved meeting two criteria: (a) Providing a secure base by helping the child feel comfortable, and (b) being involved as manifested by the attentiveness to the child and to the task." Affective Quality (Zaslow, Rabinovich, Suwalsky, & Klein, 1988) is an important construct in the cluster of *Positive Attitude*. Zaslow et al. (1988, p. 290) defined this concept as "the mother's expression of positive affect to the baby, the mother's expression of negative affect to the baby, and the degree to which mother and infant engaged in reciprocal interactions." Finally, *Stimulation* can be described as "any action on the part of the mother directed toward her baby" (Miyake, Chen, & Campos, 1985, p. 292).

The Rating Task

In a preliminary principal components analysis for categorical data (PRINCALS; Gifi, 1990), it was established that the expert raters were unidimensional in their ratings on the two rating scales (Verkes et al., 1989). The sorting data revealed five distinct clusters of parenting concepts. For each cluster, we computed the mean ratings. The five cluster means were found to differ significantly from each other: On the first rating scale, similarity to Ainsworth's sensitivity concept, the overall $F(4, 35) = 5.52$, $p = .0015$; on the second rating scale, importance for attachment development, the overall $F(4, 35) = 5.03$, $p = .003$. Multiple comparison tests, using the Tukey-HSD procedure, indicated that this overall difference between the five clusters could be ascribed to two clusters that were located at the outer part of the HOMALS configuration: Positive Attitude and Stimulation. Concepts belonging to one of these two broader concepts received significantly lower ratings on both rating scales than the concepts belonging to the other clusters.

On the basis of this outcome, we distinguished in the total set of 40 concepts between a More Optimal Group, consisting of the clusters Synchrony, Mutuality, and Emotional Support, and a Less Optimal Group of concepts, consisting of the clusters Positive Attitude and Stimulation. On the rating scales, the two global clusters were significantly different. On the similarity scale, the group means were $M = 4.10$ and $M = 3.09$ for the More Optimal and Less Optimal clusters, $F(1, 35) = 16.34$, $p = .001$. On the importance scale, the mean rating for the More Optimal Cluster was $M = 4.79$ versus $M = 4.30$ in the Less Optimal Group, $F(1, 35) = 4.89$, $p = .02$.

In sum, we started with the more general construct of maternal interactive behavior, and then distinguished, partly with the help of experts, nine different groups, each referring to a narrower construct. For each group we will conduct a separate meta-analysis. The empirical approach of conceptual analysis through expert ratings enabled us to systematically differentiate between several dimensions of parenting behavior independent of the meta-analytic results. In Study 2, we describe the methods and results of the subsequent meta-analyses.

STUDY 2

Method

Selection of the Studies

To identify studies for inclusion in the meta-analysis, we applied three search strategies: Computerized searches, manual search procedures, and consultation of other scientists working in this field.

The following computerized abstracting services were used to locate studies: Psychological Abstracts (from 1974 on), Educational Resources Information Center (from 1983 on), and Social Sciences Citation Index (from 1983 on). Furthermore, we used "World Catalog," a database of the On-Line Contents Library Center (OCLC). The following keywords were used in different combinations: attachment, childrearing practices, infant, mother, mothering, mother-child interactions, mother-child relations, parent-child relations, parenting, responsiveness, and sensitivity.

To supplement these computer searches, we used several manual search procedures. Reference lists from existing reviews were inspected (Goldsmith & Alansky, 1987; Lamb et al., 1985; Lambermon, 1991), as were reference lists of the reviewed articles. In addition, we located unpublished dissertations through a manual search of *Dissertation Abstracts International* (from 1980 on). To locate unpublished research, we also worked through several volumes of conference abstracts (ICIS, SRCD). Finally, we asked two prominent researchers in the attachment field, Drs. M. Main and J. Belsky, to suggest additional studies. They were able to mention three additional studies. One of the reviewers mentioned three additional studies (Fagot & Kavanagh, 1990; Malatesta, Culver, Tesman, & Shepard, 1989; Seifer, Schiller, Sameroff, Resnick, & Riordan, 1996). After submission of the manuscript, we came across two very recent studies (Gunnar, Brodersen, Nachmias, Buss, & Rigatuso, 1996; NICHD Early Child Care Network, 1996).

Each study had to meet three criteria for inclusion in the meta-analysis. The first criterion was that the study contained a measure of the mother's behavior toward her infant, and a measure of the infant's attachment security. The vast majority of maternal behavior measures have been derived from observations of mother-infant interaction. In a few studies, maternal behavior was measured using questionnaires or interviews with the mother. For example, Benn (1986) and Bretherton, Biringen, Ridgeway, Maslin, and Sherman (1989) assessed maternal sensitivity in an interview with mother, whereas Izard, Heynes, Chisholm, and Baak (1991) used a questionnaire to assess the mother's style of emotional expression. In the assessment of attachment security, the Strange Situation procedure was, of course, used in most studies. In some studies, the original procedure was changed because the researchers adapted the procedure to a home situation, or because they wanted to limit the infant's stress (see Capps, Sigman, & Mundy, 1994). In other studies, the Waters

and Deane (1985) Attachment Q-Sort for observers was used instead of the Strange Situation (Vaughn & Waters, 1990). We did not include studies using the Mother Attachment Q-Sort, because a recent meta-analysis clearly showed its lack of convergent and discriminant validity (van IJzendoorn, Vereijken, & Riksen-Walraven, in press).

The second criterion for inclusion was that each study should report an association between maternal behavior and infant attachment, so that effect sizes could be calculated. Even if only nonnumerical information was provided (i.e., "no relation" or "significant relation"), we included the study in the meta-analysis and estimated the effect sizes (Mullen, 1989). In some studies in which both maternal behavior and quality of attachment were assessed, however, the researchers did not compute any association between the two assessments, nor did they report on the association (see Easterbrooks & Goldberg, 1985). These studies had to be excluded.

Our third criterion concerned intervention studies. The central hypothesis in our meta-analysis refers to the association between maternal behavior—as it occurs naturally—and infant attachment. Consequently, if maternal behavior was experimentally influenced through intervention or therapy, we only included data for the nontreated control group. Thus, only those intervention studies that reported separate data for the control group(s) could be included in the meta-analysis (for a meta-analysis of intervention studies, see van IJzendoorn, Juffer, & Duyvesteyn, 1995).

Overall, we adopted a fairly liberal inclusion strategy. No studies were excluded from the meta-analysis on the basis of flawed design. Both Hedges (1986) and Mullen (1989) advise against excluding studies of low quality, because it is very difficult to assess a diffuse dimension like "study quality" directly and appropriately. Instead of using study quality as a criterion for inclusion in the meta-analysis, the meta-analyst is advised to include studies of varying quality. More detailed information on study quality should be included in the coding system, so that the influence of study characteristics on effect size can be examined in the meta-analysis. In the area of attachment research, the quality of the studies remains a hotly disputed issue. The main body of studies on attachment has been published in leading journals such as *Child Development* ($k = 21$), *Developmental Psychology* ($k = 7$), *Infant Behavior and Development* ($k = 6$), the *Monographs of the SRCD* ($k = 3$), and in other refereed journals. A number of studies, however, have been published in less prestigious journals, or remained unpublished. We have addressed the issue of quality of research in three ways. First, the overall quality of the study was estimated in terms of the quality of the publication medium. For journals, the "impact factor," defined as the average number of citations to the papers in a journal, was considered as a proxy of quality (Garfield, 1979). The impact of conference papers and dissertations was set at zero. The overall quality of the studies was then included in the meta-analyses as a moderator variable. Second, we tested the influence of publication status (published versus unpublished) on the size of the combined effects of the associations between attachment and aspects of parenting. Third, we provided combined effect sizes for associations between attachment and sensitivity separately for the total set of pertinent studies, as well as for the subset of studies using the standard assessment of attachment, that is, the Strange Situation procedure, antecedent or concurrent observational assessment of sensitivity, and nonclinical participants. These studies more closely resembled the original Baltimore study in their assessments of predictor and outcome variables.

A total of 66 studies were identified that together involved 4,176 mother-infant pairs. The meta-analytic database included 54 published articles, as well as 12 unpublished conference presentations or dissertations. A common criticism of meta-analysis is that studies with significant findings are overrepresented in the meta-analytic database because studies with significant findings are more likely to be published than studies with nonsignificant findings. It is assumed that many studies with effect sizes of zero remain unpublished in file drawers. If a meta-analyst relied exclusively on published studies, this could result in an overestimation of population effect sizes. Rosenthal (1979) identified this as the "file drawer" problem. The best solution to the file drawer problem is simply to open the drawers and include as many unpublished studies as possible (Rosenthal, 1991). Unpublished dissertations were located through a manual search of Dissertation Abstracts International (from 1980). We also searched through several conference proceedings (ICIS, SRCD) in which abstracts of presented papers were published.

Another way of dealing with the "file drawer" problem is to estimate the magnitude of the problem by calculating the minimum number of unpublished studies with null results that would be required to turn a significant meta-analytic finding into a nonsignificant one. This number of imaginary unpublished studies is called the "fail safe number" and was introduced by Rosenthal (1979). If this fail safe number is relatively small, a file drawer problem may exist. Calculation of the fail safe number is no definitive

solution to the file drawer problem. It only "establishes reasonable boundaries on the file drawer problem and estimates the degree of damage to the research conclusion that could be done by the file drawer problem" (Rosenthal, 1991, p. 104).

Calculation of Effect Sizes

In the meta-analysis, Pearson's product-moment correlation coefficient (r) was used as the effect size estimate. An effect size indicates the magnitude of the association between two variables, disregarding sample size. If a study reported means and standard deviations, one-directional t values were computed and transformed into r using Schwarzer's (1989) algorithms. If no means and standard deviations were available, the reported test statistics (t, F, or chi-square) or the one-directional p value were transformed into r with Mullen's (1989) computer program.

