CHAPTER 12

Does Maternal Responsiveness Increase Infant Crying?*

Replication of the Baltimore Study

Frans O. A. Hubbard and Marinus H. Van IJzendoorn
Center for Child and Family Studies
Leiden University
The Netherlands

Viewed from the perspective of attachment theory, crying is attributed the function of maintaining the baby's proximity to protective caregivers. Bowlby (1971) assumed that crying belonged to the class of preattachment behaviors, such as sucking and smiling, that served the human species in its evolutionary struggle for life. In this cybernetic model of attachment theory it is plausible to consider that crying is decreased by a prompt response from the caregiver, rather than being reinforced (Ainsworth & Bell, 1977; Bell & Ainsworth, 1972). If crying is considered as directed toward the goal of maintaining proximity to the protective caregiver, reaching that goal should be sufficient "reason" to terminate the crying behavior (Ainsworth, Blehar, Waters, & Wall, 1978). Infants are seen as gradually deriving a feeling of competence and mastery over the environment as a result of their perception of a prompt reaction to their crying. If infants really experience danger they anticipate short latencies in the caregiver's response. In mild distress they do not feel urged to alarm the environment instantly by crying (Bell & Ainsworth, 1972); therefore, during the last quarters of the first year of life crying

*Parts of this chapter were presented at the International Conference on Infant Studies, Washington, DC, April 1988, and at the Third International Workshop on Infant Cry Research, July 13–18, 1989, Hanasaari, Espoo, Finland. The research was supported by a grant from the National Foundation for Preventive Studies (Praeventie Fonds, Grant No. 28-859) to Marinus H. van IJzendoorn and Louis W. C. Tavecchio. We would like to thank Louis Tavecchio for his kind support and Agnes van Busschbach, Annet Foortse, Karin Haak, Gidia Jacobs, Ester Kloppers, Johan van Mill, Ina van den Oever, Letty van Oyen, Hansje Planjer, Inge Ruhaak, Marianne Thomeer, Nico Vesseur, Yvonne ten Wolde, André Wortelboer, Jeroen Zomerplaag, for their assistance in collecting, coding, and analyzing data.

behavior would decrease because of the responsive reactions of the caregiver.

The Original Study

Bell and Ainsworth's (1972) ecologically impressive and unique longitudinal study was intended to explore the idea that crying behavior was reduced by prompt reactions from the caregiver. Twenty-six middle-class mother–infant dyads were visited at home for about four hours every three weeks, during the first year of the baby's life. Although observing the mother–infant interactions, observers took notes which were later written out as a narrative account. In the narrative accounts, frequency and duration of all cry events during the home visit were coded, as were the maternal interventions. The frequency and duration measures for crying behavior contained fussing as well as long and intense crying. All measures were aggregated over three monthly intervals. Bell and Ainsworth (1972) concluded that the main results confirmed the hypothesized relation between maternal unresponsiveness and infant crying. First, crying behavior in the first half-year appeared much less stable than maternal unresponsiveness. This would exclude a purely constitutional explanation (e.g., irritability) for the differences in crying behavior at the end of the first year. Second, within-quarter correlations between crying and unresponsiveness indicated a tendency to more frequent and longer periods of crying by infants whose mothers were less responsive. Third, across-quarter correlations indicated that less responsive reactions to crying in the first quarter corresponded with more crying in a later quarter.

Controversies about the Original Study

Although the study seemed to have produced clear-cut results, it has also been criticized sharply on technical grounds, especially by Gewirtz and Boyd (1977a, 1977b). They stated, for example, that in computing the correlations across quarters, important concurrent and antecedent variables were not controlled. In their opinion it was necessary to control for unresponsiveness in the same quarter and infant crying in an earlier quarter in computing the correlations between crying in a later quarter and unresponsiveness in an earlier quarter. Data would allow the use of parametric methods, and thus for partializing variables. Ainsworth and Bell (1977), however, considered their sample too small to allow for parametric analyses such as cross-lag panel analysis or regression techniques. Furthermore, Gewirtz and Boyd (1977a) stated that there was an artificial dependence between infant and mother variables. A mother could neither respond nor ignore her infant's crying until the infant actually cried. Ainsworth and Bell (1977) agreed that within-quarter correlations between crying and unresponsiveness inevitably would be in-

flated because of confounded measures. Gewirtz and Boyd (1977a) also showed that separate analyses on inherently correlated frequency and duration measures for crying and unresponsiveness must have inflated in an unknown way the type I error probabilities. Later findings led Ainsworth and Bell (1977) to attach much more importance to the durational measures than to the frequency measures, because anxiously and securely attached dyads appeared to differ significantly only in the former measures. Lastly, Gewirtz and Boyd (1977a) pointed out that maternal responsiveness is not necessarily the inverse of maternal unresponsiveness as operationalized by Bell and Ainsworth (1972), although their conclusions appeared to imply this inverse relation. Ainsworth and Bell (1977) conceded that Gewirtz and Boyd (1977a) were technically correct, but they emphasized the psychological validity of considering maternal responsiveness the inverse of unresponsiveness. Ainsworth and Bell (1977) stated that the criticisms by Gewirtz and Boyd (1977) could be reduced to a claim of "not proven." They were confident that their conclusions would find confirmation in future research.

