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

Joosen, K.J.

4 Predicting harsh discipline of infants: Overreactive sympathetic nervous system responses to repeated infant crying predict maternal harsh discipline

Katharina J. Joosen, Judi Mesman, Marian J. Bakermans-Kranenburg, Marinus H. van IJzendoorn
Manuscript submitted for publication

Abstract

Physiological reactivity to repeated infant crying was examined as a predictor of harsh discipline use with 12-month olds in a longitudinal study with 48 mother-infant dyads. Physiological reactivity was measured while mothers listened to three blocks of infant cry sounds in a standard cry paradigm when their infants were three months old. Harsh discipline use was observed in two tasks during a home visit when the infants were 12 months old. Mothers who used harsh discipline with their 12-month-olds were compared to mothers who did not on their sympathetic and parasympathetic reactivity to the presented cry sounds. Results showed a significant interaction effect for sympathetic reactivity only. Mean Skin Conductance Levels (SCL) of harsh mothers continued to rise across the repeated exposures to cry sounds, while after an initial increase the mean SCL of the other mothers showed a steady decline. Results are discussed from a polyvagal perspective, and we suggest that harsh parenting reflects behavioral as well as physiological overreactivity to negative infant signals.

Key words: harsh discipline, physiological reactivity, infant crying, polyvagal theory

Introduction

Parental harsh discipline plays an important role in the development of problem behavior in childhood and adolescence (e.g., Bender et al., 2007; Miner & Clarke-Stewart, 2008; Prinzie, Onghena & Hellinckx, 2006). Harsh parenting can be seen as an overreactive strategy used by parents with negative, blame-oriented, and child-centered attributions about their children's (mis)behavior (Bugental & Happeney, 2004; Bugental, Lewis, Lin, Lyon, & Kopeikin, 1999). On a biological level, harsh parenting has been related to physiological hyperreactivity to (negative) infant signals (for a review see McCanne & Hagstrom, 1996). More specifically, studies have shown a stronger activation of the sympathetic nervous system (SNS) in response to infant crying among harsh or abusive parents when compared to control groups (Crowe & Zeskind, 1992; Disbrow & Caulfield, 1977; Frodi & Lamb, 1980), which may indicate that harsh parents perceive their children's behavior as more threatening (Martorell & Bugental, 2006). Porges' polyvagal theory (1995, 2001, 2011) suggests that the perception of negative infant signals as threatening, leads harsh parents to rely more on physiological defensive strategies supported by the SNS than on the functioning of the vagal brake supported by the parasympathetic nervous system (PNS). In this longitudinal study we examined sympathetic and parasympathetic activation in response to infant crying in relation to observed harsh parenting.

In the existing literature on harsh parenting, or more specifically harsh discipline, ample evidence can be found of the detrimental effects of this parenting strategy on the behavioral and emotional development of children and adolescents (for a meta-analysis see Gershoff, 2002). Children of parents who use harsh discipline have higher levels of externalizing behavior problems in childhood (e.g., Bender et al., 2007; Fine, Trentacosta, Izard, Mostow, & Campbell, 2004; O'Leary, Slep & Reid, 1999; Prinzie, Onghena, & Hellinckx, 2006), are more likely to become delinquent (Hoeve et al., 2009) and generally have lower levels of mental health (Gershoff, 2002) compared to children of non-harsh parents. Nonetheless, a controversy remains as to whether these negative consequences are related to the use of all forms of harsh or physical discipline including spanking, which has sometimes been labeled as 'discipline with reasonable force' (Baumrind, Larzelere, & Cowan, 2002).

Most empirical work on harsh discipline uses definitions including multiple disciplining or punishing behaviors. Because measures of harsh discipline generally encompass a multitude of different behaviors it is plausible that part of the explanation of the divergence in findings for effects of harsh discipline lies in the different and only partly overlapping definitions used across studies. The various harsh discipline strategies can be seen on a continuum ranging from subtle forms such as grabbing hard, to forms of corporal punishment in the middle of the continuum (such as hitting with a belt or paddle), to more extreme forms of physical punishment (such as beating or kicking causing injuries which could result in hospitalization) at the other end of this harsh discipline continuum. Thus, comparing results on child outcomes of harsh discipline is problematic when some studies include only part of the continuum (e.g., spanking) in their definitions, while other studies include the whole or a different range of harsh discipline behaviors.