Because only F tests with 1 degree of freedom in the numerator are appropriate for inclusion in meta-analysis, a contrast F with 1 degree of freedom was computed on the basis of the global F value using the Contrast Analysis procedures (Rosenthal & Rosnow, 1985). Global F tests can be converted to contrast Fs only if group means are available. We applied conservative estimation procedures if a study only reported "no significant effect" ($p = .50$), or "a significant effect" ($p = .05$) in the case of univariate relations (Mullen, 1989).

To compute combined effect sizes, each correlation coefficient was transformed to a Fisher's Z to make the sampling distribution of r more approximate to a Gauss curve. The distribution of r becomes nonlinear at the extreme ends of the scale (Mullen, 1989). Furthermore, in computing the combined effect sizes, individual effect sizes were weighted by sample size, because correlations become more stable as sample size increases and because effect sizes based on large samples deviate less from the population effect size than those based on smaller samples (Mullen, 1989; Rosenthal, 1991). As an indicator of accuracy, we computed the 95% confidence interval around the mean effect size, using Schwarzer's (1989) program.

Combining estimates of effect sizes across studies is reasonable if the studies have a common population effect size. If this is the case, estimates of effect sizes will differ only because of unsystematic sampling error. The crucial question is thus whether the underlying dataset is sufficiently homogeneous. Hedges and Olkin (1985) as well as Rosenthal (1991) advised the use of a test for homogeneity to determine to what extent effect sizes are relatively constant across studies. Regardless of whether this homogeneity test is significant, Johnson, Mullen, and Salas (1995) encouraged the meta-analyst to check for significant moderator variables that may partly account for the variation across studies (see also Rosenthal, 1995). To determine whether a study feature significantly explains a part of the variation in effect sizes, we used Rosenthal's method of focused comparison of combined effect sizes (Mullen, 1989).

Last, blocking was used to reveal moderators in the dataset. Blocking involves grouping study outcomes on the basis of a potential moderator variable. Within each level of the moderator variable, a combined effect size can be computed, and the significance of the differences can be tested (Mullen, 1989).

Multiple Outcomes within the Primary Study

In meta-analysis, the study itself is the unit of analysis. Because the meta-analytic procedures assume independence of units of analysis, we had to deal with the problem of multiple outcomes from single studies. In this meta-analysis, there were three types of multiple outcomes.

First, in several studies the same maternal behavior was assessed at multiple points in time. In these cases we computed a combined effect size over the various measurements; the combined effect size was treated as the outcome of such a study. For example, Isabella (1993) assessed maternal sensitivity at 1, 4, and 9 months of age. The association between maternal sensitivity and the Strange Situation outcome was reported for each age separately. In the meta-analysis, we took the mean outcome across the three times of assessment (see Goldsmith & Alansky, 1987).

Second, several studies examined the relation between a variety of maternal behaviors and attachment security. In the first study, we identified nine categories of predictor variables, each referring to a specific aspect of maternal behavior (see Study 1). After determining the kind(s) of behavior that had been examined in each study, we assigned each behavior to one of our nine categories of maternal behavior. If two maternal behaviors were classified in the same category, we computed the mean effect size. For example, Goldberg, Perotta, Minde, and Corter's (1986) measures of acceptance, delight, perception of the baby, and attitude were all assigned to the category Positive Attitude toward the infant. We therefore computed a mean effect size over these four behaviors.

Third, because many studies assessed different aspects of maternal behavior in their primary analyses, data from a single study were often included in several domains of maternal behavior. For example, the study of Goldberg et al. (1986) provided data for al-

most every domain of maternal behavior. In case of combining domains into the more and less optimal clusters, we computed means over the study outcomes for the domains so that each study had only one outcome in the meta-analysis.

Fourth, some studies were conducted by the same research team and included the same samples (Isabella & Belsky, 1991, and Isabella et al., 1989; Goossens, 1987, and van IJzendoorn, Kranenburg, Zwart-Woudstra, Van Busschbach, & Lambermon, 1991; Main, Tomasini, & Tolan, 1979, and Londerville & Main, 1981; Miyake et al., 1985, and Nakagawa, Lamb, & Miyake, 1992; Bates, Maslin, & Frankel, 1985, and Kiser et al., 1986, and Frankel & Bates, 1990). In these cases we also computed a mean effect size. In other words, studies involving the same sample were represented by only one effect size in the set of meta-analyses.

Coding the Variables

The outcome variable, attachment security, did not need to be coded because it is a sharply defined construct. In the majority of the studies, attachment security was assessed by means of Ainsworth and Wittig's (1969) Strange Situation procedure. In four studies, shortened versions of the Strange Situation procedure were applied (Bohlin, Hagekull, Germer, Andersson, & Lindberg, 1989; Capps et al., 1994; Lewis & Feiring, 1989; Persson-Blennow, Binett, & McNeil, 1988). A total of five studies utilized alternative assessments of attachment, like the Attachment Story Completion Task (Altman, Monk, Jones, & Sosa, 1993; Goodman, Andrews, Jones, Weissman, & Weisman, 1993) and the Attachment Q-Sort (Kerns & Barth, 1995; Pederson et al., 1990; Pederson & Moran, 1996). Recently, van IJzendoorn et al. (in press) showed that the observer Attachment Q-Sort is a reliable and valid measure of attachment security, and they argued that in several respects the Attachment Q-Sort is even preferable to the Strange Situation procedure. In the same meta-analysis, the mother Attachment Q-Sort was shown to be an invalid self-report attachment measure. Studies with the mother version of the Attachment Q-Sort were excluded, therefore, from the current meta-analyses. Furthermore, we focused in the meta-analyses on attachment security (using either the secure-insecure dichotomous variable or security ratings), and we decided not to include the traditional split between insecure attachment classifications because only part of the studies report on the difference between the insecure classifications.

We coded a number of study characteristics. First, some background variables were coded, like publication status (published, unpublished) and year of publication. Second, we coded several characteristics of the sample: sample size, whether the sample was special (i.e., whether the mother or the child suffered from mental or physical handicaps, and whether the family conditions were deviant), socioeconomic status of the sample (middle class, lower class, or heterogeneous), and whether the infants were all firstborn. The most important variables concerned study design. A total of eight design variables were coded: (1) the techniques for measuring maternal behavior (global rating scales, specific behavioral codings based on time or event sampling, or nonobservational techniques like interviews and questionnaires); (2) the exact duration of the observation, in minutes; (3) location where the mother-infant interaction was observed (home or laboratory); (4) the age of the child at the time of the observation; (5) whether the Strange Situation or an alternative measure was used to assess attachment security; (6) the age of the child at the time of assessment of attachment security; (7) time interval in months between the two assessments; and (8) whether the assessments of maternal behavior and infant attachment had been conducted completely independently.

If maternal behavior was assessed at different infant ages, we computed the mean age. For example, in Belsky et al.'s (1984) study, maternal behavior was observed at 1, 3, and 9 months of age. The Strange Situation was assessed at 12.5 months. In this case, we coded the age of the child at the maternal behavior observation (moderator 4) as 4.3 months; the time interval between predictor and outcome assessments (moderator 7) was coded as 8.2 months.

Two persons coded the moderator variables. After a training phase, we performed a reliability check in which each coder independently coded 15 studies. The mean percentage of agreement was 92%. The coding of variables such as publication year, age of the mother, parity, age of the child at predictor, and outcome assessment appeared to be straightforward (for these variables, intercoder agreement was 100%). Most problematic was the coding of the measurement technique used to assess maternal behavior (global versus specific). In this case, intercoder agreement was 67%.

Results

Central Tendency and Variability

Table 1 presents the mean weighted effect sizes for each domain of maternal behavior, as well as confidence intervals, homogeneity tests, and fail safe numbers. The data pertain to all the studies except the

Table 1 Parenting and Attachment: Meta-Analytic Findings for Each Domain of Parental Behavior

Parental Behavior	Studies (*N*)	Participants (*N*)	Effect Size (*r*)	*p*	95% Confidence Interval	Homogeneity χ^2	Fail Safe (*N*)
Sensitivity	30	1,666	.22	9.12 E-15	.18–.27	43.5[a]	861.8
Ainsworth-scale	16	837	.24	1.55 E-09	.17–.30	18.8	238.9
Contiguity of response	14	825	.10	.01	.03–.17	20.4	38.7[b]
Physical contact	9	637	.09	.04	.01–.17	7.8	14.1[b]
Cooperation	9	493	.13	.007	.03–.21	11.1	17.2[b]
More optimal group	28[c]	1,928	.19	7.77 E-08	.14–.23	63.7[a]	735.6
Synchrony	6	258	.26	.0001	.14–.37	11.9[a]	35.5[b]
Mutuality	3	168	.32	.00003	.18–.46	9.3[a]	17.6[b]
Support	22	1,664	.16	7.41 E-06	.11–.21	37.1	328.2
Less optimal group	24[c]	1,233	.19	7.88 E-10	.14–.26	30.5	377.8
Attitude	21	1,092	.18	3.02 E-08	.14–.25	29.4	273.2
Stimulation	9	422	.18	.0001	.10–.29	7.5	32.0[b]
Total set	123[c]	7,225	.17	8.24 E-24	.15–.19	198[a]	9,923.5
Random set	66	4,176	.19	1.83 E-17	.16–.22	139[a]	3,957.3

[a] Data set is heterogeneous.
[b] File drawer problem is indicated.
[c] Total without overlapping studies.

pioneering Baltimore study (Ainsworth et al., 1978) that inspired so many replication studies. Although inclusion of the Baltimore study would not change any of our results drastically, we decided to exclude this study for two reasons. First, it was in the Baltimore study that central measures for sensitivity and attachment were developed and the hypothesis about the association between attachment and sensitivity was generated and specified. Later studies can be considered as replications and extensions. Second, the Baltimore study showed very strong associations between attachment and several aspects of parenting, whereas later studies generally showed less strong relations; in many respects the Baltimore study occupies a somewhat outlying position (see Tables 3 and 4).