Existing Replication Effects

Because of the absence of a strong replication tradition in attachment research in general (van IJzendoorn & Tavecchio, 1987), and because of the very time-consuming nature of the data collection and coding involved in this specific study, few studies have been carried out to replicate even part of Bell and Ainsworth's provocative findings. First, Belsky, Rovine, and Taylor (1984), observing 60 mother-infant dyads for 45-minute periods in the first, third, and ninth month after birth, found that mothers engaging more in reciprocal interaction had babies who cried less later on in their first year of life. Maternal (un-)responsiveness to infant crying, however, was not implied in the reciprocal interaction measure. Second, Crockenberg and Smith's (1982) study involving 54 mother–infant dyads and covering the first three months showed that first-born children and children whose mothers were less responsive in their expressed attitudes fussed and cried more at three months. Maternal responsive behavior, however, did not predict infant crying. In a follow-up study, Crockenberg and McCluskey (1986) found a significantly positive relation between maternal responsiveness at three months and separation crying during the Strange Situation at 12 months, indicating that babies whose mothers had responded less quickly cried longer. It is not clear, however, whether crying during the 3-min separation episodes of the Strange Situation is a valid indicator of crying at home. Finally, Grossmann, Grossmann, Spangler, Suess, and Unzner (1985) observing 49 families for about 1 hour in the second, sixth, and tenth month after the birth of the baby, rated the mothers on a responsiveness scale and found that the more responsive the mothers were, the less their babies cried. This relation, however, is

based only upon tenth-month observations, and interquarter correlations to determine the influence of maternal responsiveness in an earlier quarter on crying behavior later on, were not reported.

Rationale for the Present Replication

We must, therefore, conclude that satisfactory empirical evidence in support of the hypothesized relation between maternal unresponsiveness and infant crying has still to be provided. We report here on a replication study in which some of the technical criticisms raised against the original study were taken into account. First, we present data on a larger sample ($N = 50$) and demonstrate the applicability of parametric analyses to control for antecedent and concurrent variables (Gewirtz & Boyd, 1977a). Second, we only analyze durational measures of infant crying and maternal unresponsiveness, thereby preventing inflation of type I error probabilities (Gewirtz & Boyd, 1977a). Third, because maternal responsiveness indeed is not necessarily inversely related to the durational measure of unresponsiveness (Gewirtz & Boyd, 1977a), we interpret the outcome of this measure only in terms of unresponsiveness. We also use a proportional measure of unresponsiveness, which is a priori the inverse of responsiveness. Lastly, we introduce a new measure for infant crying—"development of infant crying"—which appears not only to be a more valid indicator of the process Bell and Ainsworth (1972) tried to describe but also leads to much less inflated concurrent correlations with unresponsiveness. We leave aside, however, Gewirtz and Boyd's (1977a) criticisms that are based on theoretical considerations from a conditioning perspective. We do not think that studies carried out under an ethological attachment paradigm could be relevant to operant-learning theory (Gewirtz & Boyd, 1977a, 1977b). Bell and Ainsworth's (1972) design, as well as this report of our critical replication, are not intended to demonstrate presence or absence of conditioning contingencies between maternal unresponsiveness and infant crying, because data summarized in rather molar variables preclude functional analyses required by the experimental conditioning paradigm.

METHOD

Subjects

Given the limitations imposed by the restricted availability of four technical equipment sets (event recorder/F.A. audio registration unit) we were obliged to collect data from late summer 1983 to late summer 1984 for the first part of the sample ($n = 28$; a preliminary report on the first part of the study can be

found in Hubbard and Van IJzendoorn, 1987) and from fall 1986 to fall 1987 for the second part of the sample ($n = 25$). Due to the malfunctioning of technical equipment, three dyads had to be removed from our sample ($N = 50$). The sample consisted of original Dutch (i.e. nonimmigrant) families living either in, or in the neighborhood of, two Dutch cities, The Hague and Leiden. Eighty percent of the families were located through midwives (private practice), and 20 percent through municipal birth records. All infants were normal, healthy, full-term deliveries (with the exception of one caesarian delivery). The age of the babies ranged from 3 weeks (first observation) to 36 weeks (last observation). Twenty-six of the babies were boys; 24 were girls. The mean educational level for the mothers was 5.3 ($SD = 2.4$) on a scale ranging from 1 (six years of schooling) to 9 (16 years of schooling). The mean educational level of the fathers was 5.7 ($SD = 2.5$). The sample could be described as representative of young lower to middle-class families with two parents in which parental roles were traditionally allocated between spouses.

