FROM PREGNANCY TO POSTPARTUM:
BIOBEHAVIORAL PREDICTORS AND MECHANISMS OF EARLY PARENTING

Christine Hechler
From pregnancy to postpartum:
Biobehavioral predictors and mechanisms of early parenting
Promotoren:
Prof. dr. C. de Weerth
Prof. dr. J.M.A. Riksen-Walraven

Copromotor:
Dr. R. Beijers

Manuscriptcommissie:
Prof. dr. E.S. Kluwer (voorzitter)
Prof. dr. L.R.A. Alink (Universiteit Leiden)
Dr. L. Karlsson (University of Turku, Finland)
Prof. dr. R. Keizer (Erasmus University Rotterdam)
Dr. J.K. Larsen

CONTENT

Chapter 1    General Introduction                      7

Chapter 2    Young Adults’ Reaction to Infant Crying     21

Chapter 3    Prenatal Predictors of Postnatal Quality of Caregiving Behavior in Mothers and Fathers  43

Chapter 4    Prenatal and Postnatal Cortisol and Testosterone are Related to Parental Caregiving Quality in Fathers, but not in Mothers  69

Chapter 5    Association Between Fecal Microbiota and Anxiety in Pregnant Women  99

Chapter 6    Are Cortisol Concentrations in Human Breast Milk Associated with Infant Crying?  121

Chapter 7    Summary, Conclusion, and General Discussion  147

Appendices  Nederlandse samenvatting/summary in Dutch  165
Deutsche Zusammenfassung/summary in German  171
Curriculum Vitae  176
Publications  177
Acknowledgements  178
Chapter 1

General Introduction
1. General Introduction

For the past decades, numerous studies have shown that early life experiences, even during pregnancy, have a tremendous impact on infant and child development (Loman & Gunnar, 2010). Research and the resulting scientific knowledge in this area are highly relevant, because knowledge on how early experiences shape development can increase the chances of early detection of children at higher risk of adverse developmental outcomes. This knowledge can provide valuable insights for the design of prevention and intervention programs to decrease the risk of adverse developmental outcomes in infancy and childhood. While the research field of early development is growing and provides many interesting and valuable insights, many gaps in knowledge remain. The broader goal of the current thesis is to contribute scientific knowledge to the field of early development, by focusing on two specific areas that are approached through two research aims.

The first aim of this thesis is to shed light on possible prenatal predictors of the quality of parental caregiving in the first months after the birth of their infant. Given that parental caregiving affects child development from birth onwards (Feldman, Edelman, & Rotenberg, 2004), and is relatively stable over time (Else-Quest, Clark, & Tresch Owen, 2011; Hall et al., 2015), it is important to predict the quality of caregiving as early as possible, potentially already during pregnancy. The current thesis focuses on several plausible prenatal predictors of parental quality of caregiving that have not been examined yet.

The second aim of the current thesis is to examine possible biological mechanisms underlying the link between maternal stress and child outcomes. Although previous research has shown that both prenatal and postnatal maternal stress are related to child development in many ways, there is still a lack of knowledge on the precise mechanisms through which maternal stress may affect the child. The current thesis focuses on two potential biological pathways that have not been investigated previously.

This first chapter is structured as follows: section 1.1 focuses on the first aim, namely possible prenatal predictors of parental quality of caregiving. Section 1.2 focuses on the second aim, namely possible biological mechanisms through which maternal stress may affect the child. Section 1.3 summarizes the longitudinal project that has been set up to reach the goals of this thesis and to answer the research questions. The chapter ends with section 1.4 which includes the specific research questions and aims, and an overview of the variables and measures used in this thesis.

1.1 First Aim: Prenatal Predictors of Postnatal Parental Quality of Caregiving

High quality parental caregiving behavior towards the infant is important for a positive developmental trajectory. For example, high quality parenting has been linked to better self-regulatory capacities (Sroufe, 2005), lower child stress levels (Smeekens, Riksen-Walraven, & van Bakel, 2007), better cognitive performance (Klein & Feldman, 2007), and prosocial behavior (Eisenberg & Fabes, 1998).

An important question then is what explains differences in quality of parental caregiving? While research has pointed at several factors predicting caregiving quality, including demographic factors, the quality of the parental support system, parental characteristics (e.g., personality, cognitive capacities), and mental health problems (for a comprehensive review, see Thomson et al., 2014), most of these factors have been assessed after the child was born. Yet, as parental caregiving affects child development from birth onwards (Feldman et al., 2004), it is beneficial for intervention and prevention purposes to determine predictors of parental caregiving behavior already during pregnancy. Additionally, the previously described factors are also more general in nature, only predicting a small amount of variance in parental behavior (see for example Pelchat, Bisson, Bois, & Sacuier, 2003). Looking at actual parental caregiving behavior prenatally, instead of general parental factors, may be a better way of predicting how parents will interact with their infants after birth (Wachs & Gruen, 1982).

