Responsiveness to infant crying: Spoiling or comforting the baby? A descriptive longitudinal study in a normal Dutch sample

Frans O.A. Hubbard & Marinus H. van IJzendoorn
Department of Education
Leiden University

Infant crying is one of the most provocative and ubiquitous behaviors in the first year of life. It is capable of evoking strong feelings from caregivers and implies many young parents’ first confrontation with the drawbacks of child rearing. Excessive and unpredictable crying has been identified as one of the main causes of child abuse and neglect (Frodi, 1985b). Generally speaking caregivers are aiming at control and restriction of crying on a short-term basis. For that matter, the question of which caregiving style is most effective in preventing excessive crying is frequently posed in the child care literature.

Spoiling the baby

The question of whether it is possible to spoil a baby is a recurrent issue. Spoiling denotes that the baby will become demanding and dependent because of a high rate of responding to crying. Before the publication of the influential paper on 'Infant crying and maternal responsiveness' by Bell and Ainsworth (1972) it was implicitly assumed in the child care literature that on the one hand the caregiver should respond to crying to alleviate the physical distress apparently underlying crying but that on the other hand consistent prompt responses to any type of crying could foster that very crying and 'spoil' the child (cfr. Gewirtz, 1977). In the Netherlands an important magazine on child care 'Ouders van Nu' ('Modern Parents') provides a good example for trends in advising on infant crying. Before the publication of the Baltimore study on infant crying, advice was based on a medical-somatic
viewpoint (advocated by the well-known Dr. Spock) and on popular belief about the danger of spoiling the baby. Early crying behavior was considered to be predominantly correlated to biological aspects and would naturally diminish in the long run due to maturational factors. According to popular belief consistently responding to crying was a counterproductive strategy.

In the so-called Skinnerian era, popular belief was in agreement with an operant learning account of crying that emphasized the distinction between expressive crying (due to physical distress) and non-distress or operant infant crying. Operant forms of crying would have a considerable potential of controlling caregiver behavior. Consistently responding to crying would reinforce operant crying and shape infant tyrannotaurus. Empirical findings which lend apparent support to a learning operant account resulted from experimental studies with preschool children (Williams, 1959; Harth, Allen, Buel, Harris, & Wolf, 1964). From these studies it is obvious that operant crying exists in preschool children. Etzel and Gewirtz (1967) concluded that operant types of crying already emerge during the first four months of life. They claimed to have demonstrated that a high rate of operant infant crying decreases systematically when an experimenter ceases responding to crying and (counter)conditions incompatible behaviors such as eye-contacting and smiling. Furthermore, Moss and Robson (1968) found a significant negative correlation between the latency of response to fussing at one month and the frequency of fussing at three months, indicating that prompt responses to fussing stimulate later fussing behavior. These correlations were derived from observations in the home of 54 first-born infants.

**Turning point: The Baltimore study**

From their descriptive longitudinal study of 26 middle class mother-infant pairs in Baltimore, Bell and Ainsworth (1972) concluded that consistent and prompt maternal responding to infant crying is associated with a reduction in frequency and duration of infant crying in later quarters. This provocative conclusion was based on patterns of positive rank order correlations between earlier maternal unresponsiveness and later infant crying across the four quarters of the first year of life. This conclusion was coherently embedded in an extended attachment theory on crying (Hubbard & Van IJzendoorn, 1987). The extension consisted of labelling the whole universum of crying behavior in the first half year as
attachment behavior and a prompt response from the caregiver as adaptive childrearing behavior. Crying behavior interpreted as attachment behavior was considered not to be reinforced by prompt responding. On the contrary, those infants in their sample at the end of the first year, who fit the stereotype of the 'spoiled child', had mothers who were unresponsive to their crying in earlier quarters.