Procedure

All subjects were visited 12 times at home at 3-week intervals during the first nine months. The hypotheses of the study were not revealed. Visits were scheduled at the mother's convenience, the only restriction being that morning, afternoon, and evening observations were needed (unfixed order) to get a representative sample of the baby's crying behavior. Mother–infant pairs were visited regularly by one female observer, except for an occasional joint visit made by two observers for reliability checks. The last five visits on seven families were made by the first author. Visits lasted for 2 hours (first quarter) to 4 hours or more (third quarter). Observers planned to arrive given the latest expectation of the mother half an hour before the baby awoke, to install the technical equipment in a standby operating mode. Usually there was ample opportunity for interviews after the observation period. The observation period started when the baby awoke and finished when the baby fell asleep (or after 3 hours if the baby was longer awake). During the observation period the observer was obliged to play a low profile, semiparticipant role so as to be able to attend continuously to the ongoing stream of behavior recorded on the event recorder. For each quarter data of four successive observation periods was used to compute measures. The mean total observation period for the first quarter was 5.8 hours ($SD = 1.6$); second quarter: 7.1 ($SD = 1.9$); third quarter: 8.8 ($SD = 2.1$).

In order to cope adequately with the complex observational situation, the use of technical equipment was necessary. An event recorder (Epson HX 20 portable mini-computer operating on batteries) was used to code the mother–infant interactions continuously. The coding scheme of the event recorder consisted of three coding types: contextual codes (three digits); behavioral

codes (six digits), and system codes (one or two digits). An example of a typical relevant sequence is shown here in Table 12.1. The sequence starts with a contextual code A4E: the baby is actively awake; no caretaking situation, mother is in another room and can hear the baby cry. Interaction episode 1 (E1) indicates that the baby starts to cry, while mother is busy in the kitchen (ic3mbc); the next episode starts when mother enters the room and the baby still cries (meric3); E4: mother picks up/holds the baby, who is still crying; the next code is contextual, indicating the same behavioral state of the baby and close proximity; E5: the baby initiates this episode by looking at the mother (crying stops), and the mother vocalizes positively. In this way the observation session was segmented into interaction episodes dictated by behavioral changes either on the part of the mother or the baby (or both). The time table shows the start of the interaction episodes. The contextual code does not affect the time table. The task of the observer was reduced to accurately coding behavioral changes using a letter-coding scheme (see for a detailed description of the coding system Hubbard, 1989). In this chapter the focus is on the onset of maternal interventions, and not on the type of intervention.

During the observation period vocalizations of the baby and the mother (if she was in the same room) were recorded with an audio registration unit. This unit contained three components: (a) a wireless FM-transmitter/microphone combination (type Sennheiser SK 1012 (MKE 2012); (b) an Uher report 4400 portable stereo tape recorder; (c) a FM receiver (type Sennheiser/Telefunken EM 1008). In order to synchronize the event recorder and the audio-registration unit on a time-reference axis, a time-code generator was used to write the time (coded by tone pulses) on track one of the tape recorder. The vocalizations of the baby and the mother were recorded on track two. At the beginning of the visit, the time-code generator was connected up to the event recorder in order to match the two internal clocks. The audio-registration unit was placed in the living room, out of sight of the mother. The FM transmitter/microphone combination was always placed in the room where the baby was. The observer could check in the living room if the baby was awake via her earphones. When the baby awoke, the unit was started up, and could operate

Table 12.1. An example of coding interactional sequences

Code type	*Time table*	*Code*	*Behavioral episode*
	hrs:min:sec		
context	11:15:00	A4E	—
interact.	11:15:00	ic3mbc	E1
"	11:17:30	meric3	E2
"	11:18:00	mcpic3	E3
"	11:20:00	mphic3	E4
context	11:21:00	ACF	—
interact	11:21:00	illmvp	E5

for 90 minutes before the tape needed changing. In the meantime the observer was able to code the interactions of mother and baby by means of the portable event recorder.

The technical equipment was chosen to enhance the reliability and the validity of the data. Under all circumstances it could register the infant crying without being intrusive. If the baby was crying out of earshot of the observer (and often of the mother too) it was recorded. If the mother wanted the baby to sleep in the baby's room as usual during the observation session, this was possible. The observer would remain in the living room, checking unobtrusively from time to time through the earphones out of earshot of the mother, to hear if the baby was still awake. The registration of infant crying did not interfere with the mother's perceptions and consequently not with possible interventions. The job of the observer was easier than in the Baltimore study, because of the fact that the timing of relevant events was done by the event recorder and the time-code generator. Bell and Ainsworth (1972, p. 1174) remarked: "observers in their participant roles sometimes found it impossible to time accurately and hence resorted to estimates."

Infant Crying

Bell and Ainsworth (1972) had defined the central unit measure, that is, the crying episode, as any instance of a vocal distress signal (protest, fuss, or full-blown cries) not too brief to be timed (on the spot) and separated by more than a momentary pause from another instance. We took over this definition, except for the rather vague and not replicable value of the two parameters involved, minimal duration and time lag between subsequent crying. Every cry signal separated by a pause of 2 seconds from the next crying instance was coded as a crying episode.