As the quality of parental caregiving behavior has been observed to be stable till adolescence (Else-Quest et al., 2011; Hall et al., 2015), it is plausible that caregiving behavior observed prenatally is predictive of postnatal parental caregiving. We observed the quality of maternal and paternal caregiving behavior, both in the prenatal and postpartum period. Prenatal observations of parental behavior were carried out in the context of infant crying, as it is known that infant crying interrupts parental cognition (Morsbach, McCulloch, & Clark, 1986), leads to a change in mental focus (De Pisapia et al., 2013), and can elicit negative emotional reactions in the parent (Leerkes, Parade, & Gudmundson, 2011). Together, these reactions are assumed to trigger caregiving efforts (Chang & Thompson, 2011). Hence, situations involving infant crying were considered the most appropriate for triggering and observing expectant parents’ caregiving behavior. Whether the prenatal caregiving behavior predicted how parents cared for their own infant in the first weeks postpartum was determined in the current thesis.

Next to prenatal observations of parental caregiving, this thesis also examined another prenatal predictor of postpartum quality of caregiving, namely parental hormones. As recently reviewed by Bos (2017), differences in hormonal levels may be related to the quality of caregiving. We focused on the role of the steroid hormones cortisol and testosterone in parenting behavior. In mothers, higher cortisol levels postpartum have been related to more affection towards the infant (Fleming, Steiner, & Anderson, 1987),
more responsiveness to baby odor (Fleming, Steiner, & Corter, 1997) and more sympathy
towards infant crying (Stallings, Fleming, Corter, Worthman, & Steiner, 2001). In fathers, lower
testosterone levels postpartum have been related to more parental investment (Mascaro,
Hackett, & Rilling, 2013), more responsiveness to infant cues (Storey, Walsh, Quinton, &
Wynne-Edwards, 2000), and stronger sympathy in response to infant crying (Fleming, Corter,
Stallings, & Steiner, 2002). Thus, the hormones cortisol and testosterone might also be related
to the quality of parental caregiving behavior in both women and men. The present thesis
therefore investigated whether prenatal variations in cortisol and testosterone can predict
the postnatal quality of caregiving behavior in mothers and fathers.

1.2 Second Aim: Possible Biological Mechanisms through Which Maternal Stress may
affect the Child

Maternal stress during the prenatal and postpartum period is known to be closely related
to infant development, but the underlying mechanisms still need to be further unraveled.
The present thesis focuses on two biological mechanisms that have received little attention
in the past.

The first biological mechanism is related to the intestinal microbiota. Intestinal
microbiota refers to the microorganisms within the gastrointestinal tract that have a central
position in health and disease. One possible underlying mechanism for many of the child
outcomes related to maternal prenatal psychosocial stress might be that maternal prenatal
psychosocial stress could affect child development and health via the mother’s, and in turn
the child’s, intestinal microbiota (Beijers, Buitelaar, & de Weerth, 2014), see Figure 1.1.

From studies in animals and humans, we know that the infant microbiota have a
central position in infant health and disease from birth onwards (Bäckhed, 2011). The microbiota is paramount in the development of the immune system and contributes
to the maturation and function of the intestines (Dimmit et al., 2010). The microbiota also synthesizes essential vitamins and generates nutrients (Adlerberth & Wold, 2009).
Moreover, results from animal studies and a few human studies suggest that the intestinal
bacteria may be related to host cognition and behavior (Saulnier et al., 2013; Cryan
& Dinan, 2012; Luczynski et al., 2016), suggesting that the intestinal microbiota may play an
important role in both human physical as well as mental health.

The infant intestines are virtually sterile at birth, and within minutes the microbial
colonization begins (de Weerth, Fuentes, & de Vos, 2013; Grönlund, Grzeskowiak, Isolauri,
& Salminen, 2011). The microorganisms important for this colonization originate mainly
from the mother in case of vaginal deliveries (de Weerth et al., 2013; Grönlund et al., 2011).
Hence, the maternal intestinal microbiota would affect infant intestinal microbiota. The hypothesis follows that unbalances in the maternal microbiota, e.g. as a possible result
of maternal psychosocial stress, affect the microbial colonization of the infant intestines,
with possible consequences for child physical and mental health. The link between
maternal prenatal psychosocial stress and maternal microbiota, presented in the model
in Figure 1.1, was studied as part of the present thesis. This link has not been studied
before, but there is some evidence from non-pregnant animal and human studies that
psychological stress is related to altered intestinal microbiota (Bailey et al., 2011; Knowles,
Nelson, & Palombo, 2008).

The second biological mechanism through which maternal stress might be related
to infant health and development is breast milk composition. Maternal stress has been
associated with increased infant fussing and crying (DeSantis, Coster, Bigsby, & Lester,
2004; Rautava, Helenius, & Lehtonen, 1993), and this link may be mediated through
maternal milk components. Next to water, protein, carbohydrates, and immune factors,
breast milk contains several biologically active hormones, including cortisol (Miller et
al., 2013). Cortisol is freely available in plasma, and cortisol levels increase in response
to acute stressors (Dickerson & Kemeny, 2004). From plasma, cortisol is transferred to
breast milk; there is no mammary synthesis of cortisol (Hamosh, 2001). Hence, when
breastfeeding mothers are psychologically stressed, it is plausible that their milk contains
higher concentrations of cortisol and that this in turn may affect levels of infant fussing
and crying.