Besides this continuum based solely on physical aspects of harsh discipline, the broader concept of overreactive discipline which includes parenting behaviors such as yelling, threatening, commanding, and name-calling has been used in several studies (e.g. Bugental & Happaney, 2004; Chang, Schwartz, Dodge, & McBride-Chang, 2003; Fine, Trentacosta, Izard, Mostow, & Campbell, 2004; Leung & Slep, 2006; Straus & Stewart, 1999). Such overreactive parenting behaviors are central to Patterson's coercion theory. Within this theoretical framework, parents respond with escalating coercive discipline strategies such as physical and verbal harshness when faced with challenging child behavior (Patterson, 1982). There is substantial evidence that this pattern of physically and verbally harsh discipline is related to negative child outcomes through processes such as modeling and lack of positive reinforcement (e.g., Patterson, Capaldi, & Bank, 1991; Shaw, Bell,

& Gilliom, 2000). Following coercion theory and the idea of escalating discipline in multiple modalities, both physical (e.g., spanking, grabbing, pushing, and pulling) and verbal (e.g., yelling, name calling, and aggressive/irritated tone of voice) discipline strategies are included as aspects of harsh discipline in the current study.

In characterizing parents who are (at risk for) using harsh or even abusive discipline, Milner (1993, 2003) presented a four-stage model of social information processing. Across these four stages high-risk and abusive parents are described to be (1) more biased in the perception of, as well as less attentive to, their children's behavior, (2) more likely to interpret negative child behaviors as motivated by hostile intent, and to have child-centered attributions (i.e., to view this behavior as due to internal, stable, and global child factors), (3) less likely to use situational cues when evaluating their children's behavior, and (4) less likely to have well-developed skills or abilities to implement and modify child-directed responses and parenting strategies. These processes are also relevant in the study of harsh discipline, since several studies have shown significant relations between harsh discipline and child-centered, or child-blaming attributions as included in Milner's model for child abuse. For example, mothers who rated videotapes of children's negative affect in mother-infant interactions higher on child-blaming attribution also scored higher on harsh parenting (Smith & O'Leary, 1995).

From a biological perspective, individual differences in parental physiological reactivity to negative infant stimuli have also been related to differences in the quality of parenting. For example, greater heart rate (HR) reactivity to infant crying has been shown to distinguish parents at risk for abuse from controls (for a review see McCanne & Hagstrom, 1996). Furthermore, there has been empirical support for a direct relation between HR reactivity and overreactive discipline. One study observed mothers with their toddlers in two discipline tasks in a lab setting, during which HR reactivity of the mothers was measured. Overreactive discipline was measured as the degree and frequency of anger or irritation through vocal (e.g., yelling) or non-vocal (e.g., grabbing) channels during a clean-up task and a competing-demand task. Results showed that a greater HR reactivity was associated with more overreactive discipline strategies (Lorber & O'Leary, 2005). However, greater HR reactivity in response to infant crying has also been found in highly sensitive mothers (Joosen, Mesman, Bakermans-Kranenburg, Pieper, Zeskind, & Van IJzendoorn, 2012) and in mothers with a prompt response to infant crying (Del Vecchio, Walter, & O'Leary, 2009). The finding that greater HR reactivity seems to be related to both harsh and sensitive parenting may be due to the fact that activation of both the sympathetic and the parasympathetic branch of the autonomic nervous system (ANS) can lead to increases in HR (see Berntson, Cacioppo, Quigley, & Fabro, 1994). These dual mechanisms influencing HR create a problem for the interpretation of HR reactivity.

To examine the underlying mechanisms of physiological reactivity, Porges' polyvagal theory (1995, 2001, 2007, 2011) describes three subsystems of the ANS, namely the social communication circuit (involving the myelinated vagus and the PNS), the mobilization circuit (involving the SNS and fight-flight behaviors), and the most primitive circuit of immobilization (involving the unmyelinated vagus and e.g., feigning death or freezing behaviors). These three subsystems developed in three phylogenetic stages and are, therefore, hierarchically activated when responding to external stressors. The activation of a specific subsystem also depends on the perception of the environment as either safe or threatening. In a safe environment the PNS circuit for social communications is activated to promote survival by facilitating social interactions and social bonds. However, when the environment is perceived as dangerous or threatening the autonomic system switches to the more primitive circuits and activates the SNS to regulate defensive strategies through a fight-or-flight response.

Since harsh parents, given their negative and child-blaming attributions, seem to operate from a threat rather than a safety perspective, they may be expected to show physiological overreactivity of especially the SNS in response to infant crying. Activation of the SNS can be measured by an individual's electrodermal activity, operationalized as changes in SCL. Whereas HR is influenced by both branches of the ANS, electrodermal activity is exclusively regulated by the SNS (Boucsein, 1992). When the SNS is activated, sweat in the eccrine sweat glands accumulates, causing rising sweat levels in the sweat ducts (of the palmar surfaces), resulting in higher levels of skin conductance (Boucsein, 1992; Dawson, Schell, & Filion, 2000). An early study of physiological reactivity to infant signals by Frodi and Lamb (1980) indeed showed a greater increase in SCL for abusive mothers as compared to control mothers in response to infant crying. Another study including both males and females with high and low scores on the Child Abuse Potential Inventory confirmed this finding, by showing greater SCL reactivity amongst the high-Child Abuse Potential group for phonated cry sounds specifically (Crowe & Zeskind, 1992).

