



Universiteit
Leiden
The Netherlands

Harsh discipline in toddlerhood : A longitudinal study on maternal physiological and behavioral predictors

Joosen, K.J.

Citation

Joosen, K. J. (2012, November 22). *Harsh discipline in toddlerhood : A longitudinal study on maternal physiological and behavioral predictors*. Retrieved from <https://hdl.handle.net/1887/20154>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/20154>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/20154> holds various files of this Leiden University dissertation.

Author: Joosen, Katharina Jacomina

Title: Harsh discipline in toddlerhood. A longitudinal study on maternal physiological and behavioral predictors

Issue Date: 2012-11-22

Maternal sensitivity in various settings predicts harsh discipline in toddlerhood

*Katharina J. Joosen, Judi Mesman, Marian J. Bakermans-Kranenburg,
Marinus H. van IJzendoorn
Attachment & Human Development, 14, 101-117*

ABSTRACT

In a longitudinal study with 73 mothers and their second-born child, stability and main-level differences between measures of maternal sensitivity across settings and over time were examined. Furthermore, the predictability of harsh discipline by these different maternal sensitivity measures was studied. Maternal sensitivity was assessed at 3 and 6 months during bathing, free play on mother's lap and the baseline and reunion episode of the Still Face Paradigm (SFP; Tronick, Als, Adamson, Wise, & Brazelton, 1978). Harsh discipline was observed during three home visits in the second year of life. Results showed a single underlying factor for all maternal sensitivity settings at both time points and significant stability over time. Harsh discipline was predicted by maternal sensitivity at 3 months, which was fully mediated by maternal sensitivity at 6 months. Early failure to respond appropriately to infant signals is an important indicator of risk for future harsh parenting.

Keywords: maternal sensitivity, naturalistic settings, SFP, harsh discipline

INTRODUCTION

Within the framework of attachment theory, Mary Ainsworth was the first to provide a detailed description of maternal sensitivity, defined as mothers' ability to perceive child signals, to interpret these signals correctly, and to respond to them promptly and appropriately (Ainsworth, Bell, & Stayton, 1974). The extensive attention for sensitive parenting behavior is certainly not without ground, since many studies have shown significant relations between maternal sensitivity and a variety of positive outcomes such as secure attachment (Bakermans-Kranenburg, Van IJzendoorn, & Juffer, 2003; De Wolff & Van IJzendoorn, 1997), self-regulation (Eisenberg et al., 2001), social functioning (e.g., Kochanska, 2002; Van Zeijl et al., 2006), and cognitive competence (e.g., Bernier, Carlson, & Wipple, 2010; Tamis-LeMonda, Bornstein, & Baumwell, 2001). Furthermore, the absence of skills needed to respond sensitively to child signals has been linked to a risk for harsh parenting and maltreatment (Engfer & Gavranidou, 1987; Milner, 1993, 2003).

What started with Ainsworth's elaborate full-day observations developed into more efficient methods used across a variety of settings ranging from naturalistic daily routines to (semi-)structured play and stress-inducing paradigms (e.g., Alink et al., 2009; Braungart-Rieker, Garwood, Powers, & Wang, 2001; Susman-Stillman, Kalkoske, Egeland, & Waldman, 1996). However, it remains unclear what the consequences are of using one observational setting instead of another for the observation of maternal sensitivity. In the current study we aim to investigate maternal sensitivity in infancy in different situations in terms of (1) mean-level differences, (2) interrelations, (3) 3-month stability, and (4) the prediction of harsh discipline in toddlerhood.

Originally, Ainsworth's work in both Uganda (Ainsworth, 1967) and Baltimore (Ainsworth, Blehar, Waters, & Wall, 1978) was based on elaborate naturalistic observations of mother-infant dyads during all sorts of daily routines (i.e., bathing, feeding, etc.). The value of longer naturalistic observations lies in the fact that it gives a more detailed picture of the variety of maternal behaviors across different situations and under different conditions. However, due to time constraints naturalistic observations have generally been limited to single daily routines in more recent studies (such as bathing or feeding sessions, e.g., Albers, Riksen-Walraven, Sweep, & De Weerth, 2008). Nonetheless, some researchers still continue to collect extensive naturalistic observational data as well. (Highly & Dozier, 2009; Van IJzendoorn & Hubbard, 2000).

Structured play situations form an alternative and frequently used strategy in observing mother-infant interactions (e.g., Vereijken, Riksen-Walraven, & Kondo-Ikemura, 1997). Mothers generally receive instructions to interact with their infants with or without toys for a duration ranging from 5 to 15 minutes. By recreating part of the infant's daily experiences (e.g., play) this method does assure the observation of at least some dyadic interaction without the time investment of longer naturalistic observations. However, this method is not necessarily 'naturalistic' for all mothers, since not all mothers engage in this kind of dyadic play as a part of their daily routines.