To date, there are no studies examining the link between milk cortisol and infant
crying. However, previous research showed that higher cortisol concentrations in plasma
-an indicator of concentrations in breast milk- were related to a more fearful temperament
in 2-month-old infants (Glynn et al., 2007). Furthermore, higher cortisol concentrations in
breast milk were related to more temperamentally negative affectivity in female 3-month-old infants (Grey, Davis, Sandman, & Glynn, 2013) and to higher laboratory-induced fear reactivity in female 8-month-olds (Nolvi et al., 2017). The present thesis will add to the scarce literature and help further uncover a possible role of maternal milk on infant behavioral development, by investigating whether milk cortisol is associated with infant fussing and crying behavior.

1.3 The Longitudinal BINGO Study

All but one of the empirical studies presented in this thesis were part of the prospective longitudinal BINGO study that was especially set up to answer the main research questions of this thesis. BINGO is a Dutch acronym for Biological Influences on Baby’s Health and Development (in Dutch: Biologische INvloeden op baby’s Gezondheid en Ontwikkeling). The BINGO study followed expectant parents from the third trimester of pregnancy to 12 weeks after birth. Given the promising outcomes, the BINGO study was continued and is ongoing, with the most recent measurement when the infants were 1 year old and an upcoming measurement at three years of age.

Healthy expectant mothers and fathers from the region Arnhem-Nijmegen in The Netherlands were asked to participate in the BINGO study by means of flyers, personal communication, and the project’s website. At the start of the project, 88 expectant mothers and 57 of their partners signed informed consent and subsequently participated in the project. At the 12-week data collection, 79 mothers and 53 of their partners were still participating (drop-out mainly due to birth complications or personal circumstances).

The data within the BINGO study were collected at different time points. When the mothers were around 35 weeks pregnant, prenatal parental quality of caregiving, reactions to infant crying, and salivary cortisol and testosterone were measured during a lab visit. Furthermore, expectant mothers reported on psychosocial stress by means of questionnaires and collected a stool sample to analyze microbiota. When the infant was 6 weeks of age, postnatal parental quality of caregiving was measured during a home visit. When the infant was 2, 6, and 12 weeks old, mothers collected a breast milk sample in which cortisol was analyzed, and reported on infant fussing and crying behavior by means of a three-day diary. See Figure 1.2 for an overview of the variables measured in the current thesis, including the relations between the variables as investigated in the different chapters from this thesis. The chapters are summarized in the following section.

1.4 Thesis Outline

The first and second aim of this thesis were addressed in five empirical studies (Chapter 2 through 6). A schematic overview of the main research questions examined in all five studies can be found in Figure 1.2.
The studies reported in Chapters 2, 3, and 4 are related to the first aim of the thesis, i.e. investigating possible prenatal predictors of postnatal quality of caregiving behavior. In Chapter 2, cognitive and emotional reactions to infant crying were investigated in a student sample (N=120) not part of the BINGO study. This study served as a preliminary study for the study in Chapter 3, where postnatal quality of caregiving behavior at 6 weeks of infant age was predicted by parental reactions to infant crying and quality of caregiving measured at the end of pregnancy. In Chapter 4, postnatal quality of caregiving behavior at 6 weeks of infant age was predicted by using parental salivary cortisol and testosterone concentrations measured at the end of pregnancy.

The studies reported in Chapters 5 and 6 are related to the second aim of the thesis, i.e. investigating possible mechanisms that could explain how maternal prenatal and early postnatal stress may influence the behavior and development of her child. Related to prenatal stress, Chapter 5 investigated whether maternal psychosocial stress was related to maternal microbiota at the end of pregnancy. Regarding a possible influence of maternal postnatal stress on child development, Chapter 6 examined whether breast milk cortisol was related to infant fussing and crying behavior when the infant was 2, 6 and 12 weeks of age. Finally, Chapter 7 presents a general discussion of the five chapters, including a summary of the five studies, the main conclusions, and discussion points.

REFERENCES


Chapter 2

Young Adults’ Reactions to Infant Crying

ABSTRACT

Aim: To determine whether young childless adults show negative emotions and cognitive disturbances when listening to infant crying, compared to other disturbing noises, and whether negative emotions and cognitive disturbances are associated. Methods: We tested the cognitive performances and emotional reactions of 120 childless participants on a working memory task while being subjected to different disturbing noises including infant crying. Results: Participants had the least correct trials on the working memory task, and showed the most negative emotions, when hearing infant crying as compared to the other noises. Participants also showed less positive emotions when hearing infant crying as compared to working in silence. Overall, negative emotions were associated with less correct trials on the working memory task, except in the infant crying condition. Furthermore, cognitive performance and emotional reactions to infant crying were unrelated to personality characteristics. Conclusion: Negative emotions and cognitive disturbances may be general adult responses to infant crying that are not limited to parents. These results suggest a broadly present human emotional and cognitive response to infant crying, that may underlie a general predisposition to care for infants in distress.