Results of the longitudinal study were presented as unequivocal and the far reaching claims strongly influenced practical advice on dealing with infant crying. The practical appeal was obvious. Caregivers were advised to hold on to one simple rule paraphrased as "You can't spoil your baby by responding promptly to crying", and this rule seemed to be scientifically sound. In the Netherlands the magazine for young parents 'Ouders van Nu' reported in September 1973 on the interesting and appealing Baltimore findings. Further publications on the issue of 'infant crying' in 'Ouders van Nu' were solely based on the Baltimore study. In April 1977 an extensive report on the 'Baltimore study' was published once again, due to the fact that a medically oriented popular magazine 'Welzijn' ('Welfare') published an article (January 1977) on infant crying and caregiving behavior based on pre-Baltimore popular beliefs. This advice was qualified as unscientific by 'Ouders van Nu'. It would only give rise to unduly neglecting 'tiny infants' who are programmed to be cared for by responsive caregivers.

A cry for replication

Given the fact that the Bell and Ainsworth (1972) report seemed to provide a falsification of expectations under an operant learning paradigm it was criticized by operant-learning oriented researchers (Parsley & Rabinowitz, 1975; Gewirtz & Boyd, 1977a;b). The criticism by Gewirtz and Boyd (1977a;b) was most radical in questioning the internal validity of the Baltimore study. In their rejoinder, Ainsworth and Bell (1977) were obliged to admit that their main conclusion about "consistency and promptness of maternal response (being) associated with decline in frequency and duration of crying" (which was copied in the child care literature) was an illegitimate paraphrase from a technical viewpoint. They implicitly assumed that their measure for unresponsiveness was inversely related to responsiveness. That is, in correlating measures for unresponsiveness with crying inverse conclusions about the relation between responsiveness and crying seemed to be implied, as was suggested by the
title of their 1972 paper. This assumption, however, was later empirically falsified by their own data.

Furthermore Ainsworth and Bell (1977) made a distinction between the development of frequency of crying and development of duration of crying. As was requested by Gewirtz and Boyd (1977a) they computed the correlation between 'number of crying episodes responded to' and 'frequency of crying in subsequent quarters'. These correlations were not significant, falsifying their 1972 conclusion but also expectations under an operant paradigm.

Ainsworth and Bell (1977) stated that later findings had shown that anxiously and securely attached dyads appeared to differ significantly on the durational measure for crying. For that matter they attached much more importance to the durational measure. Gewirtz and Boyd's (1977a) criticism of spurious interquarter correlations due to the absence of control for antecedent infant crying and concurrent maternal unresponsiveness, was addressed with the argument that they considered their sample too small to allow for parametric analysis. Ainsworth and Bell (1977) were obliged to emphasize the need for replication: "We hold that the only satisfactory answer to the substance of the criticism that Gewirtz and Boyd raise is replication of findings with another sample" (Ainsworth & Bell, 1977, p. 1211).

Although Ainsworth and Bell (1977) did not impeach the claim that their data lend no support to their 1972 conclusion, the dominating influence on child care literature in the Netherlands was not deemphasified. Practically speaking the slogan 'You can't spoil your baby by promptly responding to crying' prevailed over warnings for spoiling. Child care literature such as popular parenting magazines were already heavily influenced by the 1972 paper, and did not take into account the complicated scientific controversy raised by the Baltimore study, nor Ainsworth and Bell's (1977) admission that their thesis about the relation between infant crying and maternal responsiveness should be considered unproven.

**Baltimore revisited**

Because of the very time consuming nature of the data collection and coding involved in the original study few studies have been carried out to replicate even part of Bell and Ainsworth's controversial findings. The results of related descriptive longitudinal observational studies done after 1977 did not point towards satisfactorily empirical evidence in support of the hypothesized relation between maternal unresponsiveness and infant crying (Belsky, Rovine
RESPONSIVENESS TO INFANT CRYING 229

& Taylor, 1984; Crockenberg & Smith, 1982; Crockenberg & McCluskey, 1986; Grossmann, Grossmann, Spangler, Suess & Unzner, 1985). A cross-sectional observational study by Landau (1982) indicated that the relation between responsiveness and fussing may also turn out to be contrary to the expectations of Bell and Ainsworth (cf. Moss, 1974). Landau found that Bedouin mothers who carried their babies all the time and were supposed to respond promptly after the onset of every little cry or fuss, had infants who fussed relatively more frequently than infants from comparison groups. This finding, however, was not considered to be inconsistent with Bell and Ainsworth's (1972) conclusions because even the most responsive Baltimore mothers did not respond to little fusses as promptly as Bedouin mothers, although the former caregivers rarely failed in responding to loud and prolonged crying (Ainsworth, pers. comm. to Landau, 1982).