Another shortcoming of structured play settings is that it limits the number of observations of maternal sensitivity to distress, whereas infant distress is particularly salient to the concepts of sensitivity and attachment (Bowlby, 1969/1982). In studies focusing on sensitivity to distress, dyads were only included in the final analyses when the infants displayed distress during a free play observation. Thus, when a non-distressing setting is used for the measurement of sensitivity to distress there is the danger of selecting subgroups of children prone to distress even in neutral situations (Leerkes, Blankson, & O'Brien, 2009; McElwain & Booth-LaForce, 2006). Nonetheless, sensitivity to distress measured in this setting does seem to be a stronger predictor of attachment security than sensitivity to non-distress (McElwain & Booth-LaForce, 2006), which is in line with findings that more sensitive nighttime responses to infant distress are related to more secure infant attachment (Aviezer, Sagi, & Van IJzendoorn, 2002; Higley & Dozier, 2009). Furthermore, sensitivity to distress has also been found to predict fewer behavioral problems and greater social competence in toddlerhood, while sensitivity to non-distress did not (Leerkes et al., 2009).

An observation paradigm that includes both non-distressing and potentially distressing interactions is the Still-Face Paradigm (SFP; Tronick, et al., 1978). Since its original development, the SFP has been used for a variety of purposes ranging from the assessment of dyadic interaction patterns to the measurement of infant coping strategies to stressful events. In general this standardized face-to-face dyadic interaction consists of three steps: (1) a baseline normal interaction, (2) the 'still-face' episode in which the adult becomes unresponsive whilst maintaining a neutral facial expression, and (3) a reunion in which normal interaction is resumed. The break in typical social interaction that takes place during the still face segment induces stress in the infant (Mesman, Van IJzendoorn, & Bakermans-Kranenburg, 2009). The classic 'still-face effect' consists of a decrease in positive affect and an increase in negative affect from baseline to still-face. In addition, the increased negative affect continues into the reunion episode ('carry-over effect'), which makes the reunion episode suitable for observing maternal sensitivity to distress. One study, for example, showed that during the reunion episode maternal involvement was related to more positive affect and less negative affect in infants (Rosenblum, McDonough, Muzik, Miller, & Sameroff, 2002).

In addition to being a predictor of various child outcomes, maternal sensitivity in infancy has also been found to predict sensitive parenting in later years (e.g. Jaffari-Bimmel, Juffer, Van IJzendoorn, Bakermans-Kranenburg, & Mooijaart, 2006). Similarly, a lack of sensitivity towards infants may be a risk for future negative parenting strategies, such as harsh discipline. In the literature on parents at risk for physical abuse, skills closely related to sensitive responsiveness play an important role. Milner (1993, 2003) describes a four-stage social information processing model in relation to physical abuse. The first stage in this model describes the fact that high-risk and abusive parents are less attentive to and less aware of their children's behavior as compared to nonabusive and low-risk parents. The second stage refers to interpretations of child signals and shows that high-risk and abusive parents are more likely to attribute hostile intent to (negative) child behavior and to view this behavior as due to internal, stable, and global child factors. The third stage of Milner's model describes a lack of adequate information integration for abusive and high-risk parents; these parents are unlikely to use situational or mitigating information in their response selection process. Furthermore, response selection of high-risk and abusive parents is limited by a more rigid use of punishment following child noncompliance. In the fourth stage, abusive and high-risk parents are less capable of monitoring and modifying their responses and parenting strategies.

These four stages show remarkable overlap with the main elements of Ainsworth's maternal sensitivity construct (i.e., awareness and interpretation of signals and appropriate responding), but are formulated in terms of the absence of these skills. Previous studies have also shown a link between maternal intrusiveness (a subscale of maternal sensitivity) and harsh or abusive parenting (Lyons-Ruth, Connell, Zoll & Stahl, 1987). The concept of intrusiveness, which can be defined as the lack of respect for the child's autonomy and (physical) interference with the child's behavior, especially during exploration, seems to be theoretically in line with the biased interpretations of, and rigid responding to (negative) child

behavior of harsh or abusive parents. In other words, intrusive as well as harsh or abusive parents seem to share a lack of empathy or understanding of the child's behavior and motives, which interferes with an appropriate interpretation of and responding to these signals during both normal interaction as well as discipline situations. Thus, there is reason to believe that a lack of sensitivity in infancy may predict harsh discipline strategies in toddlerhood.

Examining early predictors of harsh discipline is especially important since this parenting strategy has been shown to increase the risk for the development of problem behavior in (later) childhood (e.g., Bender et al., 2007; Kerr, Lopez, Olson & Sameroff, 2004; McKee et al., 2007). Beyond infancy, maternal discipline strategies become important as toddlers typically seek autonomy in the form of exploration of boundaries and noncompliance (Sroufe, 1995). If we can identify parents at risk for harsh discipline based on their parenting skills in infancy, this could enhance early preventive intervention efforts. To our knowledge there is no direct empirical evidence about the predictive value of sensitivity towards infants in relation to observed harsh discipline in toddlerhood, let alone in which situations sensitivity can best be measured to achieve optimal predictions.

In the current study we first examined whether levels of maternal sensitivity in infancy significantly differ across situations. Second, we investigated the stability of sensitivity in different situations over a 3-month period. Last, we tested whether maternal sensitivity assessed in different situations during infancy predicts harsh discipline in toddlerhood.