Greater HR reactivity in sensitive mothers has been suggested to result from activation of the PNS and not from activation of the SNS (Joosen, Mesman, Bakermans-Kranenburg, Pieper et al., 2012). The PNS, depends on the functioning of the ventral vagal complex (VVC), which originates in the nucleus ambiguous. The VVC functions as an active vagal brake by controlling cardiac output via the sino-atrial node (Porges, 2001, 2007). The degree of cardiac control by the vagal brake can be quantified by the amplitude of respiratory sinus arrhythmia (RSA; Porges, 1995). Efficient disinhibition of the vagal brake (i.e., RSA withdrawal) seems to be associated with increased behavioral and emotional control, which enables rapid mobilization as well as calm and self-soothing behavioral states in response to environmental demands (Porges, 1996). In other words, an efficiently functioning vagal brake points to physiological flexibility when reacting to environmental demands. In contrast, the negative behavioral overreactivity to infant signals seen in parents at risk for abuse (Milner, 2003) may signal a chronically overactivated SNS resulting from a *poorly* functioning vagal brake (Porges, 2001). Thus, the difference between those parents using harsh discipline and those who do not, would be the combination of an overreactive SNS with an unresponsive PNS.

To effectively distinguish these mechanisms in relation to parenting behavior, sympathetic *and* parasympathetic activation in reaction to stressors such as infant crying should be measured simultaneously to unravel the unique effects of the two branches and to identify meaningful differences between them. In the current longitudinal study we tested the hypothesis that parents with a stronger activation of the SNS in response to standardized infant cry sounds (measured when their infants were 3 months old with a cry paradigm) would be more likely to use harsh discipline with their infants at 12 months. We expect mothers who will show harsh discipline to have a greater increase in SCL in response to infant cry sounds when compared to mothers not showing harsh discipline. Concerning PNS reactivity to infant cry sounds, we expect mothers who will show harsh discipline to have significantly less RSA withdrawal than the other mothers in response to the cry sounds.

Method

Sample and procedure

Participants were recruited via the Regional Coordination Programs of the Dutch National Institute for Public Health and Environment (NIPHE). Inclusion criteria were (1) both parents were low to middle educated, (2) the first child was younger than six years, (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, and (6) 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). We added the final criterion in order to increase the chance of including both low- and high-risk mother-infant dyads very early without relying on more distal risk factors such as parental education level or stress.

Since the focus of the recruitment was to approach families from the lower socioeconomic strata, we pre-selected postal codes 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 details and were then contacted by phone to find out the educational level of both parents (1 *elementary school*, 2 *first 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.

A total of 53 families met our criteria (26 with a 'high externalizing' first child), of whom 48 families were included in this paper (three families dropped out during the first 12 months and two mothers had unreliable ECG data). Within the final sample 13% of mothers finished elementary school, 60% finished the first four years of Dutch high school, and 27% obtained a degree in vocational education or finished five or six years of high school. Maternal mean age at time of birth of their second child was 29.65 years (*SD*: 4.93, range 19-39 years). The mean age of the first children at the time of selection was 35.42 months (*SD*: 12.65, range 18-70). Of the first-born children 54% were boys, and of the second-born children 35% were boys.

All 48 families participated in a total of four home visits during the first year after the birth of the second child. Home visits were scheduled when the second child was 3, 6, 9, and 12 months old. For the current study, we used the assessments at 3 and 12 months. At the 3-month home visit mothers participated in a 20-min cry paradigm (Out, Pieper, Bakermans-Kranenburg, & Van IJzendoorn, 2010; Zeskind & Lester, 1978) during which an electrocardiogram (ECG) signal as well as the SCL were recorded. Prior to the cry paradigm mothers were asked to fill out a short questionnaire concerning possible smoking habits and their frequency of physical exercise to control for any influence on the ECG data (De Geus, Boomsma, & Snieder, 2003; Farrington, 1997; Vander, Sherman, & Luciano, 2001). At the 12-month home visit mothers were observed with both children simultaneously during two 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. Procedures and measures used in this study were reviewed and approved by the ethical committee of the Institute of Education and Child Studies at Leiden University.

Measures

Externalizing behavior. The Child Behavior Checklist for ages 1½ to 5 (CBCL/ 1½ -5; Achenbach & Rescorla, 2000; Van Zeijl et al., 2006) was used in the screening phase to assess externalizing behaviors of the first-born child and was obtained from mothers 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.