As can be derived from Table 1, the combined effect size for the association between maternal sensitivity and attachment was $r(1{,}664) = .22$ ($k = 30$, $N = 1{,}666$). However, when we adopted a more strict definition of the predictor variable and included only studies that measure sensitivity using Ainsworth's original rating scale, the effect size increased to $r(835) = .24$ ($k = 16$, $N = 837$). The correlation between maternal behavior and infant attachment was highest for the set of studies on Mutuality, $r(166) = .32$. However, this meta-analysis involved only three studies ($N = 168$). Weaker effect sizes were found in the set of studies on Contiguity of response, $r(823) = .10$ ($k = 14$, $N = 825$), and Physical Contact, $r(635) = .09$ ($k = 9$, $N = 637$). The combined effect size for Cooperation was $r(491) = .13$ ($k = 9$, $N = 493$). For Stimulation, a similar combined effect size was found: $r(420) = .18$ ($k = 9$, $N = 422$). For Positive Attitude, the effect size was $r(1{,}090) = .18$ ($k = 21$, $N = 1{,}092$), and for Emotional Support this figure was $r(1{,}662) = .16$ ($k = 22$, $N = 1{,}664$). In the set of studies on attachment and Synchrony, the combined effect size appeared to be similar to the outcome in the domain of Sensitivity: $r(256) = .26$ ($k = 6$, $N = 258$). All effect sizes were significant at the alpha = .05 level.

Besides meta-analyses on each domain, we also performed two meta-analyses on a broader conceptual level. Overall effect sizes, however, did not differ between the more optimal and the less optimal cluster: $r(1{,}926) = .19$ ($k = 28$, $N = 1{,}928$) in the more optimal group, and $r(1{,}231) = .19$ ($k = 24$, $N = 1{,}233$) in the less optimal group. On the most global level, we also computed an overall effect size of the studies, irrespective of their cluster membership: mean $r(7{,}223) = .17$ ($k = 123$, $N = 7{,}225$). This global approach increased, of course, the heterogeneity of the set of studies, and it also led to the inclusion of dependent outcomes of the same studies. Therefore, we randomly selected one effect size per study, in case of multiple outcomes. The resulting overall effect size was $r(4{,}174) = .19$ ($k = 66$, $N = 4{,}176$). In this random set of 66 effect sizes, we tested whether the combined effect size of the sensitivity cluster was significantly different from the other clusters. Because eight comparisons were made, the Bonferroni corrected alpha level was used. The Sensitivity cluster appeared to show a significantly larger effect size than the Contiguity of response cluster ($p = .004$).

The fail safe numbers indicated possible file drawer problems in several domains of maternal behavior: Contiguity of response, Physical Contact, Cooperation, Synchrony, Mutuality, and Stimulation. The combined effect sizes for these domains of parenting need to be considered carefully, as they require further corroboration. In the other domains, fail safe numbers exceeded Rosenthal's (1991) critical value ($5 \times k + 10$). These effect sizes can be considered robust. For example, 862 studies with null results in the file drawers of disappointed researchers would be required to make the effect size for the association between sensitivity and attachment nonsignificant.

Table 1 also presents the test statistic chi-square, which indicates whether the effect sizes within the domain were homogeneous. The combined effect size for heterogeneous clusters constitutes problematic estimates of the population effect size. Therefore, we tried to find homogeneous subsets of studies using disjoint cluster analysis (Mullen, 1989; Schwarzer, 1989). For the Sensitivity domain, we found two significantly disjoint clusters (alpha was set at .05). One small and heterogeneous cluster consisted of the study by Capps et al. (1994) on autistic children, and the study by Benn (1986), who used an interview to assess sensitivity. The remaining studies were homogeneous, $\chi^2(27, N = 1{,}621) = 30.6$, $p = .29$, and the combined effect size was $r(1{,}619) = .21$ ($k = 28$). Including studies similar to the Ainsworth Baltimore study in terms of (1) the use of the Strange Situation procedure, (2) the application of observational measures of sensitivity, (3) nonclinical participants, and (4) sensitivity assessments preceding or concurrent with the attachment assessment, we found a combined effect size of $r(1{,}097) = .20$ ($k = 21$), and this set of studies was homogeneous, $\chi^2(20, N = 1{,}099) = 25.2$, $p = .19$.

For the Synchrony domain, disjoint cluster analysis showed two different clusters: The study of Isabella et al. (1989) was set apart from the other studies. Isabella et al. 's (1989) study included an equal number of avoidant ($n = 10$), secure ($n = 10$), and ambivalent ($n = 10$) infants, selected from a larger sample of 51 infants. In particular, the secure group contained only the "most secure" B2 and B3 patterns of attachment (Ainsworth et al., 1978). The remaining Synchrony studies constituted a homogeneous set, $\chi^2(4, N = 228) = 3.4$, $p = .39$, and their combined effect size was $r(226) = .19$ ($k = 5$). For the domain of Mutuality studies, the disjoint cluster analysis discriminated the Smith and Pederson (1988) study from the two other studies. This latter set, however, appeared to remain heterogeneous. In the domain of Support, the disjoint cluster analysis set the Matas et al. (1978) study apart from the other studies which constituted a homogeneous set, $\chi^2(20, N = 1{,}616) = 18.4$, $p = .56$. The 21 studies in this set showed a combined effect size of $r(1{,}614) = .14$.

The more optimal group of 28 studies appeared to become homogeneous after exclusion of the Benn (1986), Isabella et al. (1989), and Matas et al. (1978) studies that were identified through the disjoint cluster analysis. The remaining set showed a combined effect size of $r(1{,}818) = .15$ ($k = 25$); $\chi^2(24, N = 1{,}820) = 22.0$, $p = .57$. In five of the domains of maternal behavior (Positive Attitude, Stimulation, Contiguity of response, Cooperation, and Physical Contact), effect sizes appeared relatively constant across studies, even though some of the meta-analyses included a large number of studies. Effect sizes in the "less optimal group" were also found to be homogeneous. We nevertheless performed moderator analyses in these domains because a nonsignificant homogeneity test "does not mean that there is no interesting variability in the research domain" (Mullen, 1989, p. 102).

Prediction

In Table 2, the moderators and their z values are reported for each subset. Again, the Baltimore study (Ainsworth et al., 1978) was not included because of its exploratory nature.

Publication status and impact factor were not significant moderators in any of the nine domains of analysis. In this respect, then, the data do not point to a file drawer or quality problem. Year of publication also was not significant in any domain. The duration of the maternal behavior assessment, the use of the Strange Situation procedure or alternative attachment measures, the independence of parenting and attachment assessments, and whether the sample consisted exclusively of firstborn infants did not emerge as significant moderators as well. In meta-analyses, sample size often is a significant moderator because smaller samples usually tend to show stronger effect sizes than larger samples (see Amato & Keith, 1991). In our case, sample size was a significant moderator in four domains: Sensitivity ($z = 2.02$), Synchrony ($z = 3.39$), Mutuality ($z = 2.36$), and Positive Attitude ($z = 2.06$). Smaller samples indeed showed stronger effect sizes. The remaining moderators will be discussed separately for each domain.

Sensitivity. In Table 3, a stem and leaf display of the effect sizes in the most important domain, Sensitivity, is presented. It shows the somewhat outlying position of the Baltimore study, as well as the Seifer

Table 2 The Effect of Moderators in Different Parenting Domains (z Values)[a]

	Sensitivity					More Optimal Cluster			Less Optimal Cluster	
Moderator	Total (k = 30)	Ainsworth Scale (k = 16)	Contiguity of Response (k = 14)	Physical Contact (k = 9)	Cooperation (k = 9)	Synchrony (k = 6)	Mutuality (k = 3)	Support (k = 22)	Positive Attitude (k = 21)	Stimulation (k = 9)
Publication	1.89	.91	.97	.97	—[b]	—[b]	1.49	.21	.61	—[b]
Impact factor	.14	.28	.82	.91	.09	1.34	.46	.26	.27	.61
Year	.97	.92	.14	.31	.52	1.72	.17	1.67	.65	.81
Size	2.02*	.28	1.48	1.76	.81	3.39***	2.36*	1.76	2.06*	.26
SES	2.41*	1.81	.41	1.10	.67	1.86	—[b]	.29	.71	.26
Clinical	.84	.17	.66	.13	2.10*	1.86	1.76	1.10	2.78**	.51
Firstborns	.28	.57	1.36	.79	.76	1.87	—[b]	.19	1.12	.11
Global	.87	—[b]	2.34*	.14	2.59**	—[b]	—[b]	1.12	2.09*	.24
Duration	.27	.99	.30	.19	.78	1.23	1.49	1.42	.46	1.17
Home/lab	.20	.65	1.26	1.65	1.63	1.33	1.49	.93	1.70	2.33*
Age	2.56**	.64	2.58**	1.19	.62	.25	2.19*	1.31	.76	1.89
SSP	1.95	1.32	1.44	1.86	.60	—[b]	—[b]	.05	.98	—[b]
Age-SSP	2.31*	1.35	.95	1.02	.68	.67	1.76	.12	.21	.88
Time interval	2.69**	.79	2.97**	1.28	.33	.40	1.76	.80	1.82	2.20*
Independent	.35	.49	—[b]	—[b]	1.14	—[b]	—[b]	.46	—[b]	—[b]

[a] All computations without the Ainsworth et al. (1978) sample.
[b] No variation in the moderator variable.
* $p < .05$; ** $p < .01$; *** $p < .001$.

et al. (1996) study with a negative correlation between sensitivity and attachment.