METHOD

Sample and procedure

Participants were recruited via midwifery practices in the western region of the Netherlands, and the Regional Coordination Programs of the Dutch National Institute for Public Health and Environment (NIPHE). Inclusion criteria were (1) the first child scored either low or high on externalizing behavior on the Child Behavior Checklist (CBCL/ 1½ -5; Achenbach & Rescorla, 2000; cutoffs based on Van Zeijl et al., 2006), (2) the first child was younger than six years old, (3) the newborn second child was neither premature nor handicapped, (4) both children and parents lived together in one household and (5) both parents were Caucasian.

Fifty-three midwifery practices handed out pamphlets about the study to pregnant women expecting a second child. Interested mothers could send in a card with their details after which they would receive a questionnaire including the CBCL and questions about demographic characteristics by mail. The NIPHE assisted in approaching families from the lower socioeconomic strata, which appeared to be underrepresented in the midwifery sample. Postal codes were selected of areas in which more than 30% of inhabitants had a low income (less than 14,200 Euro per year; as defined by the Dutch Central Bureau of Statistics). Families with a newborn infant living in these areas received a letter from the NIPHE with information about the study. Interested mothers of a second newborn infant were requested to fill out a card with their address, phone number, and the

name, gender and birth date of their newborn. Mothers were then contacted by phone to find out the educational level of both parents (1 *elementary school*, 2 *four years of Dutch high school*, 3 *vocational education or five or six years of high school*, 4 *higher professional education*, and 5 *university*). When both parents had a low (level 1 or 2) educational level or no more than one of the parents scored level 3, a home visit was planned to complete the CBCL about the first child. All mothers who completed the CBCL received a coupon with a value of 10 Euro.

The recruitment resulted in a total of 103 families (51 with a 'high externalizing' first child, and 52 with a 'low externalizing' first child). Of the 103 families originally included, a total of 73 families had sufficient data on all variables included in this paper (nine families were excluded from analyses due to missing data, and 21 families dropped out before the final home visit and were therefore excluded from the analyses for this paper). The 30 families that were excluded from the analyses did not differ from the final sample on maternal age, education, child gender, and maternal sensitivity during bathing and lap sessions (all $p > .05$). Excluded mothers were less sensitive in the still face baseline (excluded mothers: $M = 1.60$, $SD = 0.91$, included mothers: $M = 2.01$, $SD = 0.75$; $t(96) = -2.04$, $p < .05$) and reunion (excluded mothers: $M = 1.76$, $SD = 0.88$, included mothers: $M = 2.21$, $SD = 0.76$; $t(96) = -2.42$, $p < .05$) at 3 months but not at 6 months.

Within the final group, educational level of the mothers was divided as follows: 6.8% of mothers finished elementary school, 32.9% finished the first four years of Dutch high school, 28.8% obtained a degree in vocational education or finished five or six years of high school, 19.2% obtained a degree in higher professional education, and 12.3% completed at minimum a bachelor degree at university. Maternal mean age at time of birth of their second child was 31.17 years ($SD: 0.54$, range 19-40 years). The mean age of the first children at the time of selection was 31.64 months ($SD: 1.35$, range 15-70). Of the first-born children 52.1% were boys, and of the second-born children 35.6% were boys. In 41 families (56.2%) the first-born child showed low levels of externalizing behavior at the time of recruitment, and in 32 families (43.8%) the first-born showed high levels of externalizing behavior.

All 73 families participated in a total of six home visits during the first two years after the birth of the second child. Home visits were scheduled when the second child was 3, 6, 9, 12, 18 and 24 months old. During the first year, home visits included videotaping of the dyad in naturalistic situations (i.e., bathing and free play) as well as in the Still Face Paradigm (Mesman et al., 2009; Tronick et al., 1978). At 12 and 24 months mothers were observed with both children simultaneously during two discipline tasks. At 18 months mothers were observed with the second child only during discipline tasks.

In all participating families both parents signed informed consent forms. To limit loss of participants through attrition, families were given gift coupons and small presents for the children after each home visit. Families also received two DVD's with a compilation of video footage from all home visits. Procedures and measures used in this study were approved by the ethical committee of the Institute of Education and Child Studies at Leiden University.

Measures

Maternal sensitivity 3 and 6 months – naturalistic settings. Maternal sensitivity was assessed with the Ainsworth Sensitivity scale (Ainsworth et al., 1974) during free play on the mother's lap without toys (5 minutes), and during bathing (10-20 minutes). Observations took place during home visits at 3 and 6 months after the birth of the second child. All observations were rated independently by two trained coders. Inter-coder reliabilities (intraclass correlation, single rater, absolute agreement) for each pair of the six coders ranged from .75 to .92. When the scores of the two independent coders differed two or more points (on the 9-point scale) tapes were discussed and a consensus score was assigned. For smaller differences the average of both scores was used. Within the same home visit the bath and lap sessions were rated by the same coder, but observations from different home visits were coded independently.

Maternal sensitivity and intrusiveness 3 and 6 months - Still Face Paradigm. The Still Face Paradigm (SFP; Tronick et al., 1978) was used to measure maternal sensitivity and intrusiveness. The SFP consisted of three steps: (1) a baseline normal interaction (2 minutes), (2) the 'still-face' episode in which the adult becomes unresponsive while looking at the child with a neutral facial expression (1 minute at age 3 months and 2 minutes at age 6 months), and (3) a reunion in which normal interaction is resumed (2 minutes) (Mesman et al., 2009). Additionally, mothers were allowed to touch the infant as they would do so normally during baseline interaction as well as reunion, but were not allowed to touch the child during the still face segment.