INTRODUCTION

Infant crying is a highly salient signal that leads to diverse reactions in the caregiver, usually the parents. For example, listening to infant crying has been related to an increase in negative emotions in parents. Leerkes, Parade, and Gudmundson (2011) found that the emotional reaction of mothers to infant crying can range from sad to irritated to extremely hostile. Such reactions presumably compel the caregiver to react promptly and trigger appropriate caregiving efforts (Chang & Thompson, 2011).

Infant crying also appears to reach mothers who are not attending to the baby, but who are mentally focused elsewhere instead (De Pisapia et al., 2013). In line with this, a study by Morsbach, McCulloch, and Clark (1986) showed that mothers with infants below one year of age were distracted in their performance on a simple concentration task by listening to a tape of a crying infant, as opposed to machine noises. Apparently, infant crying leads to a disruption in mental focus, possibly in order to direct the attention of caregivers to infant needs.

The above leads us to speculate that the emotional reactions and cognitive disturbances as a reaction to infant crying may be complementary, and that these two processes may share the same underlying mechanism. However, to our knowledge, emotional reactions and cognitive performance while listening to infant crying have not been investigated simultaneously or in relation to each other. Also, it is unknown whether these emotional reactions and cognitive disturbances to infant crying are specific for parents or are also present in childless adults. The present study will therefore investigate the emotional reactions and cognitive performance of young childless adults while being subjected to infant crying.

Many researchers assume that psychological and physiological changes in the transition to parenthood influence and shape a person’s reactions to infant crying (e.g., Rilling, 2013; Seifritz et al., 2003; Swain, Kim, & Ho, 2011). For example, cortisol levels naturally rise throughout the second and third trimester of pregnancy (Carr, Parker, Madden, MacDonald, & Porter, 1981), and mothers with higher cortisol levels have been found to show more sympathetic responses to infant crying (Stallings, Fleming, Corter, Worthman, & Steiner, 2001). On the other hand, there are also studies showing that childless adults react similarly to parents to the sound of infant crying (e.g., Chang et al., 2011; Riem, Bakermans-Kranenburg, van Uzendoorn, Out, & Rombouts, 2012), suggesting a more general predisposition to react to infant crying, regardless of parental status. For example, listening to infant crying improved motor performance in both parents and childless adults (Parsons, Young, Parsons, Stein, & Kringelbach, 2012). According to the authors, this may facilitate caregiving behavior, hence constituting an adaptive reaction to crying. In addition, childless adults, as well as parents, react to infant
crying with physiological arousal, including increases in heart rate, skin conductance, and blood pressure (Parsons et al., 2012).

Although it can be hypothesized that negative emotional reactions and cognitive disturbances resulting from listening to infant crying might also be seen in childless adults, only one study investigated negative emotions in childless adults (Bruning & McMahon, 2009). In this study, young female students interacting with a crying infant simulator doll showed an increase in negative emotions. To date, there are no studies available investigating cognitive performance in childless adults when listening to infant crying.

In sum, the aim of this study is to investigate whether the effects of crying on emotions and cognitive performance are present in young childless adults. We will do this by simultaneously investigating the emotional reactions and cognitive performances on a working memory task in young childless adults, while being subjected to infant crying and other potentially disturbing noises. We will also investigate whether these emotional reactions and cognitive performances can be predicted by personality traits, as personality might potentially influence reactions to infant crying as well.

Zeifman (2003), for example, found that parents with high levels of empathy showed more sensitive responses to infant crying. Also, carrying out a cognitive task while being disturbed by noise represents a potentially stressful situation, where the traits of ego-resiliency, hostility, and impulsivity might be relevant. Ego-resiliency is the general ability to manage stress (Block & Block, 1980), whereas people high on hostility are less buffered against the negative effects of stressful events (Morrison & Bennett, 2009). Impulsivity denotes a tendency to act without reflecting on one's actions (Gleitman, Reisberg, & Gross, 2007), which might be especially true in stressful situations where the ability to reflect might be hindered by competing cognitively demanding tasks. By including personality characteristics, we may be able to further understand differences in individual reactions to infant crying. On the other hand, if personality characteristics fail to explain different reactions to crying, this would lend support to the notion of a more general human predisposition to care for infants in distress.

The present study will be carried out in a controlled laboratory setting. Although ecologically less valid than naturalistic observations, the laboratory standardizes situational factors over individuals, reducing the influence of possible confounding variables. Furthermore, for investigating reactions to infant crying, compared to other noises, it is necessary to control the frequency, pitch, and volume of the diverse sounds. This would be very difficult in a naturalistic setting.