Socioeconomic status appeared to be a significant moderator for the 30 studies assessing sensitivity as defined by Ainsworth et al. (1978) (z = 2.41; see Table 2). In the 18 middle-class samples, the effect size was $r(886) = .27$ (N = 888), whereas in the eight lower-class samples, this figure was $r(650) = .15$ (N = 652). Samples without clearly specified or homogeneous socioeconomic status (e.g., variability from lower to higher class) were excluded from these analyses.

Table 3 Stem and Leaf Display of Effect Sizes for the Association between Sensitivity, Contiguity, and Attachment (r)

	Leaf	
Stem[a]	Sensitivity[b]	Contiguity
.9		
.8		
.7	8[c]	
.6	5, 8	
.5	1	0
.4	1, 3, 4	8
.3	0, 2, 3, 4	
.2	0, 1, 3, 5, 6, 6, 7	3, 9
.1	2, 4, 4, 5, 6, 7, 8, 9	0, 5
.0	0, 0, 4, 4	0, 0, 1, 5, 6, 7, 8
−.1	2	7

[a] The stem contains the first digit of the correlation coefficient, whereas the leaf contains the second digit.
[b] Underscored figures indicate studies with the Ainsworth et al. (1978) rating scale for sensitivity.
[c] Ainsworth et al. (1978).

The age of the infants at the time of the sensitivity assessment was also a significant moderator (z = 2.56). Contrary to our expectation, samples with older infants (older than 1 year) showed stronger effect sizes, $r(1, 062) = .27$ (k = 11) than samples with younger infants, $r(602) = .20$ (k = 19). Also, the age of the infants at the time of the attachment assessment was a significant moderator (z = 2.31). Samples with older infants showed stronger effect sizes, $r(847) = .25$ (k = 16) than samples with younger infants, $r(817) = .19$ (k = 14). A shorter time interval between the sensitivity and attachment assessments led to stronger effect sizes (z = 2.69) in the total set of sensitivity studies.

Contiguity of response. In Table 3, a stem and leaf display of the effect sizes for the Contiguity studies is presented. Three moderators were important: Whether contiguity of response was assessed globally (z = 2.34), the age of the infant at the contiguity assessment (z = 2.58), and the time interval between the contiguity and attachment assessments (z = 2.97). More global assessments, $r(363) = .08$ (k = 8), showed weaker associations than more specific assessments, $r(458) = .15$ (k = 8). Contrary to the moderator results in the domain of sensitivity, a longer time interval, $r(352) = .16$ (k = 6), led to a stronger effect size than a shorter interval, $r(469) = .06$ (k = 9), and samples with younger infants, $r(501) = .13$ (k = 7), yielded

somewhat stronger associations than samples with older infants, $r(320) = .06$ ($k = 7$).

Cooperation. Two moderators were significant ($z = 2.10$ and $z = 2.59$, respectively): Studies with special samples revealed weaker effect sizes, $r(223) = .03$ ($k = 4$), than studies with normal samples, $r(266) = .20$ ($k = 5$). Global assessments of Cooperation yielded weaker effect sizes, $r(360) = .05$ ($k = 6$), than specific assessments, $r(357) = .32$ ($k = 2$). These results need to be interpreted with caution, however, because only nine studies were included in this dataset. For this small dataset, a stem and leaf display is not presented.

Positive Attitude. Special samples showed significantly ($z = 2.78$) weaker associations between Positive Attitude and attachment, $r(337) = .08$ ($k = 6$), than normal samples, $r(758) = .23$ ($k = 15$). Studies using global attitude assessments, $r(546) = .21$ ($k = 11$), showed significantly stronger effect sizes ($z = 2.09$) than studies with specific assessments, $r(357) = .06$ ($k = 6$).

Stimulation. Contrary to our expectation, studies on attachment and stimulation at home revealed weaker associations, $r(229) = .07$ ($k = 5$), than studies carried out in the laboratory, $r(169) = .31$ ($k = 3$). Longer intervals again led to smaller effect sizes than shorter intervals ($z = 2.20$). This set of studies was rather small.

The moderator analyses did not uncover significant moderators in the domains of Physical Contact and Emotional Support. In the domain of Synchrony, only sample size appeared to be a relevant moderator. The domain of Mutuality consisted of only three studies; the Smith and Pederson (1988) study showed the strongest effect size, and the characteristics of this study determined the outcome of the moderator analysis.

Synchrony, Mutuality, and Emotional Support. The three domains Synchrony, Mutuality, and Emotional Support were considered to be the "more optimal" cluster. In Table 4, the stem and leaf display of the effect sizes in this domain is presented. There were no negative effect sizes, and again, the Ainsworth et al. (1978) study occupied a somewhat outlying position.

We performed a moderator analysis on this cluster because its outcome would be more robust ($k = 28$). Publication status of the study ($z = 2.23$) was a significant moderator. Seven unpublished studies yielded a combined effect size of $r(861) = .11$, whereas the 21 published studies yielded a combined effect size of $r(1{,}063) = .25$.

Positive Attitude and Stimulation. The less optimal group was a combination of the domains Positive Attitude and Stimulation ($k = 24$). In Table 4, the stem and leaf display of the effect sizes in this domain is presented. The Ainsworth et al. (1978) study is somewhat of an outlier. Significant moderators were: clinical status ($z = 2.50$), time interval ($z = 2.38$), and global versus specific assessment of maternal behavior assessment ($z = 2.03$). In the 18 normal samples, the combined effect size was $r(896) = .22$, whereas in the six clinical samples it was $r(333) = .10$. Longer time intervals led to somewhat smaller effect sizes than shorter time intervals ($z = 2.38$). Thirteen studies in which maternal behavior was assessed globally showed a combined $r(681) = .22$, and the six studies in which maternal behavior was assessed in a specific way yielded a combined $r(342) = .06$.

Table 4 Stem and Leaf Display of Effect Sizes for the Association between the Less Optimal Cluster (Positive Attitude, Stimulation), the More Optimal Cluster (Synchrony, Mutuality, Support) and Attachment (*r*)

	Leaf	
Stem[a]	Less Optimal Cluster	More Optimal Cluster
.9		
.8		
.7	5[b]	7[b]
.6	6	2, 6, 6
.5		
.4	5	7
.3	1, 4, 9	1, 1, 5, 8
.2	2, 3, 3, 4, 6, 6, 7	0, 1, 2, 4, 6, 7
.1	1, 2, 3, 4, 6, 8, 8, 9	0, 2, 4, 5, 6, 7, 9
.0	0, 0, 1, 9	0, 0, 0, 0, 5, 9, 9

[a] The stem contains the first digit of the correlation coefficient, whereas the leaf contains the second digit.
[b] Ainsworth et al. (1978).

Special samples. In two domains, Cooperation and Positive Attitude, the clinical status of the samples appeared to be a significant moderator: Normal samples revealed stronger associations between attachment and parenting than did special samples. We hypothesized that relatively "mild" factors such as prematurity and adoption have no moderating effect on study outcome, whereas more severe factors like deafness, autism, and maltreatment lead to a weaker association between maternal behavior and attachment. All of the studies using special or clinical groups were combined ($k = 10$). These studies were divided into groups on the basis of whether the sample involved was "mildly" clinical or more severely clinical. However, this distinction did not prove to be a significant moderator variable ($z = .84$, $p = .20$), although the trend was clearly in favor of the hypothesis. The six mildly clinical samples (four premature

samples and two adoption samples) yielded a combined effect size of $r(265) = .26$, and the four severely clinical samples (maltreatment, cleft palate, deafness, and autism) showed a combined effect size of $r(218) = .16$.

Discussion

Maternal sensitivity, defined as the ability to respond appropriately and promptly to the signals of the infant, indeed appears to be an important condition for the development of attachment security. For the 30 pertinent studies, the combined effect size was $r(1{,}664) = .22$ ($N = 1{,}666$). Including only the 21 studies using the Strange Situation procedure in nonclinical samples, as well as applying observational sensitivity measures preceding or concurrent with the attachment assessment, we found a combined effect size of $r(1{,}097) = .20$ ($N = 1{,}099$). Applying the Hunter and Schmidt (1990) procedures of correcting results for attenuation based on the reliabilities of the measures (mean reliability of the sensitivity measures was .83, and mean reliability of the Strange Situation procedure was .81), we found a "true" population effect size for the association between attachment and sensitivity of $r(1{,}097) = .24$. The 16 studies using the original Ainsworth et al. (1974) sensitivity scale also showed a combined effect size of $r(835) = .24$ ($N = 837$). Cohen (1988, p. 82) has proposed criteria for small, medium, and large effect sizes ($d = .20$, $d = .50$, and $d = .80$, respectively) corresponding to correlations of .10, .24, and .37, respectively. According to these criteria, the size of the association between sensitivity and attachment in the replication studies is medium. After more than 25 years of research, Bowlby's (1969) important question about the role of sensitivity in the development of infant attachment can therefore be answered in the affirmative.