Tape recordings were analyzed with a time decoder which displayed the timetable recorded during the observation. Given that there was synchronization between the timetable of behavioral episodes (event recorder) and the timetable of corresponding tape recordings, it was possible for the coders to make an accurate quantitative analysis of infant crying and the interventions of the mother, using the observers' coded observations as a guideline. Six coders analyzed the vocalizations of the baby either in a couple in the beginning or alone (after some 25 analyses). They were initially trained, using a record of infant crying (Wasz-Höckert, Lind, Vuorenkoski, Partanen, & Valanné, 1968) and our own tape recordings of crying. Vocalizations were analyzed twice before being coded as crying or noncrying. In fact, the analysis of one coder was independently checked by another coder. Disagreements between coders were coded as noncrying. After the check by the second coder the onset and finish of every crying episode was noted.

Mother's Unresponsiveness

The onset and end of every intervention or intervention sequence was noted with respect to every crying episode. Because we used a sensitive wireless microphone the vocalizations of the mother were also available on track two of the taperecording. Given the time table recorded on track one, the onset of verbal interventions was noted from the taperecording. Furthermore, coders regularly inspected the accuracy of the observers with respect to the timing of maternal interventions. These checks revealed that observers were accurate by means of the event recorder (the coding of onset time took a one-button press) and that the average delay was about two seconds. In the rare case that there was a large discrepancy this was corrected by the coders using cues of tape recordings. For example the observer might be less accurate for the intervention "enters room," if a mother rushes to the baby upstairs. The onset of this nonverbal intervention was also available on tape recording because the microphone was in the baby's room. The duration of unresponsiveness equaled the time the baby cried without an intervention of the mother (mother's delay), and equaled the duration of crying if an intervention started later than two seconds after the crying episode stopped. The two-seconds criterion is only relevant for nonverbal interventions and is only a matter of precision. In fact, we adopted the Bell and Ainsworth definition of unresponsiveness. Maternal interventions were, for example, picking baby up, holding; vocalizing; changing position; offering pacifier or toy; removing noxious stimulus; entering room. The focus of this chapter is on the delay of maternal interventions, not on the type of intervention (Bell & Ainsworth, 1972).

Reliability

Intercoder reliability. The time-tabled tape recording of infant crying and maternal vocalizations made it possible to calculate duration of crying and the delay of maternal verbal interventions post hoc in such a way that intraobserver and interobserver error variance due to measurement on the spot (Bell & Ainsworth, 1972) were eliminated. The distinction between crying and noncrying was made by consensus (i.e., if there was disagreement among coders about crying, then vocalizations were excluded from the data). For vocalizations of the baby with a minimal duration of five seconds, agreement percentage for crying was 95 percent for a random sample of 60 visits out of a total of 600 visits. Given the method of consensus reliability of duration of crying is reduced to the reading of onset and finish on a time reference axis. The mean agreement among coders for duration of unresponsiveness verbal interventions was 98.4 percent for a sample of 60 visits. For nonverbal interventions mean agreement percentage was 95 percent (from a sample of 14 visits done by two observers).

Data Analysis

Several different operationalizations of crying episode were computed with a computer program (based on Fortran) by specifying the two parameters: minimal duration of a crying episode and pause between two crying instances (see Hubbard, 1989; Hubbard & Van IJzendoorn, 1987, 1991). In this chapter we focus on duration measures (Ainsworth & Bell, 1977), based on the following definition of a crying episode: a distress signal with a minimal duration of five seconds, separated by four seconds or more from another instance. For each quarter maternal and infant measures were summarized for four successive observations. These measures are supposed to be comparable to Bell and Ainsworth's (1972) measures.

First, univariate descriptive statistics for infant crying behavior and maternal unresponsiveness are described. Second, in order to replicate the results of the original study, interquarter correlations between infant crying and maternal unresponsiveness are presented. Third, cross-lag panel analysis and a partial correlation approach are applied to control for contaminating variables. Last, using a development of duration of infant crying variable, hierarchical multiple regression analyses are carried out to compute standardized regression weights between maternal unresponsiveness and development of infant crying in subsequent quarters. Sample size did not allow for inferential causal modeling by means of LISREL (Boomsma, 1982; Tanaka, 1987).

RESULTS AND DISCUSSION

Developmental Changes in Infant Crying and Maternal Unresponsiveness

First we discuss changes in duration of infant crying and maternal unresponsiveness throughout the first 9 months. We found the same overall reduction in duration of crying as was found in the Baltimore study (see also Wolff, 1987, p. 80 ff.). The median of 5.9 minutes crying per hour in the first quarter is reduced to a median of 2.9 minutes per hour in the third quarter. bell and Ainsworth (1972) reported a median of 7.7 minutes crying per hour in the first quarter and 4.4 minutes per hour in the last quarter, and observed the same wide range of more than 15 minutes per hour to almost no crying at all, with a narrowing of the range toward the end of the first year. Wolff (1987) found crying and fussing about 10% of the time he observed the infants during the first quarter of their life.