Observations of maternal sensitivity and intrusiveness during baseline interaction and reunion were coded with an adapted version of the Mother-Infant Coding System (Miller, 2000). An overall scale for *Sensitivity* was used, with scores on a 4-point subscale defined as: 0 *No sensitivity*, 1 *Minimal or low sensitivity*, 2 *Mixed or moderate sensitivity*, and 3 *Predominant or high sensitivity*. A 4-point subscale for *Intrusiveness* was defined as: 0 *No intrusiveness*, 1 *Minimal intrusiveness*, 2 *Mixed or moderate intrusiveness*, and 3 *Predominant or high intrusiveness*. The SFP at ages 3 and 6 months were independently coded by two trained coders (inter-coder reliabilities i.e., intraclass correlation, single rater, absolute agreement for *Sensitivity*: .69 for *Intrusiveness*: .75). Baseline interaction and reunion within one assessment were scored by the same person. All tapes were double coded for *Sensitivity* since this scale was considered the more difficult one. When the two scores were not identical, a final score was given by two expert coders who were reliable on the initial reliability set (inter-coder reliability: .79).

Harsh discipline 2nd year. Maternal discipline was observed during all home visits in the second year (12, 18 and 24 months). Observations consisted of two tasks: a don't-touch task and a clean-up task. In the don't-touch task mothers were presented with a bag full of attractive toys. They were instructed to take the toys out of the bag, place them in front of the child and to make sure the child would not touch the toys. After 2 minutes (3 minutes at 24 months) the child was allowed to play only with the least attractive toy: a simple stuffed animal. After another 2

minutes the child was allowed to play with all the attractive toys as well as a bag of extra toys. Following 15 minutes of free play with all toys (not coded for this study), mothers received an empty bag with the instruction that all toys had to be cleaned up by the child. Instructions specified that she could help and encourage the child as she would normally do, but that it was important for the observation that the child would clean up as much as possible. The clean-up session was finished when all toys were back in the bag or when ended by the experimenter after a maximum of 10 minutes.

The observations were coded with an adapted version of the discipline rating scales used by Verschueren, Dossche, Marcoen, Mahieu and Bakermans-Kranenburg (2006). Adaptations included a further division of the harsh discipline scale in a physical and a verbal subscale. Discipline strategies were rated as *'harsh physical discipline'* when mothers used unnecessary physical force, either when preventing their child from touching a forbidden object or when forcing the child to clean up (i.e., slapping, grabbing the child, pulling an arm too hard, grabbing toys from the child, pinching an arm, or grabbing/holding face of the child). The action also had to cause a noticeable physical impact on the child (e.g., body movement, facial or vocal expression of shock or discomfort). Scores for harsh physical discipline were assigned using a 5-point rating scale. Because observing harsh physical discipline in relatively short episodes is rare, the scale was defined in a way that would allow more subtle as well as blatant harsh acts to be included: 1 *no physically harsh acts*, 2 *a hint of harshness, but not severe or unclear impact on child*, 3 *at least one harsh act but not as physical punishment and not used to emphasize a verbal command*, 4 *either more than one harsh act or a single act of physical punishment (e.g., slapping) or emphasizing a command (e.g., by grabbing the child's face)*, and 5 *more than one harsh act of which at least one act of physical punishment; mother has clearly lost control*.

'Harsh verbal discipline' was defined as irritation and anger in the tone of voice (i.e., impatient, irritated, angry voice, yelling, screaming.). Scores for harsh verbal discipline were also defined on a 5-point scale as follows: 1 *no harsh verbal discipline*, 2 *mild irritation*, 3 *irritation and anger*, 4 *obvious irritation and or anger on more than one occasion*, and 5 *almost constant irritation and/or anger*.

At both the 12 and 24-month home visits both children participated simultaneously with their mother in the discipline tasks (during the 18-month visit only the second child participated). Discipline strategies of the mother directed to each child were scored independently. In the current analyses only the mother's discipline strategy towards her second child was used in the analyses. Inter-coder reliabilities (intra-class correlations, single rater, absolute agreement) of six coders ranged from .76 to 1.00 for *'Harsh physical discipline'* and from .78 to 1.00 for *'Harsh verbal discipline'*. At each time point, the maximum score for each scale was used (so either the score for the 'do' or the 'don't' situation). Maximum scores from all three time points were then averaged, resulting in average harsh physical and harsh verbal discipline scores during the second year. When families missed one of three home visits, the maximum scores of the remaining two home visits were averaged instead ($n = 7$). A total score for harsh discipline was calculated by aggregating the averaged scores during the second year for both

the physical and the verbal subscale. Fifty-nine percent of the sample showed at least a hint of physical harshness during one of the observation episodes in the second year. Of those, 10 mothers showed at least one clear physically harsh act. Fifty-one percent of the sample showed at least some verbal overreactivity during one of the observation episodes in the second year. Of those, 3 mothers showed at least irritation and anger. Combining the two types of harshness, 55% of the sample showed at least some physical or verbal harshness.