Cohen's (1988) criteria are rather arbitrary, however, even according to the originator (Cohen, 1962). Several authors argue that criteria for the statistical magnitude of effect sizes are not equivalent to criteria for their theoretical or practical importance (Abelson, 1995; Prentice & Miller, 1992; Sechrest & Yeaton, 1982). Rosenthal (1990) criticized simplistic interpretation of the squared correlation coefficient in terms of percentage of explained variation: "From undergraduate days on we have been taught that there is only one proper thing to do when we see a correlation coefficient: We must square it" (Rosenthal, 1990, p. 775). This approach ignores, among other things, the ceiling effect imposed by measurement error. Rosenthal and Rubin's (1982) Binomial Effect Size Display (BESD) is an alternative interpretation of effect sizes. The BESD depicts an effect size (r) in terms of the improvement rate that is attributable to the predictor variable. Applying this approach to the findings of the current meta-analyses, the correlation of $r = .24$ for those studies closely resembling the original Baltimore study represents an improvement in security from 38% to 62%; that is, infants whose mothers respond sensitively to their signals improve their chance of developing a secure relationship from 38% to 62%, whereas infants whose mothers are less sensitive decrease their chance of developing a secure relationship from 62% to 38%. This improvement rate can hardly be considered trivial in a theoretical or practical sense, in particular when we compare this effect size with famous examples from medical research, such as the widely used heart failure reducing drugs Propranolol ($r = .04$) and aspirin ($r = .03$) (Gage, 1996; Rosenthal, 1991).

We therefore cannot agree with Goldsmith and Alansky's (1987) conclusion that there is only a weak association between sensitivity and attachment. In 12 studies, they found an overall effect size of $r = .16$ (without the Ainsworth et al. [1978] study), which is remarkably similar to our overall combined effect size across all study outcomes, $r(7{,}223) = .17$. First, on the basis of the BESD approach, we are inclined to be more impressed with this overall effect size than Goldsmith and Alansky (1987) were. Second, our combined effect size represents somewhat stricter replication studies of the original Baltimore study. Goldsmith and Alansky (1987) derived their combined effect size from 12 studies in which a variety of parenting measures had been examined in relation to attachment security. These disparate measures were combined in a single meta-analysis. For example, behavioral categories such as the frequency of maternal looking or vocalizing were combined with Ainsworth's rating scale for sensitivity. We showed that Contiguity of response is significantly less strongly associated with attachment security than Sensitivity, and that the moderator analysis in the former domain leads to opposite outcomes, in particular with respect to infants' age and time interval between assessments. Hedges (1986) argued that the results of meta-analyses using broad constructs may obscure important differences among narrower constructs subsumed under the broad construct. He recommended including broad constructs in the meta-analysis, but distinguishing narrower constructs in the data and presentation of results.

Third, moderately strong—and even weak—correlations may nevertheless indicate powerful causal mechanisms. The current meta-analyses are based on

correlational studies—like the Baltimore study—and it is therefore impossible to derive causal conclusions from its outcome (Cook et al., 1992; Miller & Pollock, 1994; Stroebe & Diehl, 1991). Although most studies on attachment and sensitivity are predictive in the sense that earlier assessments of sensitivity were correlated with subsequent attachment assessments, a requirement of causality is the absence of a third factor explaining the association between sensitivity and attachment. In their meta-analysis of attachment intervention studies, van IJzendoorn et al. (1995) showed that interventions are effective in enhancing maternal sensitivity, and that, in particular, short-term interventions focusing on maternal sensitivity can enhance infants' attachment security. The combined effect size for the effectiveness of the attachment interventions was $d = .48$, which is a medium effect (Cohen, 1988). In fact, the short-term interventions may be considered to be a minimal manipulation of the predictor that still accounts for significant variance in the criterion. Prentice and Miller (1992) argue that this is a plausible reason for interpreting effect sizes as important even when they are small. Combined with the outcome of the current meta-analyses, this attests to the important, but not exclusive, causal role of sensitivity in the development of infant attachment (Ainsworth et al., 1978; Bowlby, 1969).

GENERAL DISCUSSION

Although our results appear to support the "orthodox" position that maternal sensitivity is an important condition of attachment security, the outcome of the Baltimore study itself cannot be considered to be replicated (Goldsmith & Alansky, 1987; Lamb et al., 1985). In the Baltimore study, an effect size of $r(21) = .78$ was found for the association between sensitivity and attachment, which is rather different from the combined effect size of the replication studies, $r(1,097) = .24$. Correcting the Baltimore correlation for predictor and criterion unreliability (.89 and .95, respectively) yields a corrected $r(21) = .85$. Logically, a correlation of this impressive magnitude seems to indicate close similarity of the constructs and/or the assessments of attachment and sensitivity rather than an association between independent variables. Furthermore, in a small sample, confidence boundaries around estimated correlations are broad, and outlying observations may be rather influential. Even replication studies with similar longitudinal and intensive designs have failed to replicate the exceptionally strong and striking results of the Baltimore study (e.g., Grossmann et al., 1985; Hubbard & van IJzendoorn, 1991). Paradoxically, the strength of the Baltimore results may have inspired many researchers to document the association between sensitivity and attachment, and at the same time its exploratory design and the size of its results prevented them from strictly replicating the original effect size. Without the Baltimore study, the solid scientific fact of a moderately strong causal association between sensitivity and attachment would not have been established.

The current meta-analyses qualify the original Baltimore results in yet another way. Sensitivity cannot be considered to be the exclusive and most important factor in the development of attachment. Several domains of maternal interactive behavior showed effect sizes that were similar to those for the domain of Sensitivity. For example, Mutuality and Synchrony were quite strongly associated with attachment security, $rs(166, 256) = .32$ and .26, respectively, as were Stimulation, Positive Attitude, and Emotional Support. Contrary to our expectation, the combined effect size of the more optimal cluster of studies on aspects of parenting more closely resembling sensitivity, $r(1,926) = .19$, did not differ from the outcome of the less optimal cluster, $r(1,231) = .19$. That is, aspects of parenting only indirectly related to the sensitivity concept appear to play a similar role in the development of attachment. The modest correlations between the various aspects of parenting and sensitivity leave room for unique and additional influences on attachment. In six studies, sensitivity was correlated with six other aspects of parenting: mean $r(358) = .34$ ($N = 360$). The original concept of sensitivity may not capture the only mechanism through which the development of attachment is shaped (van IJzendoorn, 1995), and studies combining the promising measures may provide more insight into the additional explanatory value of these alternative approaches over and above sensitivity. Sensitivity has lost its privileged position as the only important causal factor. A multidimensional approach of parenting antecedents should replace the search for the unique contribution of sensitivity. It should be noted that in the current set of attachment studies the concept of parenting is virtually limited to the general domain of parental warmth and acceptance, rather than parental management and control. It is therefore unclear whether these latter aspects of parenting would also contribute to the development of attachment, in particular, after the first year of life.

Contrary to our expectations, the duration of home observations was not related to the magnitude of the association between sensitivity and attachment. It did not appear to matter whether studies were conducted in the laboratory or in the home, except in the

case of maternal stimulation. The use of the standard Strange Situation procedure did not lead to different effect sizes compared to the use of alternative attachment measures. The intensive naturalistic design of the Baltimore study does not seem to be essential for the strength of the association between parenting and attachment found in the replication studies. It should be kept in mind, however, that only few replications indeed used an exactly similar design. Furthermore, in samples with younger infants, somewhat weaker associations between sensitivity and attachment tended to be found. Contrary to our hypothesis based on the Baltimore study, we suggest post hoc that these results are theoretically congruent with Bowlby's (1973) view of the development of attachment as contextually labile and flexible in the early years. Because the security of attachment is a characteristic of the dyad more than of the infant in the early years, the development of attachment can easily change direction when family life circumstances, childrearing arrangements or maternal sensitivity change (Thompson, in press; Thompson & Lamb, 1983; Thompson, Lamb, & Estes, 1982). Sensitivity may be an important condition of attachment security only when it remains stable across time, which may occur only in a stable social context (Lamb et al., 1985; Sroufe, 1988).

In this respect, it is important to note that the association between maternal behavior and infant attachment is significantly weaker in studies of lower-class or clinical samples. The measures of maternal behavior and attachment security have been developed and validated in nonclinical, middle-class samples, and they may be less valid in lower-class and clinical samples. We also want to suggest that the formation of attachment relationships under complex lower-class or clinical conditions may not be adequately explained in a monocausal and linear way. In their quasi-experimental study of family-based and communal kibbutzim, Sagi and his colleagues (Sagi, van IJzendoorn, Aviezer, Donnell, & Mayseless, 1994) showed that maternal sensitivity may be overridden by an unfavorable childrearing arrangement in which infants have to sleep away from home, and therefore often develop insecure attachments. In a similar vein, it may be expected that the strains and stresses of lower-class life or the problems presented by clinical conditions may overburden potentially sensitive mothers. Davies and Cummings (1994), for example, suggested that unresolved marital conflicts may have a profound negative effect on the children's emotional security even if the bond with the mother is balanced and her interaction style toward the child is sensitive. They proposed to study the development of emotional (in-)security from a family-wide perspective (Cummings & Davies, 1996). The transactions between social context or clinical conditions, on the one hand, and attachment on the other need more careful study to determine the role of sensitivity, and other aspects of parenting and family life in the development of attachment security more precisely (Belsky & Cassidy, 1994; Egeland & Erickson, 1993; Sameroff & Chandler, 1975; Thompson, in press). In attachment theory, a move to the level of context may be necessary, so that the interaction between maternal sensitivity and the accumulation of stresses and risk factors in lower-class or clinical groups can be taken into account.