We hypothesize that (1) childless adults make more mistakes on the working memory task (i.e. fewer correct trials), and (2) show more negative emotions and (3) less positive emotions, while being subjected to infant crying as compared to other disturbing noises. We also expect that (4) cognitive performance and emotional reactions while listening to infant crying will be predicted by personality traits. Finally, we expect that (5) cognitive performance and emotional reactions will be associated, i.e. more negative, and less positive, emotions will be related to fewer correct trials on the working memory task.

**METHOD**

**Participants**

In total, 122 students from the Radboud University in Nijmegen, The Netherlands, participated in this study. Exclusion criteria were: being pregnant, being a parent, drug and alcohol abuse. One participant became ill during testing and was therefore excluded from all analyses. A second participant’s data was incomplete due to technical failure and had to be excluded as well. The final sample included 120 university students (90 female, $M=21.62$ years, $SD=3.24$ years). Most participants were Dutch (61.70%; German: 35.00%; other: 3.30%).

**Design and Procedure**

Participants completed online questionnaires on demographics and personality before the lab visit. Subsequently, participants were invited to come to the lab and were not informed of the goal of the present study. Upon arrival, participants were fitted with a heart rate monitor to measure their heart rate. Participants did the N Back task (Kirchner, 1958) that measured working memory. Afterwards, participants continued with four other tests that are not within the scope of this paper.

Participants did the N Back task five times in a row: first in silence, and then four times while hearing noises through a headphone: a crying infant, a dinner conversation, traffic noises, and an excavator (the order of the noises was counter-balanced). These noises were chosen and matched for frequency, pitch and volume, and then pilot tested beforehand on 30 childless university students. Repeated-measures analyses with post hoc comparisons of the pilot test results showed that the noises were relatively equally loud and disturbing. Nevertheless, the crying noise was rated as more irritating ($p=.006$) and more disturbing ($p=.014$) than the excavator noise. A fifth noise, a barking dog, was not used in the present study because several participants from the pilot study reported being scared by this noise, and the score on aversiveness was higher compared to the other noises (all $ps<.05$).
Measures and Instruments

Cognitive performance. To measure cognitive performance while listening to infant crying, compared to other noises, the visual version of the N Back task was used (Kirchner, 1958). The N Back task measures an individual’s working memory. The participant is presented with a 3 by 3 grid, in which every 2s one of the cells turns red. The participant is required on each trial to indicate whether the cell that just turned red is the same cell that turned red n trials before. The factor n can be adjusted for different levels of difficulty. We used an n of 2, which corresponds to a low level of difficulty (Doumas, Rapp, & Kramer, 2009), which is in line with a low level of difficulty used by Morsbach et al. (1986). After a practice run, participants had to do this test five times, each time with a different noise. Each noise condition lasted 90s and had 35 trials. Cognitive performance was calculated by the number of correct trials per noise condition (i.e. Silence, Infant crying, Dinner conversation, Traffic, Excavator).

Emotions. After each of the five of the five N Back tasks, participants rated how strongly they felt positive and negative emotions (Leerkes et al., 2011) on a scale from 1 = not at all to 4 = very much. The questionnaire contained two positive emotions, namely happy and amused, and four negative emotions, namely bothered, frustrated, angry, and irritated. Exploratory factor analysis with Principal Axis Factoring as the extraction method and Promax rotation showed two clear factors per N Back condition. In all N Back conditions, the two positive emotions loaded on one factor, and the four negative emotions loaded on another. Cronbach’s alphas were acceptable to good for the positivity and negativity factor ranging from α = .62 (Positivity in the Silence condition) to α = .86 (Negativity in the Dinner conversation condition). Therefore, per noise condition, a positivity factor (sum of scores of positive emotions) and a negativity factor (sum of scores of negative emotions) were calculated and used as dependent variables in the analyses.

Personality.

Ego-resiliency. Ego-resiliency was measured with a Dutch translation of the Ego-resiliency scale (Letzring, Block, & Funder, 2005). This scale consists of 14 statements that have to be answered on a scale from 1 = does not apply at all to 4 = applies very strongly. An example item is “I quickly get over and recover from being startled.” Adding up all 14 items leads to a sum score for ego-resiliency. A higher score on this scale means a higher level of ego-resiliency. Internal reliability of this scale for this sample was acceptable (Cronbach’s alpha = .72).

Empathy. Empathy was measured with a Dutch translation of the Interpersonal Reactivity Index (Davis, 1980). This questionnaire consists of 28 statements that can be answered on a scale from 0 = does not describe me well to 4 = describes me very well. The scores on all 28 items are added up, with a higher score indicating a higher level of empathy. Internal consistency for this sample was acceptable (Cronbach’s alpha = .66).