The duration of maternal unresponsiveness also decreased in the first nine months from a median of 4.0 minutes per hour (Bell & Ainsworth, 1972; 3.8

minutes per hour) to 1.9 minutes per hour in the third quarter (Bell & Ainsworth, 1972, reported the same median for the fourth quarter). The medians for all measures were close to the means, indicating that the distribution are approximately normal. A repeated measures analysis of variance with age as factor confirmed the trend of decreasing duration of infant crying and maternal unresponsiveness during the first 9 months. For each quarter, analyses of variance with parity and sex as factors and duration of unresponsiveness and crying as dependent variables were computed. These analyses yielded no significant results.

Stability of infant crying and maternal unresponsiveness. A significant (auto-) correlation for duration of crying was found only between the first and second quarter ($r = .64$; $p < .001$). Bell and Ainsworth (1972) reported a significant correlation between the third and fourth quarter. In our case, results pointed toward instability of individual differences in duration of crying after the second quarter. Duration of maternal unresponsiveness showed about the same pattern. Bell and Ainsworth's (1972) data showed stability for the whole year.

Relation between infant crying and maternal unresponsiveness. We also found significant positive interquarter correlations between earlier maternal unresponsiveness and later infant crying. If these correlations were not spurious, they would indicate that babies who cried longer had mothers who were more unresponsive to their infants' crying in earlier quarters. Only the correlation between the first and second quarter ($r = .53$), and the first and third quarter ($r = .28$), appeared to be significant, whereas Bell and Ainsworth (1972) reported also significant figures for the last two interquarter correlations. On the basis of this correlational pattern Bell and Ainsworth (1972) concluded that those mothers who tended to delay response in one quarter would see their baby cry longer in the next one. In our study, this conclusion would certainly apply for the first half year, with the same tendency in the third quarter.

Controlling for antecedent and concurrent variables. Because intraquarter (synchronous) correlations between crying and unresponsiveness were very high, and the autocorrelations for crying and unresponsiveness were also considerable, statistical control for contaminating variables had to be made (Gewirtz & Boyd, 1977a). First, a cross-lag panel analysis was carried out. Our data allowed for cross-lag panel analysis because synchronous (or intraquarter) correlations did not differ significantly, and because a stable difference existed between parallel autocorrelations, indicating a stationary process (Cook & Campbell, 1979; Kenny, 1975). The reliability of synchronous panels was also comparable, because the same procedures were used to

measure unresponsiveness and crying. The largest difference between the three cross-lag correlation pairs was small (.06) and not significant, indicating that the alleged causal relation between unresponsiveness in an earlier quarter and crying in a later quarter is spurious.

Second, crying in an earlier quarter was partialled from the association between crying in a later quarter and unresponsiveness in an earlier quarter, to confirm our cross-lag panel analysis using a different statistical approach. For the first two quarters the partial correlation between unresponsiveness and crying was .03 (n.s.), and for the last two quarters this partial correlation was −.19 (n.s.). The significantly positive bivariate correlations between earlier unresponsiveness and later crying disappeared after controlling for duration of crying in the earlier quarters. Partialing unresponsiveness in the second quarter and in the third quarter also yielded nonsignificant partial correlations between earlier unresponsiveness and later crying. We have to conclude, that Bell and Ainsworth's (1972) hypothesis about the relation between unresponsiveness and crying is not supported by the partial correlations, nor by the results of the cross-lag panel analysis.

Development of crying. Because Bell and Ainsworth's (1972) measure for duration of infant crying did not take into account individual differences in development of crying, we computed a variable "development of duration of infant crying," defined as the log-transformed proportional decrease or increase of crying for two successive quarters.[1] This measure was used as a criterion variable to answer the question whether infants with more unresponsive mothers decrease their crying faster or slower than infants with less unresponsive mothers. In fact, it was this developmental question that Bell and Ainsworth (1972) tried to address in their Baltimore study, because they already had registered how infants in general decreased their crying during

[1]The development measure is not a difference score, but a log-transformation of the quotient of crying during the second quarter divided by crying during the first quarter, and so forth. Change scores have been criticized because of their lesser reliability, their inability to take initial differences into account, and their artificially high negative correlation with initial scores (Lord, 1963; O'Connor, 1970; Visser, 1982). Rogosa, Brandt, and Zimovski (1982; see also Zimmerman & Williams, 1982) showed, however, that under certain conditions this criticism is unwarranted. In our case, the reliability of the original variables is very high, because they were registered in a purely mechanical way on four different occasions per quarter. Statistical regression to the mean would, therefore, occur within those four different times of measurement. Furthermore, the developmental measure constitutes a change score based upon a quotient that takes initial differences into account, and through the log transformation the measure is symmetrically distributed. The correlations between the developmental measure and the initial score were in our sample −.31 ($p < .05$) for the first two quarters, and −.07 (n.s.) for the last two quarters (cf. Zimmerman & Williams, 1982). These correlations are quite modest, they can be considered "facts of life" (Rogosa et al., 1982), and in a multiple regression approach the significant correlation can be controlled for.

the first year. Besides representing the developmental changes more adequately, our developmental measure has the advantage to be correlated less strongly with concurrent maternal unresponsiveness (.43 and .69 for the second and third quarter respectively), thereby avoiding the multicollinearity problem. A hierarchical multiple regression analysis on development of crying in the first half year was carried out. Unresponsiveness in the second quarter and crying in the first quarter were entered into the equation first, and unresponsiveness in the first quarter was entered thereafter.