Even in "normal," nonclinical groups, sensitivity plays an important but not exclusive role in the emergence of attachment security. In cognitive development, genetics may constitute a ceiling effect for the influence of environmental factors, for which less than 50% of the variation in individual differences appears to be left (Plomin, 1994). Attachment security, however, does not seem to be genetically determined in any comparable way. The first, small-scale twin studies on attachment reported only 30%–50% congruence of attachment classifications (Minde, Corter, Goldberg, & Jeffers, 1990; Szajnberg, Skrinjaric, & Moore, 1989; Vandell, Owen, Wilson, & Henderson, 1988). A secondary analysis of the available twin data ($N = 56$) did not support the idea of a genetic basis for individual differences in attachment security (Ricciuti, 1993). Only the variable representing the split between the Strange Situation subclassifications A1-B2 and B3-C2 showed significant genetic influence. This may be due to temperament (e.g., emotional reactivity), which has been shown to be associated with this split (Belsky & Rovine, 1987). Although more research in this area is needed, we suggest that attachment security is especially liable to nongenetic, environmental influences. Nevertheless, behavior genetics may inspire a move toward the contextual level in a specific sense. One of the most intriguing findings of behavior genetics is the crucial role of nonshared environment for child development. Unrelated adoptive siblings appear to develop quite differently, even though they are raised in the same, shared family environment. For example, the correlation between IQ of adoptive siblings is approaching zero when they grow older (Plomin, 1994). In the case of parenting and attachment, the concept of nonshared environmental influences may, for example, lead to more emphasis on the family system and on life events. Although parents may interact equally sensitively with both siblings at the same age, the older sibling also experiences the parents inter-

acting sensitively with the younger sibling—which is a unique and potentially powerful experience that never has been studied thoroughly (Dunn & Plomin, 1990). Life events such as parental loss of attachment figures may affect siblings in different ways, depending on their age. If parents suffered bereavement within 2 years after the birth of a sibling, this sibling—but not the other siblings—may develop dissociative tendencies that may be related to insecure-disorganized attachment (Hesse & van IJzendoorn, 1996; Liotti, 1992). After more than 25 years of research on the important dimension of sensitivity, a move to the level of (nonshared) context should inspire the next wave of studies on the antecedents of attachment security.

ACKNOWLEDGMENTS

Support for this study was provided by a PIONEER award from the Netherlands' Organization for Scientific Research (NWO: PGS 59-256) and by a fellowship from the Netherlands Institute for Advanced Study in the Humanities and Social Sciences (NIAS) to Marinus van IJzendoorn. The authors gratefully acknowledge the thoughtful comments of members of the editorial board and reviewers on an earlier draft of this paper.

ADDRESSES AND AFFILIATIONS

Corresponding authors: Marianne S. De Wolff and Marinus H. van IJzendoorn, Center for Child and Family Studies, Leiden University, P.O. Box 9555, NL-2300RB Leiden, The Netherlands; e-mail: vanijzen@rulfsw.leidenuniv.nl.

REFERENCES[1]

Abelson, R. P. (1995). *Statistics as principled argument.* Hillsdale, NJ: Erlbaum.

Ainsworth, M. D. S., Bell, S. M., & Stayton, D. J. (1974). Infant-mother attachment and social development: "Socialization" as a product of reciprocal responsiveness to signals. In M. P. M. Richards (Ed.), *The integration of a child into a social world* (pp. 99–135). Cambridge: Cambridge University Press.

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

Ainsworth, M. D. S., & Wittig, B. A. (1969). Attachment and exploratory behavior of one-year-olds in a Strange Situation. In B. M. Foss (Ed.), *Determinants of infant behaviour* (Vol. 4, pp. 11–136). London: Methuen.

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis.* Beverly Hills, CA: Sage.

Altman, S., Monk, C., Jones, P., & Sosa, L. (1993, March). *Children's working models of attachment at 42 months and maternal attitudes and behavior.* Poster presented at the meetings of the Society for Research in Child Development, New Orleans, LA. [7].*

Amato, P. R., & Keith, B. (1991). Parental divorce and the well-being of children: A meta-analysis. *Psychological Bulletin, 110,* 26–46.

Anisfeld, E., Casper, V., Nozyce, M., & Cunningham, N. (1990). Does infant carrying promote attachment? An experimental study of the effects of the increased physical contact on the development of attachment. *Child Development, 61,* 1617–1627. [1, 2].*

Antonucci, T. C., & Levitt, M. J. (1984). Early prediction of attachment security: A multivariate approach. *Infant Behavior and Development, 7,* 1–18. [3, 5].*

Bates, J. E., Maslin, C. A., & Frankel, K. A. (1985). Attachment security, mother-child interaction, and temperament as predictors of behavior-problem ratings at the age of three years. In I. Bretherton & E. Waters (Eds.), Growing points in attachment theory and research (pp. 167–193). *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209). [2, 7, 8].*

Belsky, J., & Cassidy, J. (1994). Attachment: Theory and evidence. In M. Rutter, D. Hay, & S. Baron-Cohen (Eds.), *Developmental principles and clinical issues in psychology and psychiatry* (pp. 373–402). Oxford: Blackwell.

Belsky, J., & Rovine, M. (1987). Temperament and attachment security in the Strange Situation: An empirical rapprochement. *Child Development, 58,* 787–795.

Belsky, J., Rovine, M., & Taylor, D. G. (1984). The Pennsylvania Infant and Family Development Project, III: The origins of individual differences in infant-mother attachment: Maternal and infant contributions. *Child Development, 55,* 718–728. [2, 5, 7, 9].*

Benn, R. K. (1986). Factors promoting secure attachment relationships between employed mothers and their sons. *Child Development, 57,* 1224–1231. [1, 8].*

Bohlin, G., Hagekull, B., Germer, M., Andersson, K., & Lindberg, L. (1989). Avoidant and resistant behaviors as predicted by maternal interactive behavior and infant temperament. *Infant Behavior and Development, 12,* 105–117. [1, 2, 3, 4].*

Booth, C. L., Rose-Krasnor, L., & Rubin, K. H. (1991). Relating preschoolers' social competence and their mothers' parenting behaviors to early attachment security and high-risk status. *Journal of Social and Personal Relationships, 8,* 363–382. [7].*

Bornstein, M. H. (Ed.). (1989). Maternal responsiveness: Characteristics and consequences. *New Directions for Child Development.* San Francisco: Jossey-Bass.

1. References marked with an asterisk indicate studies included in the meta-analyses. We also indicated after these references the numbers of the meta-analytic domains for which the particular study has provided data. The numbers refer to the following meta-analytic domains: (1) Sensitivity, (2) Responsiveness, (3) Physical Contact, (4) Cooperation, (5) Synchrony, (6) Mutuality, (7) Support, (8) Attitude, and (9) Stimulation.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment* (rev. ed.). Harmondsworth: Penguin.

Bowlby, J. (1973). *Attachment and loss: Vol. 2. Separation: Anxiety and anger.* Harmondsworth: Penguin.

Bretherton, I. (1985). Attachment theory: Retrospect and prospect. In I. Bretherton & E. Waters (Eds.), Growing points in attachment theory and research (pp. 3–38). *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209).

Bretherton, I., Biringen, Z., Ridgeway, D., Maslin, C., & Sherman, M. (1989). Attachment: The parental perspective. *Infant Mental Health Journal, 10,* 203–221. [1].*

Burchinal, M. R., Bryant, D. M., Lee, M. W., & Ramey, C. T. (1992). Early day-care, infant-mother attachment, and maternal responsiveness in the infant's first year. *Early Childhood Research Quarterly, 7,* 383–396. [7, 8].*

Butcher, P. R., Kalverboer, A. F., Minderaa, R. B., Van Doormaal, E. F., & Ten Wolde, Y. (1993). Rigidity, sensitivity, and quality of attachment: The role of maternal rigidity in the early socio-emotional development of premature infants. *Acta Psychiatrica Scandinavica, 88*(Suppl. 374), 4–38. [1, 8].*

Caldwell, B. M., & Bradley, R. H. (1984). *Administration manual. Home Observation for Measurement of the Environment.* Little Rock: Department of Psychology, University of Arkansas.

Camfield, E., Brownell, C., Taylor, P., Day, N., Brown, E., & Kratzer. L. (1991, April). *Mediators of relations between maternal behavior at 4 months and infant attachment at 12 months.* Poster presented at the meetings of the Society for Research in Child Development, Seattle. [2, 7].*

Capps, L., Sigman, M., & Mundy, P. (1994). Attachment security in children with autism. *Development and Psychopathology, 6,* 249–261. [1].*

Cohen, J. (1962). The statistical power of abnormal social psychology research: A review. *Journal of Abnormal and Social Psychology, 65,* 145–153.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Cook, T. D., Cooper, H., Cordray, D. F., Hartman, H., Hedges, L. V., Louis, T. A., & Mosteller, F. (1992). *Meta-analysis for explanation: A casebook.* New York: Russell Sage.

Cooper, H., & Hedges, L. V. (Eds.). (1994). *The handbook for research synthesis.* New York: Russell Sage.

Cox, M. J., Owen, M. T., Henderson, V. K., & Margand, N. A. (1992). Prediction of infant-father and infant-mother attachment. *Developmental Psychology, 28,* 474–483. [3, 7, 8].*

Crockenberg, S. B. (1981). Infant irritability, mother responsiveness, and social support influences on the security of infant-mother attachment. *Child Development, 52,* 857–865. [2].*

Cummings, E. M., & Davies, P. T. (1996). Emotional security as a regulatory process in normal development and the development of psychopathology. *Development and Psychopathology, 7,* 123–139.

Davies, P. T., & Cummings, E. M. (1994). Marital conflict and child adjustment: An emotional security hypothesis. *Psychological Bulletin, 116,* 387–411.

Del Carmen, R., Pedersen, F., Huffman, L. C., & Bryan, Y. E. (1993). Dyadic distress management predicts subsequent security of attachment. *Infant Behavior and Development, 16,* 131–147. [2, 8, 9].*

Dunn, J., & Plomin, R. (1990). *Separate lives: Why siblings are so different.* New York: Basic.

Easterbrooks, M. A., & Goldberg, W. A. (1985). Effects of early maternal employment on toddlers, mothers, and fathers. *Developmental Psychology, 21,* 774–783.