Data inspection revealed a single outlier which was winsorized by adding the difference between the two next highest values and adding this difference to the next highest value (with a $z < 3.29$), thereby replacing the outlying score (Tabachnik & Fidell, 2001). A logarithmic transformation was calculated for the final total harsh discipline score to correct for positive skewness (Tabachnik & Fidell, 2001).

Externalizing behavior. The Child Behavior Checklist for ages 1½ to 5 (CBCL/1½ -5; Achenbach & Rescorla, 2000; Van Zeijl et al., 2006) was used to assess externalizing behaviors of the first-born child and was obtained from mothers via the preliminary questionnaire by mail or during the selection home-visit. Mothers indicated whether their child displayed any of the 100 behavioral descriptions in the last 2 months on a 3-point scale (0 *not true*, 1 *somewhat or sometimes true*, and 2 *very or often true*). The distribution of CBCL Externalizing Problems scores reported by Van Zeijl et al. (2006) was used to determine cut-off scores in identifying low versus high levels of externalizing behaviors. Families were selected for the low externalizing group when scores on externalizing behavior of the oldest child did not exceed the lowest 50% of scores in the norm group of the same age. Families were placed in the high externalizing group when scores were within the highest 25% of scores in the norm group of the same age. The two groups did not differ significantly on sensitivity and harsh discipline towards their second child.

Difficult temperament at 3 months. Child temperament of the second child was measured with the Infant Characteristics Questionnaire (ICQ; Bates, Freeland, & Lounsbury, 1979) completed by the mother. The ICQ was translated into Dutch and found reliable by Kohnstamm (1984). The Dutch ICQ contains 33 items, describing concrete behaviors in well-defined situations. Because in the larger project the ICQ (at some of the home visits) was used in combination with the aforementioned CBCL/1½-5, five items in the ICQ were not included in the questionnaire due to content-overlap between items of both instruments (Van Zeijl et al., 2006). Furthermore, four items were removed due to lack of applicability on young infants, resulting in a final 24-item temperament questionnaire (Van Zeijl et al., 2006). The first 22 items were rated on a 5-point scale, ranging from 0 *not true* to 4 *true*. However, item 23 (fussy/irritated behavior) was rated on a 5-point scale (ranging from 0 *never* to 5 *10 or more times a day*), while item 24 (general perceived difficultness of infant) was rated on a 7-point scale (ranging from 0 *very easy* to 7 *very difficult*). A total score for difficult temperament was calculated based on standardized scores of all 24 items. Missing values for eight participants were replaced by the total standardized mean score of the group. Internal consistency

(Cronbach's alpha) of the scale for difficult temperament was .77. Data inspection revealed a single outlier which was winsorized by adding the difference between the two next highest values and adding this difference to the next highest value (with a $z < 3.29$), thereby replacing the outlying score (Tabachnik & Fidell, 2001).

RESULTS

Maternal sensitivity and intrusiveness across time and situations

Paired *t*-tests were conducted to examine differences in means across the various time-points and measures. In the naturalistic settings maternal sensitivity was significantly lower during free play on the lap than during bathing at both 3 and 6 months. During the SFP mothers were significantly more sensitive and less intrusive during the reunion than during the baseline at 3 months but not at 6 months. Last, maternal sensitivity was significantly lower during the 6 months visit than during the 3 month visit across all settings, while maternal intrusiveness levels rose from 3 to 6 months. Means across time and situations are presented in Table 3.1.

Stability of maternal sensitivity across time

Pearson product-moment correlation coefficients for all sensitivity measures at 3 months and 6 months are presented in Table 3.1. With the exception of the Lap observations, all sensitivity measures were positively and significantly correlated over time. Especially the SFP baseline intrusiveness measure showed high stability.

To check for a possible influence of infant temperament on maternal sensitivity, the relation between all sensitivity variables and the ICQ at 3 months was examined. Except for a significant negative correlation between 6 month SFP baseline intrusiveness and difficult temperament ($r = -.30, p < .05$), no significant relations were found. Furthermore, difficult temperament at 3 months was not significantly related to the discipline outcome measure. Difficult temperament was thus not included in further analyses.

Table 3.1

Differences in means and stability of sensitivity and intrusiveness across time and situations (N = 73)

	3 months	6 months	<i>t</i>	Stability (<i>r</i>)
Sensitivity Bath	7.00 (0.83)	6.25 (1.18)	4.04**	.32**
Sensitivity Lap	6.14 (1.27)	5.35 (1.35)	5.25**	.22
Bath vs Lap <i>t</i> -value	-5.96**	-6.05**		
Sensitivity SFP baseline	2.01 (0.75)	1.64 (0.79)	3.91**	.45**
Sensitivity SFP reunion	2.21 (0.76)	1.64 (0.81)	5.35**	.35**
Baseline vs Reunion <i>t</i> -value	-2.57*	0.00		
Intrusiveness SFP baseline	1.30 (1.00)	2.01 (0.79)	-6.73**	.51**
Intrusiveness SFP reunion	0.99 (0.95)	1.96 (0.82)	-8.82**	.44**
Baseline vs Reunion <i>t</i> -value	3.72**	0.56		

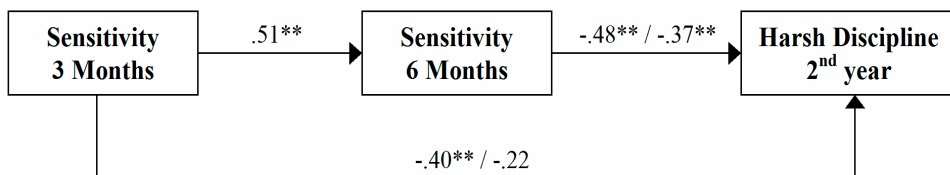
* $p < .05$, ** $p < .01$.