From this regression it can be derived that controlling for earlier quarter crying and concurrent unresponsiveness (Gewirtz & Boyd, 1977a), resulted in a significantly negative standardized bèta weight for the relation between first quarter unresponsiveness and development of crying in the first half year. This result did not support Bell and Ainsworth's (1972) hypothesis: Mothers who tended to be more unresponsive in the first quarter had babies who decreased their crying relatively more in the first half year. To test whether this outcome was due to the subgroup of infants who cried relatively short (cf. Ainsworth, personal communication; Landau, 1982), the sample was divided in two groups using the median for duration of crying as criterion. In the subgroup of infants crying relatively short ($n = 25$), the bèta weight for unresponsiveness in the second quarter was .69 ($p < .001$), for crying in the first quarter $-.46$ ($p < .01$), and for unresponsiveness in the first quarter $-.35$ ($p < .05$). In the subgroup of infants who cried relatively long ($n = 25$), bèta weights were .92 ($p < .001$), $-.28$ (n.s.); and $-.39$ ($p < .05$) respectively. In both subgroups, therefore, essentially the same relation between first quarter unresponsiveness and development of crying in the first half year was found as in the total group.

A hierarchical multiple regression analysis was also carried out on development of infant crying in the last two quarters. Unresponsiveness in the third quarter, development of crying in the first two quarters, and infant crying in the first quarter were entered into the regression equation first, unresponsiveness in the second quarter was entered thereafter, and unresponsiveness in the first quarter was entered last. Maternal unresponsiveness in the second quarter and the first quarter was not significantly related to development of infant crying during the last two quarters. This result does not support Bell and Ainworth's (1972) hypothesis, but also does not confirm the reverse relation between unresponsiveness and development of crying that appeared to exist in the first half year. Multiple regressions in the two subgroups of infants who cried short or long yielded the same result.

To address the question, raised by Gewirtz and Boyd (1977a), whether responsiveness is inversely related to unresponsiveness, the same analyses were carried out with a proportional measure for unresponsiveness. The percentage duration of crying ignored—which is, of course, the the inverse of the percentage of duration of crying responded to—did not correlate with first

quarter duration of crying ($r = -.01$). Outcomes of multiple regression analyses with this proportional measure for unresponsiveness were parallel to the results we reported upon above: a significantly negative bèta weight ($-.28$; $p < .05$) was found for the relation between first quarter maternal unresponsiveness and development of infant crying during the first half year, indicating that mothers who tended to respond more promptly, had infants who decreased their crying relatively less in the first half year of life. No such relation could be established for the development of infant crying after the second quarter.

CONCLUSIONS

According to Ainsworth and Bell (1977) an independent replication, preserving the intensive naturalistic longitudinal aspects of their exploratory study, was necessary before the dispute with their critics could be settled. The present study is such an effort. The same design has been used in which the mean total duration of the observations in every family is about seven to eight hours per quarter. Furthermore, more reliable and refined methods of observing, coding, and analyzing data in a larger sample have been applied. It is, therefore, interesting to see how far our descriptions of development in infant crying and maternal unresponsiveness converge with those of the original study, in view of the fact that our study has been carried out in another country more than fifteen years after the original one. Although Bell and Ainsworth (1972) collected data via the method of narrative accounts, for which reliability figures were lacking (Lamb, Thompson, Gardner, & Charnov, 1985), descriptive results did not differ markedly from those of our study. It is also evident that our bivariate interquarter correlations between infant crying and maternal unresponsiveness are consistent with the main conclusion of the original study: those mothers who tend to delay their response in one quarter, will see their babies cry longer in the next one (Bell & Ainsworth, 1972).

A cross-lag panel analysis, however, indicated that the correlations between unresponsiveness in an earlier quarter, and crying in a later quarter have to be considered as spurious. Partialling crying in an earlier quarter from the association between crying in a later quarter and unresponsiveness in an earlier quarter also resulted in nonsignificant interquarter correlations. Therefore, we have to conclude that Bell and Ainsworth's (1972) outcome is not replicated in our study. Although Ainsworth and Bell (1972, 1977) interpreted infant crying as a developmental phenomenon, differences in duration of crying at a certain point in time do not appear to reflect adequately differences in development of crying. Therefore, a new measure for development of infant crying was introduced operationalizing differences in individual "growth" of crying behavior. Using this new measure as criterion variable in multiple regression analyses we found a negative relation be-

tween earlier unresponsiveness and later development of crying in the first half year. Antecedent infant crying and concurrent maternal unresponsiveness were simultaneously controlled for. For the development of crying from the second to the third quarter we did not find a relation with earlier unresponsiveness. The data for the first half-year indicate that mothers who tend to delay their response relatively longer in the first quarter have infants who decrease their crying relatively more in the next quarter. This outcome was replicated in the two subgroups of infants who cry relatively long or short. Results were also the same if a somewhat more "precise" measure of crying episodes was used (i.e., every instance of a distress signal separated by two seconds or more from another).