Cry Paradigm. The cry perception stimuli and paradigm (Out et al., 2010) were administered using a laptop with E-prime software. Mothers listened to three blocks of each three cry sounds that varied in fundamental frequency within each block. Cry stimuli were derived from the spontaneous crying of a healthy 2-day old, full birth-weight and full term female infant and were recorded midway between scheduled feedings. A 10-sec portion of the sustained period of crying, containing seven expiratory sounds, was selected for presentation. The seven cry expirations had a mean duration of 1055 msec (range: 0.6195 to 1899 msec) and a mean Peak F_0 of 452.6 Hz (range 425.2 to 515.6 Hz).

To provide cry stimuli with a wide range of fundamental frequencies, the original cry (averaging approximately 500 Hz) was digitally altered to increase the fundamental frequency while holding temporal and other spectral aspects of the cry constant. Two new 10-sec cry stimuli were created by digitally increasing the original cry by approximately 200 and 400 Hz, respectively, resulting in two new cry sounds with an overall Peak F_0 of 714.5 Hz (700 Hz Cry) and 895.8 Hz (900 Hz Cry; see Out et al., 2010). Each of the three blocks of cry sounds contained each of the three cry sounds (i.e., 500 Hz, 700 Hz, 900 Hz) presented in a random order.

The cry stimuli were presented at a constant volume through Sennheiser HD202 headphones. Following a 4-min baseline period and a practice trial in which the 500 Hz cry was presented, mothers listened to the three blocks of three cry sounds each. The paradigm was concluded with a 2-min recovery period. Following every presented cry sound mothers answered four questions on perceived urgency (not aroused – aroused, not aversive – aversive, healthy – sick and not urgent – urgent) on a 5-point rating scale (Zeskind & Lester, 1978; Zeskind & Marshall, 1988). Given the high internal consistencies (alphas ranging from .77 to .82), answers to the four perceived urgency questions were averaged to form an overall perceived urgency score for each of the three frequencies. Missing values for one participant (2%) were replaced with the standardized mean scores of the group for the overall perceived urgency per frequency. During the entire cry paradigm mothers' ECG and SCL signals were recorded continuously.

RSA and SCL. ECG and SCL signals were measured with an ambulatory monitoring system (VU-AMS5fs; TD-FPP, Vrije Universiteit, Amsterdam, the Netherlands). For the ECG signal, three disposable pre-gelled Ag-AgCl electrodes (ConMed, New York, USA) were used, which were placed below the right collar bone 4 cm to the right of the sternum, 4 cm under the left nipple and at the lateral right side. The full ECG signal was stored at a 16-bit sampling rate. SCL was measured with two Ag-AgCl electrodes, which were filled with Biopac isotonic paste. After washing both hands with a mild soap solution, the electrodes were attached to the volar surfaces of the medial phalanges of the middle and index finger of the nondominant hand. The SCL signal was stored on the VU-AMS5fs at 10 samples per second (10 Hz). Both physiological responses were synchronized to the cry sounds using a marker button on the AMS device. The experimenter pushed the button two seconds before the stimulus was presented, leaving markers that allowed for accurate labeling of each cry sound.

Inter beat interval time series (IBIs) were derived by visual peak-detection of the R-wave through accompanying VU-AMS5fs software packages. Each recorded ECG complex was inspected and corrected by hand when necessary. RSA was measured through calculation of the root mean square of successive differences (RMSSD) of inter beat intervals for each of the labeled segments (baseline, recovery and each of the cry presentations). Mean SCL was also calculated for each of these labeled segments.

For mean RMSSD, missing values for two participants were replaced with the mean scores for each labeled segment. For mean SCL, missing values for four participants were replaced with the mean scores for each labeled segment. Outliers were only detected for mean RMSSD and were winsorized for three participants to values corresponding to a standardized value of 3.29, whilst preserving the participant's response pattern. Mean SCL and RMSSD levels were aggregated within each of the three blocks (i.e., average of three consecutive episodes of 10 seconds).

Smoking and Physical exercise. Mothers were asked to fill out a short questionnaire prior to the cry paradigm in which they answered whether and how much they had smoked prior to the home visit that day (0 *no*, 1 *yes one or two*, and 2 *more than two*). Furthermore, mothers answered how often they had engaged in physical exercise in the week prior to the home visit (0 *no*, 1 *yes once or twice*, and 2 *more than twice*). For both smoking as well as physical exercise, scores were dichotomized creating a smoking (scores of 1 and 2) and a non-smoking (score of 0) group as well as an exercise and non-exercise group. Missing values for one participant (2%) were replaced with the modus for both variables. Groupings for smoking and physical exercise are presented in Table 4.1.