Hostility and impulsivity. Hostility and impulsivity were measured with subscales of the Quick Big Five (Gerris et al., 1998). This questionnaire consists of 30 characteristics (e.g., irritable, negligent, uneasy) and participants have to indicate whether these characteristics appropriately describe them. Answers can be given on a scale from 1 = not at all true to 7 = completely true. Of these 30 characteristics, the sum of scores on the 6 hostility characteristics was taken to indicate level of hostility. The sum of scores on the 6 impulsivity characteristics was taken to indicate the level of impulsivity. Higher scores indicated more hostility and impulsivity, respectively. Confirmatory factor analysis with Principal Axis Factoring as the extraction method and Promax rotation confirmed the subscales of the Quick Big Five (Gerris et al., 1998). The internal consistencies for these two constructs in this sample were good (Cronbach’s alpha = .86 and .81, respectively).

Control variables

Five potential control variables were included in the analyses: sex (0 = male, 1 = female), nationality (0 = Dutch, 1 = other), duration of romantic relationship, experience with caring for children, and depression. Duration of romantic relationship was operationalized as length of the present romantic relationship in months (singles received a score of 0). Caregiving experience was indicated on a scale from 1 = none to 5 = a lot. Depression was measured with a Dutch translated version of the CES-D (Radloff, 1977). This questionnaire consists of 20 statements that have to be answered on a scale from 0 = rarely or never to 3 = mostly or always. A higher score indicates more depressive symptoms. Internal consistency for the CES-D in this sample was excellent (Cronbach’s alpha = .90).

Statistical analyses

All variables were controlled for outliers and normality. Number of correct trials for each of the five N Back noise conditions and duration of relationship were not normally distributed and therefore log transformed. Outliers were defined as scores below -2.50 and above 2.50 SD from the mean, and subsequently replaced with the lowest or highest score, respectively, within the -2.50 to 2.50 SD range. The log-transformed number of correct trials for every N Back noise condition had outliers on the low end. We replaced each four scores within the Silence and Infant crying noise condition, six scores within the Dinner conversation noise condition, five scores within the Traffic noise condition, and seven scores within the Excavator noise condition. The following N Back noise conditions contained outliers in the high end for negative emotions: two scores in the Infant crying noise condition, three scores within the Dinner conversation noise condition, and each one score within the Traffic and Excavator noise condition. Furthermore, Ego-resiliency had two outliers on the high end, Hostility and Depression each had one outlier on the
high end, and the log-transformed Duration of relationship had two outliers on the high end that were replaced. Results from the main analyses with the outliers included in the analyses were comparable to the results from the main analyses with the outliers replaced.

Three repeated measures analyses of variance with noise condition of the N Back task as the within-subject factor were done: one for the number of correct trials per N Back noise condition, one for negative emotions per N Back noise condition, and one for positive emotions per N Back noise condition. To test the predictive ability of personality characteristics on cognitive performance under influence of infant crying, three repeated measures ANOVAs (i.e., number of correct trials, negative emotions, positive emotions) were done with noise condition as the within subject factor, personality characteristics as covariates, and controlling for all confounders.

To test whether emotions could explain variance in the number of correct trials, regression analyses were done per N Back noise condition, with negative and positive emotions during the N Back task as the predictors, the number of correct trials as the dependent variable, and controlling for all confounders. Regression analyses were performed on the full (i.e. data previously deemed outliers not changed), untransformed dataset. A check for assumption revealed non-normally distributed residuals in all analyses. Therefore, regression analyses with bootstrapping were performed. As a result, confidence intervals are reported.

RESULTS

Preliminary analyses

Concerning the heart rate data, due to technical difficulties we were only able to read the data of 52 participants (42.6%). We therefore decided to exclude heart rate data from the analyses. Descriptive statistics of the predictors, confounders, and outcome variables can be found in Table 1. Table 2 presents the correlations between the outcome measures, personality characteristics and confounders for the Infant Crying condition.

Table 1
Descriptive Statistics of the Predictors, Confounders and Outcome Measures