Principal component analysis for maternal sensitivity measures

We conducted principal component analyses for Time 1 and Time 2 separately, including all sensitivity and intrusiveness measures. On the basis of the Scree test a one-factor solution (54% and 51% of the variance explained at Times 1 and 2 respectively) was selected. Factor loadings were high, ranging from $|.46|$ to $|.88|$ at 3 months and from $|.39|$ to $|.84|$ at 6 months (factor loadings for intrusiveness were negative). Based on the factor loadings, all sensitivity scores were aggregated into one overall sensitivity variable. For each time point, standardized scores for bath, lap, and sensitivity during baseline and reunion of the SFP were added after which standardized intrusiveness scores for SFP baseline and reunion were subtracted. Cronbach's alphas for the aggregated sensitivity measures were $.79$ at 3 months and $.77$ at 6 months. Stability of the overall sensitivity factor between Time 1 and Time 2 was high ($r = .58, p < .01$).

Prediction of harsh discipline from sensitivity in infancy Correlations between sensitivity in the first year and harsh discipline in the second year were strongly negative and significant at both 3 months ($r = -.50, p < .01$) and 6 months ($r = -.57, p < .01$). Mothers who were less sensitive in interaction with their infant in the first 6 months after birth used more harsh discipline in the second year. The effect size (R^2) of the relationship between maternal sensitivity in infancy and harsh discipline in the second year was $.36$ (regression analysis with both 3-month and 6-month sensitivity predicting harsh discipline).

A mediation model was tested to investigate whether the relation between sensitivity at 3 months and harsh discipline in the second year was mediated by sensitivity at 6 months. Maternal education and age were both positively associated with maternal sensitivity (education at 3 months $r = .33, p < .01$ and 6 months $r = .39, p < .01$; age at 3 months $r = .28, p < .05$ and 6 months $r = .26, p < .05$), and negatively related to harsh discipline (education, $r = -.39, p < .01$, age $r = -.40, p < .01$). To control for the influence of maternal age and education, we computed standardized residuals for the sensitivity and harsh discipline scores. The mediation model was tested using Baron and Kenny's (1986) four conditions. Results confirmed all four conditions. First, sensitivity at 3 months significantly predicted sensitivity at 6 months, $R^2 = .26, F(1,71) = 24.33, p < .01$.



$^{**} p < .01$.

Figure 3.1. Predicting maternal harsh discipline in the second year from sensitivity in infancy ($N = 73$)

Second, sensitivity at 3 months was significantly related to harsh discipline, $R^2 = .16$, $F(1,71) = 13.77$, $p < .01$. Third, sensitivity at 6 months was also predictive of harsh discipline, $R^2 = .23$, $F(1,71) = 21.25$, $p < .01$. Fourth and last, when adding sensitivity at 6 months, sensitivity at 3 months was no longer significantly related to harsh discipline ($\beta = -.22$, $p > .05$), whereas 6-month sensitivity did remain significant, indicating full mediation (Final model: $R^2 = .27$, $F(2,70) = 12.61$, $p < .01$). This was confirmed by a Sobel test ($z = -3.38$, $p < .01$). Standardized B for each step are presented in Figure 3.1, with the standardized B after each slash presenting the values in step 4 of the test for mediation. Thus, less sensitivity at 3 months predicted less sensitivity at 6 months, which in turn predicted more use of harsh discipline in the second year.

Sensitivity in different settings in relation to harsh discipline

Pearson product-moment correlation coefficients were calculated for the relation between each of the separate sensitivity settings in infancy and harsh discipline during the second year. For this purpose composite variables were created in which scores from the 3 and 6 month home visits were aggregated for bath, lap, and the SFP baseline and reunion (sensitivity minus intrusiveness). Again, standardized residuals were used to control for the influence of maternal education and age. All correlations of sensitivity in all separate settings with harsh discipline were significant ($p < .01$) and negative (for lap $r = -.43$; bath $r = -.35$; SFP baseline $r = -.47$; SFP reunion $r = -.34$). Thus, in each of the settings lower sensitivity during infancy was related to more use of harsh discipline during the second year. To further examine the influence of intrusiveness and sensitivity in the SFP on harsh discipline, two composite variables were calculated combining baseline and reunion scores for both time points. Both these composite variables were significantly related to harsh discipline in the expected directions (i.e., positive for intrusiveness $r = .42$ and negative for sensitivity $r = -.40$). Because sensitivity and intrusiveness scores for the separate settings were highly intercorrelated and all loaded on one factor, multivariate analyses were not conducted.