Maternal sensitivity. Maternal sensitivity was assessed with the Ainsworth Sensitivity rating 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 the 3-month home visit and only involved the mother with her second-born infant. Arrangement were made to ensure that the first-borns

Table 4.1.
Background and physiological data of harsh (n = 10) and non-harsh mothers (n = 38)

	Observed harsh discipline						
	Harsh mothers			Non-harsh mothers			
	n (%)	*M*	*SD*	*n* (%)	*M*	*SD*	*F*
Smoking during day of homevisit (yes)	5 (10.4)			8 (16.7)			
Physical excercise during week prior to homevisit (yes)	4 (8.3)			24 (50.0)			
Maternal sensitivity at 3 months		6.23	0.84		6.51	1.00	
Age in years		30.47	5.54		29.44	4.81	
Mean baseline SCL		9.41	3.11		9.66	3.07	
Mean SCL during block 1		11.11	3.70		10.85	3.64	1.50
Mean SCL during block 2		11.32	4.10		10.66	3.67	4.96*
Mean SCL during block 3		11.39	3.90		10.57	3.56	5.70*
Mean SCL during recovery		11.06	3.73		9.96	3.10	13.46**
Mean baseline RMSSD		35.22	14.98		51.64	25.73	
Mean RMSSD during block 1		35.96	15.07		45.50	21.88	0.14
Mean RMSSD during block 2		38.15	16.40		46.06	20.23	0.82
Mean RMSSD during block 3		34.28	14.95		45.91	21.74	0.02
Mean RMSSD during recovery		33.69	18.31		47.67	23.01	0.03
Total	10 (20.8)			38 (79.2)			

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

were not at home, or would play with a research assistant in a separate room. All observations were rated independently by two trained coders, but bath and lap sessions were rated by the same coder. Intercoder 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 by 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. One missing value was replaced with the overall mean of the group and one outlying value was winsorized to a value corresponding to a standardized value of -3.29.

Harsh discipline. Maternal discipline was observed during the home visit at 12 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 children and to make sure neither child would touch the toys. After 2 minutes the children were allowed to play only with the least attractive toy: a simple stuffed animal. After another 2 minutes the children were 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 both children. She could help and encourage the children as she would normally do, but was told that it was important that the children would do as much of the cleaning up as possible. The clean-up session was finished when all toys were back in the bag or when ended by the experimenter after 10 minutes.

The observations were coded with an adapted version of the discipline rating scales used by Verschueren et al. (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 the 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 rated on 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 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., grabbing 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 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.* However, *'harsh verbal*

discipline' scores showed too little variance at 12 months to include in further analyses, thus only *harsh physical discipline' scores* are included in this study.

At the 12-month home visit both children participated simultaneously with their mother in the discipline tasks. Discipline strategies of the mother directed to each child were scored independently. In the current analyses only the mother's harsh discipline score towards her second child was used. Intercoder reliabilities of six coders ranged from .76 to 1.00 for *'harsh physical discipline'*. The highest score of the two situations observed (either of the don't-touch task or of the clean-up task) was used as indicator of the severity of maternal harsh physical discipline. Using this score we created two groups based on a cut-off score of 3. This cutoff distinguished mothers who performed at least one act of clear physical harsh discipline (n = 10) from mothers who did not (n = 38).

Results

Preliminary analyses

Sample distributions, means and standard deviations are presented in Table 4.1. Mothers in the harsh discipline group did not differ significantly from the non-harsh group of mothers on smoking, physical exercise, or age (all ps > .05). Smoking, physical exercise, and maternal age neither had a significant influence on baseline levels of RMSSD, nor were these variables significant covariates in any of the repeated measures analyses for RMSSD. Therefore, analyses are reported without these variables. All repeated measures analyses included a baseline value for SCL or RMSSD, respectively, as the reference point for reactivity and the Greenhouse-Geisser Epsilon was used to correct for violation of sphericity.

Perception of urgency.

A repeated measures analysis for perceived urgency of the three cry sounds (500, 700 and 900 Hz) showed that all mothers perceived the 700 as well as the 900 Hz cry sound as significantly more urgent than the 500 Hz cry sound, $F(1.75, 80.49) = 48.15$, $p < .01$, partial $\eta^2 = .51$. No significant differences in perceived urgency were found between the 700 and 900 Hz cry sound. Mothers in both groups also did not differ significantly in their ratings of perceived urgency for all three cry frequencies.

SCL

Mean SCL response patterns for both groups are presented in Table 4.1 and Figure 4.1. Separate analyses of covariance showed higher mean SCL for the harsh group of mothers compared to the non-harsh group in block 2 ($F(1,45) = 4.96$, $p < .05$, partial $\eta^2 = .10$), block 3 ($F(1,45) = 5.70$, $p < .05$, partial $\eta^2 = .11$), and the recovery ($F(1,45) = 13{,}46$, $p < .01$, partial $\eta^2 = .23$) when corrected for baseline SCL values. A repeated measures analysis showed a significant rise in mean SCL from baseline to the three cry sound blocks ($F(2.41, 111.04) = 23.57$, $p < .01$, partial $\eta^2 = .34$) as well as a significant decrease in mean SCL from block 3 to recovery ($F(1, 46) = 10.29$, $p < .01$, partial $\eta^2 = .18$). A significant interaction effect between SCL reactivity over time and harsh discipline was found, $F(2.41, 111.04) = 4.10$, $p < .05$, partial η^2 =