Egeland, B., & Erickson, M. F. (1993). Attachment theory and findings: Implications for prevention and intervention. In S. Kramer & H. Parens (Eds.), *Prevention in mental health: Now, tomorrow, ever?* (pp. 21–50). Northvale, NJ: Jason Aronson.

Egeland, B., & Farber, E. A. (1984). Infant-mother attachment: Factors related to its development and changes over time. *Child Development, 55,* 753–771. [1].*

Erickson, M. F., Sroufe, L. A., & Egeland, B. (1985). The relationship between quality of attachment and behavior problems in preschool in a high-risk sample. In I. Bretherton & E. Waters (Eds.), Growing points in attachment theory and research (pp. 147–166). *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209).

Everitt, B. (1974). *Cluster analysis.* London: Heineman Educational Books.

Fagot, B. I., & Kavanagh, K. (1990). The prediction of antisocial behavior from avoidant attachment classifications. *Child Development, 61,* 864–973. [7, 8]*

Fracasso, M. P., Busch-Rossnagel, N. A., & Fisher, C. B. (1994). The relationship of maternal behavior and acculturation to the quality of attachment in Hispanic infants living in New York city. *Hispanic Journal of Behavioral Sciences, 16,* 143–154. [1].*

Frankel, A., & Bates, J. E. (1990). Mother-toddler problem solving: Antecedents in attachment, home behavior, and temperament. *Child Development, 61,* 810–819. [7].*

Frodi, A., Grolnick, W., & Bridges, L. (1985). Maternal correlates of stability and change in infant-mother attachment. *Infant Mental Health Journal, 6,* 60–67. [1].*

Gage, N. L. (1996). Confronting counsels of despair for the behavioral sciences. *Educational Researcher, 25,* 5–15

Garfield, E. (1979). *Citation indexing: Its theory and application in science, technology, and humanities.* New York: Wiley.

Gewirtz, J. L., & Boyd, E. F. (1977). Does maternal responding imply reduced infant crying? A critique of the 1972 Bell and Ainsworth report. *Child Development, 48,* 1200–1207.

Gifi, A. (1990). *Non-linear multivariate analysis.* New York: Wiley.

Goldberg, S., Perotta, M., Minde, K., & Corter, C. (1986). Maternal behavior and attachment in low-birth-weight twins and singletons. *Child Development, 57,* 34–46. [1, 2, 3, 4, 7, 8, 9].*

Goldsmith, H. H., & Alansky, J. A. (1987). Maternal and infant predictors of attachment: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 55,* 805–816.

Goodman, G., Andrews, T., Jones, S., Weissman, J., & Weissman, A. (1993). *Maternal play behavior and children's representations as mediators of developmental outcomes among high-risk dyads.* Manuscript in preparation. [1].*

Goossens, F. A. (1987). Kwaliteit van de hechtingsrelatie van jonge kinderen aan hun moeder, vader en crèche-leidster. [Quality of infants' attachment to father, mother, and caregiver in daycare]. *Nederlands Tijdschrift voor de Psychologie, 42,* 308–320. [1].*

Grossmann, K., Grossmann, K. E., Spangler, G., Suess, G., & Unzner, L. (1985). Maternal sensitivity and newborns' orientation responses as related to quality of attachment in northern Germany. In I. Bretherton & E. Waters (Eds.), Growing points in attachment theory and research (pp. 233–268). *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209). [1].*

Gunnar, M. R., Brodersen, L., Nachmias, M., Buss, K., & Rigatuso, J. (1996). Stress reactivity and attachment security. *Developmental Psychobiology, 29,* 191–204. [1].*

Hedges, L. V. (1986). Issues in meta-analysis. *Review of Research in Education, 13,* 353–399.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hesse, E., & van IJzendoorn, M. H. (1996). *Parental loss of close family members within two years preceding or following participants' birth is related to absorption.* Unpublished manuscript, Center for Child and Family Studies, Leiden University.

Hoeksma, J. B., & Koomen, M. M. Y. (1991). *Development of early mother-child interactions and attachment.* Unpublished doctoral dissertation, Free University of Amsterdam, The Netherlands. [1, 6].*

Hubbard, F. O. A., & van IJzendoorn, M. H. (1991). Maternal unresponsiveness and infant crying across the first nine months: A naturalistic longitudinal study. *Infant Behavior and Development, 14,* 299–312.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Beverly Hills, CA: Sage.

Isabella, R. A. (1993). Origins of attachment: Maternal interactive behavior across the first year. *Child Development, 64,* 605–621. [1, 8, 9].*

Isabella, R. A., & Belsky, J. (1991). Interactional synchrony and the origins of infant-mother attachment: A replication study. *Child Development, 62,* 373–384. [5].*

Isabella, R. A., Belsky, J., & Von Eye, A. (1989). Origins of infant-mother attachment: an examination of interactional synchrony during the infant's first year. *Developmental Psychology, 25,* 12–21. [5]*

Izard, C. E., Heynes, O. M., Chisholm, G., & Baak, K. (1991). Emotional determinants of infant-mother attachment. *Child Development, 62,* 906–917. [8].*

Johnson, B. T., Mullen, B., & Salas, E. (1995). Comparison of three meta-analytic approaches. *Journal of Applied Psychology, 80,* 94–106.

Juffer, F. (1993). *Verbonden door adoptie: Een experimenteel onderzoek bij gezinnen met een adoptiebaby.* [Attached through adoption: An experimental study on attachment and competence in families with an adopted baby]. Amersfoort, The Netherlands: Academische Uitgeverij. [1, 4].*

Kerns, K. A., & Barth, J. M. (1995). Attachment and play: Convergence across components of parent-child relationships and their relations to peer competence. *Journal of Social and Personal Relationships, 12,* 243–260 [3].*

Kiser, L. J., Bates, J. E., Maslin, C. A., & Bayles, K. (1986). Mother-infant play at six months as a predictor of attachment security of thirteen months. *Journal of the American Academy of Child Psychiatry, 25,* 68–75. [6].*

Krentz, M. S. (1982). *Qualitative differences between mother-child and caregiver-child attachment of infant in family day care.* Unpublished doctoral dissertation, California School of Professional Psychology. [2, 7].*

Kveton, E. M. (1989). The quality of infant-parent attachment at 17 months as related to the quality of parent-infant interaction and prenatal parenting attitudes (Doctoral dissertation, Virginia Consortium for Professional Psychology, 1989). *Dissertation Abstracts International, 50,* 5884B. [1].*

Lamb, M. E., Hopps, K., & Elster, A. B. (1987). Strange Situation behavior of infants with adolescent mothers. *Infant Behavior and Development, 10,* 39–48. [2, 5, 8, 9].*

Lamb, M. E., Thompson, R. A., Gardner, W., & Charnov, E. L. (1985). *Infant-mother attachment: The origins and developmental significance of individual differences in Strange Situation behavior.* Hillsdale, NJ: Erlbaum.

Lambermon, M. W. E. (1991). *Video of folder? Korte- en langetermijn effecten van voorlichting over vroegkinderlijke opvoeding.* [Video or booklet? Short-term and long-term effects of information about early childhood education]. Unpublished doctoral dissertation, University of Leiden, The Netherlands. [1, 7].*

Lederberg, A. R., & Mobley, C. E. (1990). The effect of hearing impairment on the quality of attachment and mother-toddler interaction. *Child Development, 61,* 1596–1604. [4, 7, 8, 9].*

Lewis, M., & Feiring, C. (1989). Infant, mother, and mother-infant interaction behavior and subsequent attachment. *Child Development, 60,* 831–837. [2, 3].*

Lieberman, A. F., Weston, D. R., & Pawl, J. H. (1991). Preventive intervention and outcome with anxiously attached dyads. *Child Development, 62,* 199–209. [7, 9].*

Liotti, G. (1992). Disorganized/disoriented attachment in the etiology of the dissociative disorders. *Dissociation, 5,* 196–204.

Londerville, S., & Main, M. (1981). Security of attachment, compliance, and maternal training methods in the second year of life. *Developmental Psychology, 17,* 289–299. [4].*

Lyons-Ruth, K., Connell, D. B., Zoll, D., & Stahl, J. (1987). Infants at social risk: Relations among infant maltreatment, maternal behavior, and infant attachment behavior. *Developmental Psychology, 23,* 223–232. [1, 3, 4, 8].*

Main, M. (1990). Cross-cultural studies of attachment organization: Recent studies, changing methodologies, and the concept of conditional strategies. *Human Development, 33,* 48–61.

Main, M., Tomasini, L., & Tolan, W. (1979). Differences among mothers of infants judged to differ in security. *Developmental Psychology, 15,* 472–473. [1, 3, 8].*

Malatesta, C. Z., Culver, C., Tesman, C., & Shepard, B. (1989). The development of emotion expression during the first two years of life. *Monographs for the Society for Research in Child Development, 54*(1–2, Serial No. 219). [8].*

Mangelsdorf, S., Gunnar, M., Kestenbaum, R., Lang, S., & Andreas, D. (1990). Infant proneness to distress temperament, maternal personality, and mother-infant attachment: Associations and goodness of fit. *Child Development, 61,* 820–831. [7].*

Matas, L., Arend, R. A., & Sroufe, L. A. (1978). The relationship between quality of attachment and later competence. *Child Development, 49,* 547–556. [7].*

Meij, J. (1991). *Sociale ondersteuning, gehechtheidskwaliteit en vroegkinderlijke competentie-ontwikkeling.* [Social support, quality of attachment, and early development of competence in infants]. Unpublished doctoral dissertation, University of Nijmegen, The Netherlands. [1].*

Miller, N., & Pollock, V. E. (1994). Meta-analytic synthesis for theory development. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 457–483). New York: Russell Sage.