<table>
<thead>
<tr>
<th>Outcome measures:</th>
<th>Mean ± SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct trials:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silence</td>
<td>32.41 ± 3.06</td>
<td>20</td>
<td>35</td>
</tr>
<tr>
<td>Infant crying</td>
<td>30.67 ± 2.44</td>
<td>20</td>
<td>34</td>
</tr>
<tr>
<td>Dinner conversation</td>
<td>33.07 ± 2.60</td>
<td>22</td>
<td>35</td>
</tr>
<tr>
<td>Traffic</td>
<td>32.14 ± 2.83</td>
<td>18</td>
<td>35</td>
</tr>
<tr>
<td>Excavator</td>
<td>32.84 ± 2.57</td>
<td>22</td>
<td>35</td>
</tr>
<tr>
<td>Negative Emotions:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silence</td>
<td>6.42 ± 2.76</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Infant crying</td>
<td>7.73 ± 2.86</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Dinner conversation</td>
<td>7.36 ± 3.02</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Traffic</td>
<td>6.60 ± 2.56</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Excavator</td>
<td>6.45 ± 2.48</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Positive Emotions:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silence</td>
<td>4.52 ± 1.24</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Infant crying</td>
<td>4.12 ± 1.57</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Dinner conversation</td>
<td>4.46 ± 1.51</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Traffic</td>
<td>3.96 ± 1.37</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Excavator</td>
<td>3.90 ± 1.36</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Personality Characteristics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empathy</td>
<td>94.01 ± 9.83</td>
<td>70</td>
<td>112</td>
</tr>
<tr>
<td>Ego-Resiliency</td>
<td>40.42 ± 4.43</td>
<td>29</td>
<td>51</td>
</tr>
<tr>
<td>Hostility</td>
<td>22.60 ± 6.00</td>
<td>8</td>
<td>37</td>
</tr>
<tr>
<td>Impulsivity</td>
<td>27.90 ± 6.44</td>
<td>13</td>
<td>41</td>
</tr>
<tr>
<td>Confounders:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex (female)</td>
<td>75.00%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of relationship (in months)</td>
<td>15.39 ± 25.73</td>
<td>0</td>
<td>188</td>
</tr>
<tr>
<td>Experience with children</td>
<td>2.54 ± 1.07</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Depression</td>
<td>11.58 ± 8.08</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>Nationality (Dutch)</td>
<td>61.70%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2
Correlations Between Outcome Measures, Personality Characteristics and Potential Confounders for the Crying Infant N-Back task

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Correct trials Infant crying</td>
<td>-23*</td>
<td>.14</td>
<td>-.10</td>
<td>-.02</td>
<td>-.02</td>
<td>-.11</td>
<td>.07</td>
<td>-.06</td>
<td>-.12</td>
<td>-.10</td>
<td></td>
</tr>
<tr>
<td>2. Negative emotions Infant crying</td>
<td>-.45**</td>
<td>.05</td>
<td>-.06</td>
<td>.24**</td>
<td>.06</td>
<td>.04</td>
<td>.08</td>
<td>-.03</td>
<td>.18</td>
<td>.16</td>
<td></td>
</tr>
<tr>
<td>3. Positive emotions Infant crying</td>
<td>-.18</td>
<td>.01</td>
<td>-.10</td>
<td>-.12</td>
<td>-.13</td>
<td>.04</td>
<td>.18*</td>
<td>-.05</td>
<td>-.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Empathy</td>
<td>-.11</td>
<td>.32**</td>
<td>.19*</td>
<td>.24**</td>
<td>-.02</td>
<td>.09</td>
<td>.16</td>
<td>.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Ego-Resiliency</td>
<td>-.35**</td>
<td>-.06</td>
<td>-.04</td>
<td>-.08</td>
<td>-.03</td>
<td>.31**</td>
<td>.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Hostility</td>
<td>.18</td>
<td>.15</td>
<td>.02</td>
<td>.01</td>
<td>.39**</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Impulsivity</td>
<td>.21*</td>
<td>.09</td>
<td>.09</td>
<td>.05</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Sex</td>
<td>.24**</td>
<td>-.07</td>
<td>-.18</td>
<td>.13</td>
<td>-.06</td>
<td>.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Duration of relationship</td>
<td>-.19*</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Experience with children</td>
<td>-.19*</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Depression</td>
<td>.23*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 12. Nationality | Note. ** p < .01, * p < .05

As can be seen in Table 2, negative emotions during the Infant crying N Back correlated negatively with the number of correct trials during the Infant crying N Back. Hence, the more negative emotions were felt, the more mistakes were made. Negative and positive emotions during the Infant crying N Back were also positively correlated, indicating a similar manner within all remaining N Back conditions (not shown in the table).
when listening to Infant crying, as opposed to Silence ($F_{1,119}=37.38$, $p<.001$, $\eta^2_a=.239$), the Dinner conversation ($F_{1,119}=188.42$, $p<.001$, $\eta^2_a=.613$), Traffic ($F_{1,119}=68.24$, $p<.001$, $\eta^2_a=.364$) and Excavator ($F_{1,119}=199.82$, $p<.001$, $\eta^2_a=.627$) noises. Analyses showed that the four personality characteristics did not have an effect on cognitive performance in any of the noise conditions (all $p's>.05$), and that there were no interactions between personality characteristics and noise condition (all $p's>.05$).

**Negative emotions.** A repeated measures analysis of variance with N Back noise condition as a within-subject factor, and Infant crying noise condition as the reference category, showed a significant difference in negative emotions over the noise conditions ($F_{4,116}=8.86$, $p<.001$, $\eta^2_a=.234$; Fig. 2). Post hoc analyses revealed that participants reported more negative emotions during the Infant crying compared to the Silence, Traffic, and Excavator noise conditions ($F_{1,119}=22.32$, $p<.001$, $\eta^2_a=.158$; $F_{1,119}=21.04$, $p<.001$, $\eta^2_a=.150$; and $F_{1,119}=31.29$, $p<.001$, $\eta^2_a=.208$, respectively). Participants reported comparable negative emotions during the Infant crying and the Dinner conversation noise conditions ($F_{1,119}=2.09$, $p=.151$, $\eta^2_a=.017$). The four personality characteristics did not have an effect on negative emotions in any of the noise conditions (all $p's>.05$), and there were no interactions between personality characteristics and noise condition (all $p's>.05$).