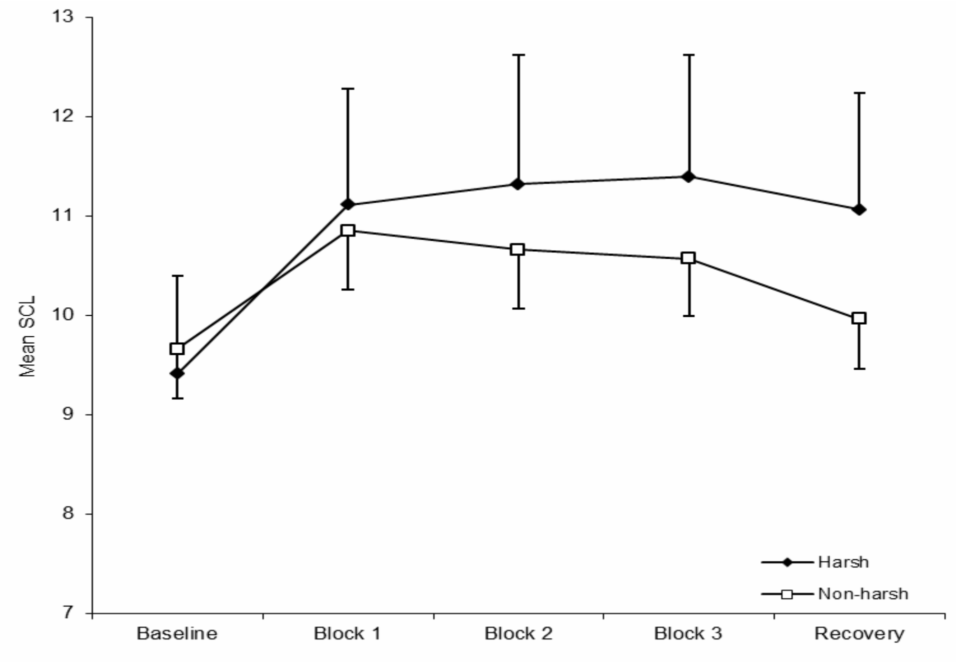

Figure 4.1. SCL reactivity to infant crying *(M, SE)* in harsh (*n* = 10) and non-harsh (*n* = 38) groups of mothers

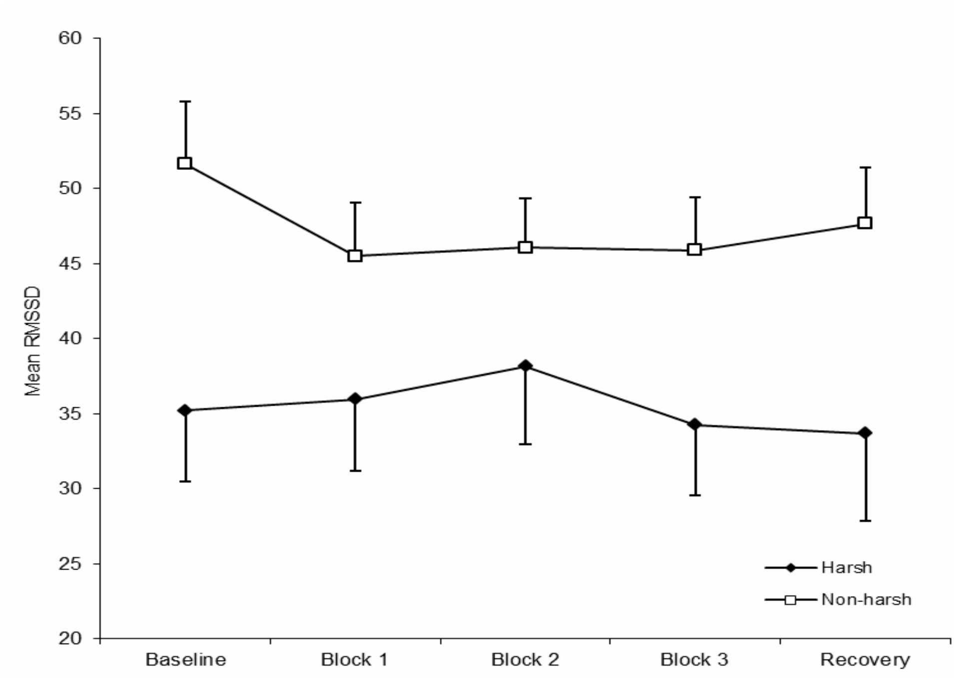

Figure 4.2. RSA in response to infant crying (*M, SE*) in harsh (*n* =10) and non-harsh groups of mothers (*n* = 38)