The outcome of our critical replication study makes clear that the relation between maternal unresponsiveness and infant crying is somewhat more complex than originally was suggested. The technical critisms of Gewirtz and Boyd (1977a) on the Baltimore study appear to be supported. The critical replication study does of course not imply a refutation of attachment theory as a research program (Van IJzendoorn & Tavecchio, 1987), but indicates the necessity of adapting the theory to our unexpected outcome. Maybe current attachment theory should take Bowlby's (1971, p. 374) contention more seriously into account that mothers may condition certain infant behaviors. The conditioning paradigm has demonstrated experimentally that mothers are able to reinforce certain infant crying behaviors (Petrovich & Gewirtz, 1985). We propose, therefore, that in future studies the model of "differential responsiveness" should be tested. This model implies that only severe distress vocalizations (e.g., the pain cry) should be conceptualized as evolutionary "biased" attachment behavior, requiring a prompt response. The development of mild distress vocalizations ("operant" or "instrumental" cries) may solely be explained by proximate causes, or in terms of conditioning. Bowlby (1971, p. 347) already considered crying as a graded signal (Murray, 1979), carrying different information depending upon context, intensity, and rhythm (Thompson & Lamb, 1984), and he did not think it necessary for mothers to react promptly to all crying behaviors (cf. Ainsworth, 1973). A delayed response to mild distress may enable the infant to learn to cope with the kind of situations in which mild distress arises (Landau, 1982). Furthermore, simultaneously ignoring fussing behavior but being responsive to other behaviors such as exploratory play might diminish mild distress vocalizations (Lester, 1985). Only a minor percentage of infant crying behavior consists of pain or panic cries (Wolff, 1987) and might be considered attachment behavior. The model of differential responsiveness may, therefore, explain why in our study mothers who delayed their response to infant crying had children who tended to cry less later in the first half year of life. This model, however, can only be tested thoroughly in experimental studies discriminating between different types of infant distress vocalizations.

Our replication of the Baltimore study is, of course, not an exact replication

(see Chapter 3). Our study has been carried out in another culture some decades after the original study. Nevertheless, the raw data of the replication study did converge with those of the Baltimore study. It appears that a secondary analysis of the original study would have resulted in the same conclusions. A replication in another culture at a different point in time, however, also shows the remarkable robustness of the development of infant crying and its relation to maternal unresponsiveness. In this respect, the behavioral sciences do not always have to deal with phenomena that are more elusive than those studied in the natural sciences (see Chapter 5). Although our replication produced an unexpected outcome, it confirms at least the possibility of replicating social phenomena across a considerable cultural and historical span.

REFERENCES

Ainsworth, M. D. S. (1973). Systems for rating maternal care behavior. In E. G. Boyer, A. Simon, & G. R. Karafin (Eds.), *Measures of maturation. An anthology of childhood observation instruments*. Philadelphia: Research for Better Schools.

Ainsworth, M. D. S., & Bell, S. M. (1977). Infant crying and maternal responsiveness: A rejoinder to Gewirtz & Boyd. *Child Development, 48*, 1208–1216.

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.

Bell, S. M., & Ainsworth, M. D. S. (1972). Infant crying and maternal responsiveness. *Child Development, 43*, 1171–1190.

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 contribution. *Child Development, 55*, 718–728.

Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Causality structure-prediction* (Part I, pp. 149–173). New York: North-Holland.

Bowlby, J. (1971). *Attachment and loss. Volume I: Attachment*. Harmondsworth, UK: Penguin.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation. Design and analysis issues for field settings*. Chicago: Rand McNally.

Crockenberg, S., & McCluskey, K. (1986). Change in maternal behavior during the baby's first year of life. *Child Development, 57*, 746–753.

Crockenberg, S. B., & Smith, P. (1982). Antecedents of mother-infant interaction and infant irritability in the first 3 months of life. *Infant Behavior and Development, 5*, 105–119.

Gewirtz, J. L. (1961). A learning analysis of the effects of normal stimulation, privation and deprivation on the acquisition of social motivation and attachment. In B. M. Foss (Ed.), *Determinants of infant behavior* (pp. 213–299). London: Methuen.

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

Gewirtz, J. L., & Boyd, E. F. (1977b). In reply to the Rejoinder to our critique of the 1972 Bell and Ainsworth report. *Child Development, 48*, 1217–1218.