We report here briefly on a replication study in which some of the technical criticisms raised against the original study were taken into account (Hubbard & Van IJzendoorn, 1988). The sample consisted of 50 non-immigrant Dutch families. All infants were normal, healthy, full-term deliveries (one caesarian delivery). The age of the babies ranged from three weeks (first observation) to 36 weeks (last observation). Twenty-six of the babies were boys, twenty-four were girls. The sample could be described as representative of young lower to middle class families with parental roles traditionally allocated between spouses.

Procedure

All subjects were visited by a female observer twelve times at home at three weeks intervals, during the first nine months of life. Visits lasted for two hours (first quarter) to four hours or more (third quarter). The observation period started when the baby awoke and finished when the baby fell asleep (or after three hours if the baby was awake longer). In order to cope with the complex observational situation infant crying and maternal verbal interventions were recorded with a wireless FM-audio registration unit on audio tape independently of the observer. The wireless microphone was always placed in the vicinity of the infant. The audio registrational unit was chosen to enhance the reliability and the validity of the data. It could register infant crying under all circumstances without being intrusive. Additionally the observer coded the observation session by event sampling with a portable event recorder. Audiotaped data and event-recorded data were synchronized by recording a corresponding
timetable on audio tape. The observer played a low profile semi-participant role and was not supposed to code crying episodes on the spot and to register time markers as in the Baltimore study. This procedure was criticized as lacking reliability (Lamb, Thompson, Gardner, & Charnov, 1985).

Tape recordings were analyzed by six coders who noted the onset and end of every crying episode. Infant vocalizations were analyzed twice before being coded as crying or non-crying. Disagreements between coders were coded as non-crying. A crying episode was considered to have been ignored by the mother if the intervention started later than two seconds after the episode stopped. The duration of unresponsiveness equalled the time the baby cried without an intervention of the mother (mother's delay), and equalled the duration of crying if an intervention started later than two seconds after the crying episode stopped. In this paper we will focus on a molar definition of the crying episode: a crying instance with a minimal duration of five seconds separated by four seconds or more from another instance. This molar measure is supposed to be more comparable to the Bell and Ainsworth measure than our molecular operationalizations (Hubbard & Van IJzendoorn, 1987). Agreement percentage for crying was 95% and for duration of unresponsiveness 98.4% for a random sample of 60 visits (95% for non-verbal interventions based on 14 visits).

Results

In this chapter we will focus on the interquartile correlations between (un)responsiveness and crying and their validity (see Hubbard (in prep.), and Hubbard & Van IJzendoorn (1988) for further details). Developmental trends for crying and unresponsiveness across quarters were consistent with the Baltimore sample. The mean rate of crying was constant over quarters but babies cried twice as long in the first quarter (6 minutes per hour) compared with the third quarter. Furthermore, we found correlations between maternal unresponsiveness in the first quarter on the one hand, and duration of crying in the second ($r = .53; p < .001$), and the third quarter ($r = .28; p < .05$) on the other hand. That is, mothers who were more unresponsive in an earlier quarter, seemed to have babies who cried longer in subsequent quarters. Bell and Ainsworth (1972) derived their provocative conclusion from about the same correlational pattern.
However, it can be demonstrated that the correlations between quarters are spurious because of the high correlations between unresponsiveness and crying in the same quarter (81, 96, and 90 for the first, second, and third quarter respectively), and the rather high stability of crying (64) and unresponsiveness (55) from the first to the second quarter. A cross-lagged panel analysis appeared to be a valid approach to entangle real from spurious correlations, considering some characteristics of our data (Hubbard & Van IJzendoorn 1988). The differences between the cross-lag correlation pairs were not significant (largest differences 06), indicating that a causal interpretation was not warranted. Crying in an earlier quarter could have elicited unresponsiveness in a later quarter as much as the reverse. Similar results were found for duration of responsiveness and crying.