.08. Mothers in the harsh group showed a significantly different response pattern of SCL than the non-harsh group from baseline to block 3 ($F(1,46) = 5.22$, $p < .05$, partial $\eta^2 = .10$) and from baseline to the recovery ($F(1,46) = 13.79$ $p < .01$, partial $\eta^2 = .23$). While the mean SCL of the mothers in the non-harsh group declined after the first block of crying, harsh mothers' mean SCL continued to rise across the cry paradigm, resulting in significantly different response patterns from baseline to block 3 and remaining differences during recovery. Maternal sensitivity at 3 months was not a significant covariate when added to the analyses, and when controlling for 3-month sensitivity the interaction between SCL reactivity over time and harsh discipline remained significant, $F(2.38, 107.13) = 3.51$, $p < .05$, partial $\eta^2 = .07$.

RSA

Mean RMSSD response patterns for both groups are presented in Table 4.1 and Figure 4.2. Separate analyses of covariance for each block and the recovery did not show any significant differences between means for the harsh and non-harsh group when correcting for baselines values (test-statistics are presented in Table 4.1). A repeated measures analysis for RMSSD showed no significant main effect of RMSSD reactivity over time ($F(3.04, 139.68) = 0.87$, $p > .05$, partial $\eta^2 = .02$), nor an interaction effect between RMSSD reactivity over time and harsh discipline, $F(3.04,139.68) = 1.38$, $p > .05$, partial $\eta^2 = .03$. Thus mothers in the harsh group did not show a different RMSSD response pattern to the cry sounds compared to the other mothers. Maternal sensitivity at 3 months was not a significant covariate in the repeated measures analysis, $F(3.08,138.51) = 1.27$, $p > .05$, partial $\eta^2 = .03$.

Discussion

This study investigated differences in physiological reactivity to cry sounds between mothers observed to use harsh discipline and mothers who were not. Specifically, we hypothesized that the mothers using harsh discipline would have a stronger reactivity of the SNS in response to cry sounds compared to other mothers. Furthermore, we expected mothers using harsh discipline to have a weaker RSA withdrawal compared to other mothers in response to cry sounds, due to the absence of a well-functioning vagal brake (Porges, 2001). Our results confirmed that mothers in the harsh discipline group showed more reactivity of the SNS, as shown by greater increases in SCL compared to mothers in the non-harsh group, especially after repeated exposure to cry sounds across the paradigm. Mean SCL of mothers using harsh discipline continued to rise across the cry paradigm, while the mean SCL of mothers not using harsh discipline declined after an initial increase in response to the first block of cry sounds. Where there was no significant difference in mean SCL baselines between the two groups, mothers in the harsh group continued to show increased SCL leading to significant mean level differences between groups from the second block of crying onwards.

Nonetheless, *both* groups did show an initial activation of the SNS in response to the first block of cry sounds, shown by an increase in mean SCL from baseline to block 1 for all mothers. Furthermore, group differences in mean SCL were not yet significant during the first block of cry presentations as shown by the analysis of covariance for block 1. It thus seems that not the immediate activation of the SNS in response to infant crying differentiates between the two groups of mothers, but that only the enduring activation of the SNS in response to repeated cry sounds is linked to later observations of harsh parenting. These results are in line with results from a study that found a later peak for Shaken Baby Syndrome hospitalizations (10-13 weeks of age) than for the peak of the crying curve (5-6 weeks of age). This delay in peak for Shaken Baby Syndrome hospitalization could be due to the stress that is caused for some (at-risk) parents when they are confronted with persistent high crying levels that endure beyond the peak of the normal crying curve (Barr, Trent, & Cross, 2006). Thus, it seems that especially repeated exposure to crying is related to the experience of psychological stress in parents using harsh discipline, which in combination with negative, blame-oriented, and child-centered attributions (Bugental & Happeney, 2004; Bugental et al., 1999), leads to unremitting activation of their SNS.

Concerning our hypothesis on RSA withdrawal, results did not confirm a weaker RSA withdrawal in response to the cry sounds for the mothers using harsh discipline compared to the mothers who did not. A possible explanation for this unexpected finding could be the use of RMSSD as a quantification of RSA. In the present study, very short segments were used for the calculation of RMSSD: per block we averaged three segments of 10 seconds (the duration of each cry sound), each divided by a resting period of one minute. Unfortunately, RSA estimates in general have been found to be less reliable when measurement intervals are short (Berntson et al., 1997). Furthermore, RMSSD as a measure of RSA is known to have a large range of acceptable scores (see manual VU-AMS.5fs on http://www.vu-ams.nl/support/previous-ams-versions/manuals/amsgra), which in this study was shown by large standard deviations for the group averages per block. It is possible that these large standard deviations, in combination with the small group sizes (especially for the harsh group), were responsible for the lack of significant differences between the two groups of mothers on RMSSD levels.