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 of attachment theory and research. Monographs of the Society for Research in Child Development, 50*(1-2, Serial No. 209), 233–256.

Epilogue

We have reached the end of our book; a book that was meant to investigate one of the canons of traditional research methodology: that experimenters in both the natural and social sciences should strive for replication or replicability. In the course of the book it has become clear—we believe—that the traditional belief that replication involves isomorphic copying of the original studies is ill-founded The introduction to this book and the chapters in the first section prepared the ground for this conclusion. The contributors to Part I dealt with questions concerning the nature, function, and possibility of replication. It was argued that there are no ahistorical or cross-cultural rules about what constitutes replication and that exact replication is impossible. Rather, researchers should strive for theoretically interesting conceptual replications. Such replications provide the fuel for the never-ending debate about the validity of conclusions drawn on the basis of research data. In Part II the problem of method effects formed a major topic of discussion. It was argued that studies may not be replicable as the results are not method-invariant. Several suggestions for the prevention of method effects were given and a major new instrument to detect method effects—meta-analysis—was described in great detail. The authors agree that exact replication is either impossible or fruitless and that replication efforts should be theory-driven. In Part III several attempts to replicate interesting original studies have been described. The main conclusion of the careful replications of the Istomina study was that it may well be an impossible task for any experimenter to exactly reconstruct functionally similar experimental settings across cultures or historical periods. The replication studies related in the other chapters deliberately introduced variations in the original research

Hubbard, F. O. A. (1989). *Infant crying and maternal unresponsiveness.* Leiden, The Netherlands: DSWD Press.

Hubbard, F. O. A., & Van IJzendoorn, M. H. (1987). Maternal unresponsiveness and infant crying. A critical replication of the Bell & Ainsworth study. In L. W. C. Tavecchio & M. H. van IJzendoorn (Eds.), *Attachment in social networks. Contributions to the Bowlby-Ainsworth attachment theory* (pp. 334–375). Amsterdam: Elsevier Science Publishers (North Holland).

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

Kenny, D. A. (1975). Cross-lagged panel correlations: A test for spuriousness. *Psychological Bulletin, 82,* 887–903.

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.

Landau, R. (1982). Infant crying and fussing. *Journal of Cross-cultural Psychology, 13,* 427–443.

Lester, B. M. (1985). Introduction: There's more to crying than meets the ear. In B. M. Lester & C. F. Z. Boukydis (Eds.), *Infant crying. Theoretical and research perspectives* (pp. 1–28). New York: Plenum Press.

Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change.* Madison: University of Wisconsin Press.

Murray, A. D. (1979). Infant crying as an elicitor of parental behavior: an examination of two models. *Psychological Bulletin, 86,* 191–215.

O'Connor, E. F. (1970). *Extending classical test theory to the measurement of change.* Unpublished dissertation, University of California, Los Angeles, CA.

Petrovich, S.B., & Gewirtz, J.L. (1985). The attachment learning process and its relation to cultural and biological evolution: Proximate and ultimate considerations. In M. Reite & T. Field (Eds.), *The psychobiology of attachment and separation* (pp. 259–286). New York: Academic Press.

Rogosa, D., Brand, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92,* 726–748.

Sroufe, L. A. (1985). Attachment classification from the perspective of infant-caregiver relationships and infant temperament. *Child Development, 56,* 1–14.

Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58,* 134–146.

Thompson, R. A., & Lamb, M. E. (1984). Individual differences in dimensions of socio-emotional development in infancy. In R. Plutchik & H. Kellerman (Eds.), *Emotion: theory, research and experience (Vol. 2). Emotions in early development* (pp. 87–114). New York: Academic Press.

Van IJzendoorn, M. H., & Tavecchio, L. W. C. (1987). The development of attachment theory as a Lakatosian research program: Philosophical and methodological aspects. In L. W. C. Tavecchio & M. H. van IJzendoorn (Eds.), *Attachment in social networks. Contributions to the Bowlby-Ainsworth attachment theory* (pp. 3–31). Amsterdam: Elsevier Science Publishers.

Visser, R. (1982). *On quantitative longitudinal data in psychological research.* Unpublished doctoral dissertation, University of Leiden, The Netherlands.

Wasz-Höckert, O., Lind, J., Vuorenkoski, V., Partanen, T., & Valanné, E. (1968). The infant cry: A spectrographic and auditory analysis. *Clinics in Developmental Medicine, 29,* Lavenham, Suffolk: Spastics International Medical Publications.

Wolff, P. H. (1969). The natural history of crying and other vocalizations in early infancy. In B.M. Foss (Ed.), *Determinants of infant behaviour IV* (pp. 81–109). London: Methuen.

Wolff, P. H. (1987). *The development of behavioral states and the expression of emotions in early infancy. New proposals for investigation.* Chicago: University of Chicago Press.

Zimmerman, D. W., & Williams, R. H. (1982). Gain scores in research can be highly reliable. *Journal of Educational Measurement, 19,* 149–154.