Further analyses of our data, however, indicated that mothers may spoil their babies if they respond promptly to all crying and fussing. First, we computed a variable 'development of duration of crying', defined as the log transformed proportional decrease or increase of crying for two successive quarters, to take individual differences in development of crying into account, and to avoid the multicollinearity problem (the developmental measure correlated with concurrent maternal unresponsiveness 43 and 69 for the second and third quarter respectively). A hierarchical multiple regression analysis on development of crying during the first half year was carried out. Second-quarter unresponsiveness and first quarter crying were entered the equation first, and thereafter first quarter unresponsiveness was entered. In this way, we controlled for earlier quarter crying and concurrent unresponsiveness. The regression analysis did not support the Bell and Amsworth hypothesis, because a significantly negative standardized beta weight of 48 was found for the relation between first quarter unresponsiveness and development of crying. Division of the sample in infants crying relatively long (N=25), and infants crying relatively short (N=25) yielded essentially the same results. For development of crying between the second and third quarter, we did not find a significant relation between earlier unresponsiveness and later development, indicating that mother-infant interactions in the first half year are most powerful in shaping infants crying behavior.

Secondly, the results of analyses on the measures for frequency of crying were quite clear cut. Although Amsworth and Bell (1977) preferred the durational measures, we think that from a practical viewpoint unpredictable frequent crying may be more annoying to caregivers than (predictable) prolonged crying. Across quarters we found a significantly negative correlation between percentage of episodes ignored in the first quarter and frequency of crying in the
second quarter. A cross-lag panel analysis warranted the conclusion that mothers who ignored more crying had babies who cried less frequently in the next quarter. Similar results were found for the relation between number of episodes responded to and frequency of crying. These results contradict the Bell and Ainsworth (1972) hypothesis and are consistent with results reported by Moss and Robson (1968) and Moss (1974). Mothers who responded less frequently to infant crying had infants who cried less frequent in the next quarter (Hubbard & Van IJzendoorn, 1988; Hubbard, 1989).

Discussion and conclusion

According to Ainsworth and Bell (1977) and independent replication, preserving the intensive naturalistic longitudinal aspects of their study was necessary to test their hypothesis. The present study is such an effort and the outcome makes clear that the technical criticism of Gewirtz and Boyd (1977a) on the Baltimore study is supported. The findings of this study do not confirm the Bell and Ainsworth (1972) hypothesis and, therefore, do not warrant the popular advice based on the Baltimore study.

Maybe current attachment theory should take Bowlby's contention that mothers may condition crying behaviors (Etzel & Gewirtz, 1967) more seriously into account. We propose, therefore that in future studies the model of 'differential responsiveness' should be tested which implies that only severe distress vocalizations should be conceptualized as 'evolutionary biased' attachment behavior. The development of mild distress vocalizations may solely be explained by proximate causes, that is, in terms of conditioning. Although research in clinical populations, i.e. highly irritable babies from very low socio-economic status families (Van den Boom, 1988) seems to indicate that more responsive reactions lead to less crying, our model would imply that in a normal population parents should try to strike a balance between responding promptly to severe distress vocalizations and ignoring mild distress vocalizations such as fussing. We suggest that even the most responsive mothers of highly irritable babies are only able to respond promptly to the frequently occurring severe distress vocalizations, ignoring other types of crying and fussing, just because of natural restrictions in available time and energy. In a normal population, infants will show severe distress less often, and therefore overconcerned parents would have the opportunity to respond to almost all distress vocalizations, including fussing. They would, then, run the risk of conditioning mild distress
vocalizations, and seeing their babies cry more during the first year of life. Advocating a prompt response to any form of crying in a normal population using the slogan that parents can not spoil their baby is therefore scientifically not warranted. Before the model of differential responsiveness is thoroughly tested, we cannot present an alternative guideline, replacing the ill-founded slogan. For the time being, caregivers should rely on their intuition ('intuitive parenting', Papousek & Papousek, 1987), rather than trust on scientifically premature advices in determining how to deal with their infants' crying.

Note

1. This paper is based on research supported by a grant from the National Foundation for Preventive Studies (Praeventie Fonds, grant number 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 Mil, Ina van den Oever, Letty van Ooyen, Hansje Planjer, Inge Ruhaak, Marianne Thomecr, Nico Vesseur, Yvonne ten Wolde, André Wortelboer, Jeroen Zomerplaag, for their assistance in collecting, coding, and analyzing the data.

References