It could also be that varying levels of sensitivity in the second year after birth have dampened the degree of RSA withdrawal in the non-harsh group. Previous findings have shown stronger RSA withdrawal for highly sensitive mothers (Joosen et al., 2012a) as well as a significant relation between low levels of early maternal sensitivity and the use of harsh discipline in toddlerhood (Joosen, Mesman, Bakermans-Kranenburg, & Van IJzendoorn, 2012). Therefore, we also controlled for the level of early maternal sensitivity (at 3 months after birth) in the current study. Nonetheless, since we did not include later levels of maternal sensitivity (during the second year) when creating the two groups, it is possible that the mothers in the non-harsh group included highly sensitive mothers as well as mothers with medium or low levels of sensitivity. Therefore, even though the lines in Figure 4.2 do suggest an absent RSA withdrawal for the mothers in the harsh group, the non-significance of this result may be due to a mix of mothers with various levels of sensitivity within the non-harsh group. In a larger sample one could distinguish between a highly sensitive and a moderately sensitive non-harsh group. We would predict that this would uncover significantly less RSA withdrawal in mothers showing harsh discipline compared to highly sensitive mothers who do not show harsh discipline.

In sum, our results have shown endured persistent activation of the SNS in response to repeated infant cry sounds by mothers using harsh discipline when compared to other mothers. However, the cry paradigm did not evoke significantly different PNS response patterns over time between the two groups of mothers. From a polyvagal perspective, the current results for SNS reactivity seem to confirm that repeated cry sounds are experienced by mothers in the harsh group as more psychologically stressful and possibly threatening than by the other mothers, maybe causing both chronic overactivation of the SNS and a behavioral overreactivity (i.e., harsh discipline).

It has been suggested that negative parenting responses to crying are especially related to the aversive perception of hyperphonated (high pitched) cry sounds. Zeskind and Shingler (1991), for example, found that abusive parents perceived hyperphonated cries more similar to their own abused infant's cries than the parents in the comparison group. In the present study differences in physiological reactivity to infant crying could not be attributed to group differences in perception of urgency of the three frequencies. Our results are in line with a study showing that *both* low-Child Abuse Potential and high-Child Abuse Potential adults rated hyperphonated cries as significantly more urgent, arousing, distressing, aversive and sick-sounding than the phonated cries (Crowe & Zeskind, 1992). Inconsistencies in findings regarding the perception of cries with varying pitches could be related to the items used to rate the perception of the different cries, since only one of these items (not aversive – aversive) seems to distinguish between harsh and non-harsh parents. However, a lack of empathy in response to presentations of a crying infant has been shown for mothers with a high risk for physical abuse when compared to low risk mothers (Milner, Halsey, & Fultz, 1995). It would be interesting to investigate whether dimensions such as empathetic responsiveness or negative appraisals of infant crying, instead of perception ratings, are related differently to variations in cry pitch for parents who use harsh discipline and parents who do not.

Some limitations of the current study should be mentioned. First, the small sample sizes of especially the harsh group could have influenced the reliability of the results. Results on RSA withdrawal may have been affected by the small group sizes in combination with the large standard deviations for RMSSD mean levels per segment. Second, we based our split between the groups of mothers on the more subtle signs for harsh discipline, which resulted in a total of only 10 mothers in the harsh group. Home observations of moderate to severe harsh discipline are not to be expected in a non-clinical population, since any mother would be at her best in front of a camera. Therefore, when using home-observations cut-offs for harsh parenting can only be based on the milder and subtler signs of harsh discipline, registering therefore only the tip of the iceberg. Last, it could be seen as a limitation that we did not control for respiration rate in our estimation of RSA. Some studies have shown that RSA magnitude can be altered purely as a function of respiratory changes, even in conditions of mental activation or emotion (Grossman & Taylor, 2007). In a comparative study between different quantification methods for RSA using data collected with a similar ambulatory device, RMSSD levels were proven as a reliable and valid estimate of RSA (Goedhart, Van der Sluis, Houtveen, Willemsen, & De Geus, 2007). Since all participants were seated during the entire cry paradigm (including baseline measurement), and did not speak during cry presentations, changes in respiration rate due to physical activity or talking could only have had a minimal effect on our estimation of RSA.

In conclusion, the current study shows a sympathetic overreactivity to repeated cry sounds among mothers observed to use harsh discipline when compared to mothers who did not. Mean SCL of harsh discipline using mothers continued to rise across the repeated exposure to cry sounds, while after an initial increase the mean SCL of the other mothers showed a steady decline. Therefore, we suggest that harsh parenting reflects behavioral as well as physiological overreactivity to negative infant signals.