DISCUSSION

The results from this study showed significant 3-month stability of maternal sensitivity within observation settings, as well as mean-level differences across different settings. Furthermore, the results showed that lower levels of maternal sensitivity at 3 as well as 6 months predicted more harsh discipline use in toddlerhood. Virtually all settings showed significant stability of maternal sensitivity from 3 to 6 months. This stability was not due to coder stability since sensitivity observations at both time points were independently coded. Existing empirical work has shown mixed results concerning the degree to which maternal sensitivity can be seen as a stable factor during infancy and later development. Several studies have found significant stability over time (e.g., Ainsworth et al., 1978; Kemppinen, Kumpulainen, Raita-Hasu, Moilanen, & Ebeling, 2006; Vereijken et al., 1997) while other studies have failed to do so (e.g., Lohaus, Keller, Ball, Voelker, & Elben, 2004). Time periods between measurements in these studies

varied from 9 months (Lohaus et al, 2004), up to almost 24 months (Kemppinen et al., 2006). Furthermore, the time of the first measurement also varied strongly from an age of 6 to 8 weeks old (Kemppinen et al., 2006) up to 14 months old (Vereijken et al., 1997). It has been suggested that part of the non-stability in maternal sensitivity may be attributed to changing levels of experienced family stress and social support (Belsky & Fearon, 2002; Pianta, Sroufe, & Egeland, 1989). However, these factors seem more likely to influence the stability of maternal sensitivity over longer time periods (up to 42 months; Pianta et al., 1989).

Besides significant stability of maternal sensitivity over time, we also found significant mean-level differences over time and between settings. First, within each setting maternal sensitivity declined significantly from 3 to 6 months. Between 3 and 6 months infants become more active participants in the dyadic interaction through, for example, the development of laughter and the social smile (Sroufe, 1995). Thus, there are more signals for a mother to pick up on and adequately respond to, and therefore also more signals that can be missed; possibly resulting in lower sensitivity scores. For the SFP, we also found a significant increase in intrusiveness over time during the baseline interaction and the reunion. An increase in intrusiveness over time may also be linked to the development of infant laughter, 'encouraging' mothers to engage their infants in more intrusive game playing. Thus, the drop in maternal sensitivity scores over time may be partly due to increasing intrusiveness.

Significant differences were also found across the various settings for the observation of maternal sensitivity within the same home visit, which could suggest that some settings create circumstances in which it is easier for mothers to respond sensitively, while other settings prove to be more challenging. Our results showed, for example, that mothers were more sensitive during bathing than during free play at both time points. This could be explained by the degree of routine and practice mothers have within certain settings and the accompanying infant signals. Bathing is a daily and therefore well-practiced routine for most mothers, which may not necessarily be the case for free play on mother's lap. For example, during home visits several mothers mentioned how special and unusual it was for them to have their baby on their lap just for free play. In addition, free play without toys is less defined by a clear-cut task than bathing, which could prove to be more difficult for some mothers. Furthermore, infants will be increasingly likely to look around for exploration due to the development of gaze shifting (Hunnius & Geuze, 2004). This may result in less undivided attention for the mother's attempts to establish interaction, which could be especially challenging for mothers during free play. Therefore, a possible explanation for lower sensitivity scores during free play as compared to bathing could be that mothers try to force a playful interaction (possibly by means of intrusive game playing) instead of responding adequately to the infants attempts to explore.

A second between-setting mean-level difference was found for the baseline and reunion of the SFP at 3 months. Mothers were more sensitive and less intrusive during the reunion than during the baseline interaction of the SFP. This could be caused by higher levels of infant negative affect during the reunion, which is carried over from the still face segment (Mesman et al., 2009). Infant

crying provides a clear signal for the mother and can hardly be missed in a face-to-face setting, which makes it easier to perceive the signal and to respond to it with at least some sort of soothing behavior. In contrast, during the baseline interaction, infants generally show higher levels of positive affect and lower levels of negative affect as compared to the reunion episode (Mesman et al., 2009). Therefore, during the baseline interaction mothers do not necessarily have clear signals (such as infant distress) to respond to and might have to be attuned to more subtle infant signals (i.e., of exploration). However, it remains unclear why we did not find these mean level differences for sensitivity and intrusiveness between baseline and reunion at 6 months. According to a recent meta analysis, this result can not be attributed to a less pronounced increase in negative affect during the SFP, since no age differences were found for the 'carry-over effect' of negative affect from still face to reunion (Mesman et al., 2009).

Results of the factor-analysis showed that sensitivity in all settings had high factor loadings on a single factor at both time points. This supports the idea that in the various sensitivity settings the same underlying construct of sensitive parenting is measured. Also, both sensitivity and intrusiveness loaded on the same factor, suggesting a strong overlap between these constructs. However, this does not eliminate the potential importance of separate subscales for the measurement of maternal sensitivity. For instance, even though higher levels of intrusiveness generally lead to lower scores on maternal sensitivity ratings, this does not mean that a low score on maternal sensitivity automatically indicates high levels of intrusiveness. Lower scores on maternal sensitivity can also be due to an overall absence of responses to infant signals, which would indicate low intrusiveness. Distinguishing between insensitive-passive mothers and insensitive-intrusive mothers may be highly relevant for specific processes and outcomes.