Minde, K., Corter, C., Goldberg, S., & Jeffers, D. (1990). Maternal preference between premature twins up to age four. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 367–374.

Miyake, K., Chen, S. J., & Campos, J. J. (1985). Infant temperament, mother's mode of interaction, and attachment in Japan: An interim report. In I. Bretherton & E. Waters (Eds.), Growing points in attachment theory and research (pp. 276–297). *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209). [2, 4, 7, 9].*

Mullen, B. (1989). *Advanced BASIC meta-analysis.* Hillsdale, NJ: Erlbaum.

Nakagawa, M., Lamb, M. E., & Miyake, K. (1992). Antecedents and correlates of the Strange Situation behavior of Japanese infants. *Journal of Cross-Cultural Psychology, 23,* 300–310. [1, 4, 7, 8].*

NICHD Early Child Care Network. (1996, April). *Infant child care and attachment security: Results of the NICHD study of early child care.* Symposium presented at the International Conference on Infant Studies, Providence, RI. [7].*

Pederson, D. R., & Moran, G. (1996). Expressions of the attachment relationship outside of the Strange Situation. *Child Development, 67,* 915–967. [1].*

Pederson, D. R., Moran, G., Sitko, C., Campbell, K., Ghesquire, K., & Acton, H. (1990). Maternal sensitivity and the security of infant-mother attachment: A Q-sort study. *Child Development, 61,* 1974–1983. [1].*

Persson-Blennow, I., Binett, B., & McNeil, T. F. (1988). Offspring of women with nonorganic psychosis: Antecedents of anxious attachment to the mother at one year of age. *Acta Psychiatrica Scandinavica, 78,* 66–71. [2, 3, 8].*

Plomin, R. (1994). The Emanuel Miller Memorial Lecture 1993: Genetic research and identification of environmental influences. *Journal of Child Psychology and Psychiatry, 35,* 817–834.

Prentice, D. A., & Miller, D. A. (1992). When small effects are impressive. *Psychological Bulletin, 112,* 160–164.

Ricciuti, A. E. (1993). Child-mother attachment: A twin study. *Dissertation Abstracts International, 54,* 3364.

Roggman, L. A., Langlois, J. H., & Hubbs-Tait, L. (1987). Mothers, infants, and toys: Social play correlates of attachment. *Infant Behavior and Development, 10,* 233–237. [5, 9].*

Rosenboom, L. G. (1994). *Gemengde gezinnen, gemengde gevoelens? Hechting en competentie van adoptiebaby's in gezinnen met biologisch eigen kinderen.* [Mixed families, mixed feelings? Attachment and competence of adopted babies in families with biological children]. Unpublished doctoral dissertation, University of Utrecht, The Netherlands. [1, 7].*

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Rosenthal, R. (1990). How are we doing in soft psychology? *American Psychologist, 45,* 775–777.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183–192.

Rosenthal, R., & Rosnow, R. L. (1985). *Contrast analysis: Focused comparisons in the analysis of variance.* New York: Cambridge University Press.

Rosenthal, R., & Rubin, D. B. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin, 92,* 500–504.

Sagi, A., van IJzendoorn, M. H., Aviezer, O., Donnell, F., & Mayseless, O. (1994). Sleeping out of the home in a kibbutz communal arrangement: It makes a difference for infant-mother attachment. *Child Development, 65,* 992–1004.

Sameroff, A. J., & Chandler, M. J. (1975). Reproductive risk and the continuum of caretaking casualty. In F. D. Horowitz (Ed.), *Review of child development research* (Vol. 4, pp. 187–244). Chicago: University of Chicago Press.

Schneider Rosen, K., & Rothbaum, F. (1993). Quality of parental caregiving and security of attachment. *Developmental Psychology, 29,* 358–367. [7, 8].*

Schwarzer, R. (1989). *Meta-analysis programs.* Unpublished manual, University of Berlin, Department of Psychology.

Sechrest, L., & Yeaton, W. H. (1982). Magnitudes of experimental effects in social science research. *Evaluation Review, 6,* 579–600.

Seifer, R., Schiller, M., Sameroff, A. J., Resnick, S., & Riordan, K. (1996). Attachment, maternal sensitivity, and infant temperament during the first year of life. *Developmental Psychology, 32,* 12–25. [1].*

Shaw, D. S., & Vondra, J. I. (1995). Infant security and maternal predictors of early behavior problems: A longitudinal study of low-income families. *Journal of Abnormal Child Psychology, 23,* 335–357. [2].*

Smith, P. B., & Pederson, D. R. (1988). Maternal sensitivity

and patterns of infant-mother attachment. *Child Development, 59,* 1097–1101. [4, 6].*

SPSS, Inc. (1990). *SPSS categories.* Chicago: SPSS, Inc.

Sroufe, L. A. (1988). The role of infant-caregiver attachment in development. In J. Belsky & T. Nezworski (Eds.), *Clinical implications of attachment* (pp. 18–40). Hillsdale, NJ: Erlbaum.

Stifter, C. A., Couleham, C. M., & Fish, M. (1993). Linking employment to attachment: The mediating effects of maternal separation anxiety and interactive behavior. *Child Development, 64,* 1451–1560. [1, 4].*

Stroebe, W., & Diehl, M. (1991). You can't beat good experiments with correlational evidence: Mullen, Johnson, and Salas' meta-analytic misinterpretations. *Basic and Applied Social Psychology, 12,* 25–32.

Szajnberg, N. M., Skrinjaric, J., & Moore, A. (1989). Affect attunement, attachment, temperament, and zygosity: A twin study. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 249–253.

Teti, D. M., Nakagawa, M., Das, R., & Wirth, O. (1991). Security of attachment between preschoolers and their mothers: Relations among social interaction, parenting stress, and mothers' sorts of attachment Q-set. *Developmental Psychology, 27,* 440–447. [7].*

Thompson, R. A. (in press). Early socio-personality development. In W. Damon (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development.* New York: Wiley.

Thompson, R. A., & Lamb, M. E. (1983). Security of attachment and stranger sociability in infancy. *Developmental Psychology, 19,* 184–191.

Thompson, R. A., Lamb, M. E., & Estes, D. (1982). Stability of infant-mother attachment and its relation to changing life circumstances in an unselected middle-class sample. *Child Development, 53,* 144–148.

Ungerer, J. A., Dolby, R., Waters, B., Barnett, B., Kelk, N., Lewin, V., & Blaszczynski, A. (1990). The early development of empathy: Self-regulation and individual differences in the first year. *Motivation and Emotion, 14,* 93–106. [8].*

Vandell, D. L., Owen, M. T., Wilson, K. S., & Henderson, V. K. (1988). Social development in infant twins: Peer and mother-child relationships. *Child Development,* 59, 168–177.

van den Boom, D. C. (1988). *Neonatal irritability and the development of attachment: Observation and intervention.* Unpublished doctoral dissertation, University of Leiden, The Netherlands. [1].*

Van der Kloot, W. A., & Van Herk, H. (1991). Multidimensional scaling of sorting data: A comparison of three procedures. *Multivariate Behavioral Research, 26,* 563–581.

van IJzendoorn, M. H. (1990). Attachment in Surinam-Dutch families: A contribution to the cross-cultural study of attachment. *International Journal of Behavioral Development, 13,* 333–344. [1].*

van IJzendoorn, M. H. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin, 117,* 387–403.

van IJzendoorn, M. H., Juffer, F., & Duyvesteyn, M. G. C. (1995). Breaking the intergenerational cycle of insecure attachment: A review of the effects of attachment-based interventions on maternal sensitivity and infant security. *Journal of Child Psychology and Psychiatry, 36,* 225–248.

van IJzendoorn, M. H., Kranenburg, M. J., Zwart-Woudstra, H. A., Van Busschbach, A. M., & Lambermon, M. W. E. (1991). Parental attachment and children's socio-emotional development: Some findings on the validity of the Adult Attachment Interview in the Netherlands. *International Journal of Behavioral Development, 14,* 375–394. [1].*

van IJzendoorn, M. H., & Kroonenberg, P. M. (1988). Cross-cultural patterns of attachment. A meta-analysis of the Strange Situation. *Child Development, 59,* 147–156.

van IJzendoorn, M. H., Vereijken, C. M. J. L., & Riksen-Walraven, J. M. A. (in press). Is the Attachment Q-sort a valid measure of attachment security in young children? In B. E. Vaughn & E. Waters (Eds.), *Patterns of secure base behavior: Q-sort perspectives on attachment and caregiving.* Hillsdale, NJ: Erlbaum.

Vaughn, B. E., & Waters, E. (1990). Attachment behavior at home and in the laboratory: Q-sort observations and Strange Situation classifications of one-year-olds. *Child Development, 61,* 1965–1973.

Verkes, R. J., Van der Kloot, W. A., & Van der Meij, J. (1989). The perceived structure of 176 pain descriptive words. *Pain, 38,* 219–229.

Waters, E., & Deane, K. E. (1985). Defining and assessing individual differences in attachment relationships: Q-methodology and the organization of behavior in infancy and early childhood. In I. Bretherton & E. Waters (Eds.), Growing points in attachment theory and research (pp. 39–40). *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209).

Wille, D. E. (1991). Relation of preterm birth with quality of infant-mother attachment at one year. *Infant Behavior and Development, 14,* 227–240. [6, 7, 8].*

Zaslow, M. J., Rabinovich, B. A., Suwalsky, J. T. D., & Klein, R. P. (1988). The role of social context in the prediction of secure and insecure/avoidant infant-mother attachment. *Journal of Applied Developmental Psychology, 9,* 287–299. [3, 7, 8].*