As a second aim we investigated whether sensitivity during the first 6 months would predict harsh discipline use in toddlerhood. Results showed that maternal sensitivity at 3 months predicted harsh discipline use in the second year. This relationship was completely mediated by maternal sensitivity at 6 months. Mothers who were less sensitive at 3 months were also less sensitive at 6 months, which predicted more harsh discipline use during the second year. This finding extends an earlier finding by Engfer and Gavranidou (1987), which showed that maternal sensitivity on the maternity ward directly after birth was a significant predictor of ideas and attitudes on punishment use at 4 and 18 months after birth. To our knowledge there have been no previous longitudinal studies linking extensive observations of maternal sensitivity in infancy directly to observations of harsh discipline in toddlerhood.

Theoretical definitions of maternal sensitivity can be fit easily into the social information processing model proposed by Milner (1993, 2003), which is often referred to in the harsh discipline and child abuse literature. The various stages in the social information processing model all define a lack of parenting skills and behaviors that are central to the concept of maternal sensitivity. One aspect of maternal sensitivity that has been specifically linked to harsh discipline use is the degree of maternal intrusiveness during interaction (Lyons-Ruth et al., 1987). Maternal intrusiveness as well as the use of harsh discipline seems

to be the result of a lack of respect and/or empathy for the child's autonomy and behavioral motivations. Parents who do not perceive or even ignore their infants' signals and interfere (physically) with their infants' behavior could also be more likely to misread their children's motivations during toddlerhood due to a lack of empathy and child-centered attributions; leading in turn to harsh discipline (Dadds, Mullins, McAllister & Atkinson, 2003; Dopke & Milner, 2000; Milner 1993, 2003; Milner, Halsey, & Fultz, 1995). Nonetheless, in our results the relation between intrusiveness and harsh discipline was not stronger than the relation between the other sensitivity measures and harsh discipline. Therefore, the results point to a more general lack of perspective-taking which hampers the capability for maternal sensitivity and nonintrusive responsiveness in the first months after birth and leads to the use of harsh discipline in toddlerhood. Since maternal sensitivity can be relatively easily observed during the first months after birth, this would provide excellent opportunities for early identification of those mothers who are possibly at risk for using harsh discipline in toddlerhood. This in turn would give room for timely interventions to prevent the development of more negative dyadic interactions.

The current study has several limitations. First, a different measure was used to rate maternal sensitivity and intrusiveness in the SFP as compared to the bath and free play sessions. Contrary to the SFP, there were no separate subscales for sensitivity and intrusiveness in the bath and free play sessions, which may have limited the comparability of the different settings. Nevertheless, a single underlying factor was shown for all included sensitivity measures at both time points. Second, within time points sensitivity during bath and free play were rated by the same coder, as were the baseline and reunion of the SFP. Therefore, it is possible that mean levels of sensitivity within time points were influenced by same-coder stability. Furthermore, significant mean-level differences in maternal sensitivity at a certain age could be less reliable due to the use of one coder for both settings. However, same coder-stability would have sooner resulted in a lack of significant mean-level differences between the settings. This was not the case; we found significant mean-level differences between bath and free play sessions as well as between the baseline and reunion of the SFP within time points. Last, we did not find a significant relation between infant difficult temperament and maternal sensitivity, which was found in previous studies (Kivijärvi, Räihä, Kaljonen, Tamminen & Piha, 2005; Pederson et al., 1990; Seifer, Schiller, Sameroff, Resnick & Riordan, 1996). This could have been due to the use of questionnaires filled out by the mothers, since it has been shown that observed infant temperament is not always strongly related to mother-reported infant temperament (Seifer, Sameroff, Barrett & Krafchuk, 1994).

The current study is the first to show that maternal sensitivity in various observation settings refers to a single underlying construct. Nevertheless, sensitivity as observed in specific settings may be more predictive of specific parenting or child outcomes. It may therefore be advantageous to include multiple observation settings. In addition, the assessment of paternal sensitivity may require different observation settings than in studies focusing only on mothers (Feldman, Gordon, Schneiderman, Weissman & Zagoory-Sharon, 2010;

Grossmann, Grossman, Kindler & Zimmermann, 2008; see for a meta-analysis Lucassen et al., 2011).

Our study is also the first to show that a lack of sensitivity in infancy predicts harsh discipline in toddlerhood. We have also highlighted the similarities between the social information processing model of abusive parenting and the sensitivity hypothesis in attachment theory. The two models not only converge in terms of their description of the process of response selection, but also in their views on the origins of such processes. Both models refer to beliefs about parenting that have been formed before the child has even been born and which are based at least partly on the parents' childhood experiences with their own caregivers. And in both models, it is not the childhood experiences themselves that cause insensitive or abusive parenting, but rather the extent to which such experiences have been resolved or remain salient and/or traumatic (Milner et al., 2010; Hesse, 2008). Thus the two models appear to be rather similar in their explanation of insensitive and abusive parenting, even though they come from quite different theoretical backgrounds.

Despite remarkable similarities in the theoretical backgrounds of both research areas, there are few studies exploring the nature of the link between early lack of sensitivity and later harsh discipline. Future research should also identify protective factors that break this pattern of maladaptive parenting. Finally, early identification of families at risk for developing negative parenting strategies, such as harsh discipline, could provide mental health care professionals with more time for efficient intervention strategies (Heckman, 2006), thereby improving the prognosis of young infants at risk.