**Positive emotions.** A repeated measures analysis of variance with N Back noise condition as the within-subject factor, and Infant crying noise condition as the reference category, showed a significant difference in positive emotions over the noise conditions ($F_{4,116}=10.66$, $p<.001$, $\eta^2_a=.269$; Fig. 3). Post hoc analyses revealed that participants reported significantly more positive emotions during the Silence and Dinner conversation N Backs compared to the Infant crying N Back ($F_{1,119}=8.14$, $p=.005$, $\eta^2_a=.064$; $F_{1,119}=4.35$, $p=.039$, $\eta^2_a=.035$, respectively). Participants reported comparable positive emotions during the Infant crying, the Traffic and Excavator noise conditions ($F_{1,119}=1.34$, $p=.250$, $\eta^2_a=.011$; $F_{1,119}=2.36$, $p=.128$, $\eta^2_a=.019$, respectively).

Ego-resiliency, hostility, and impulsivity were unrelated to positive emotions in all of the noise conditions (all $p's>.05$). Empathy was related to positive emotions in all noise conditions ($F_{1,109}=6.13$, $p=.015$), and this did not differ per condition ($F_{4,106}=1.19$, $p=.319$). Participants higher on empathy had less positive emotions throughout the noise conditions, as compared to participants lower on empathy. A positive interaction of ego-resiliency and noise condition ($F_{1,106}=3.82$, $p=.006$) was further investigated with univariate ANOVAs. These analyses showed that ego-resiliency was unrelated to positive emotions in all of the individual noise conditions.
Hypothesis 5: Cognitive performance and emotional reactions. We hypothesized that cognitive performance and emotional reactions are associated. Table 3 shows the results of the regression analyses predicting number of correct trials by positive and negative emotions. As can be seen in the table, for every N Back noise condition, except the Infant crying condition, negative emotions were inversely related to the number of correct trials. The amount of explained variance for the whole model varied from 6.90% in the Infant crying N Back to 26.60% in the Silence noise condition.

Table 3
Negative and Positive Emotions as Predictors of Correct Trials (Controlled for Confounders) per N Back Sound Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Total b</th>
<th>95% Confidence Interval</th>
<th>R²</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(b+bias) lower</td>
<td>lower upper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Sound</td>
<td>.310 .266</td>
<td>- .008 - .382</td>
<td>- .657 - .096</td>
<td></td>
</tr>
<tr>
<td>Negativity*</td>
<td>-.374 -.008</td>
<td>-.382 -.657</td>
<td>-.008 -.657</td>
<td></td>
</tr>
<tr>
<td>Positivity</td>
<td>.158 .010</td>
<td>.148 -.374</td>
<td>.010 -.374</td>
<td></td>
</tr>
<tr>
<td>Infant crying</td>
<td>-.201 -.001</td>
<td>-.202 -.437</td>
<td>.021</td>
<td></td>
</tr>
<tr>
<td>Negativity</td>
<td>-.354 .002</td>
<td>-.352 -.528</td>
<td>-.199</td>
<td></td>
</tr>
<tr>
<td>Positivity</td>
<td>.019 .003</td>
<td>.012 -.298</td>
<td>.019</td>
<td></td>
</tr>
<tr>
<td>Dinner conversation</td>
<td>.229 .181</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negativity*</td>
<td>-.354 .002</td>
<td>-.352 -.528</td>
<td>-.199</td>
<td></td>
</tr>
<tr>
<td>Positivity</td>
<td>.019 .003</td>
<td>.012 -.298</td>
<td>.019</td>
<td></td>
</tr>
<tr>
<td>Traffic</td>
<td>-.019 -.003</td>
<td>-.22 -.298</td>
<td>.047</td>
<td></td>
</tr>
<tr>
<td>Negativity</td>
<td>-.298 .000</td>
<td>-.298 -.523</td>
<td>.079</td>
<td></td>
</tr>
<tr>
<td>Positivity</td>
<td>.105 .007</td>
<td>-.098 -.389</td>
<td>.199</td>
<td></td>
</tr>
<tr>
<td>Excavator</td>
<td>-.444 .002</td>
<td>-.446 -.659</td>
<td>.240</td>
<td></td>
</tr>
<tr>
<td>Negativity*</td>
<td>-.444 .002</td>
<td>-.446 -.659</td>
<td>.240</td>
<td></td>
</tr>
<tr>
<td>Positivity</td>
<td>-.049 .000</td>
<td>-.049 -.315</td>
<td>.233</td>
<td></td>
</tr>
</tbody>
</table>

Note. Confidence intervals not containing zero are significant. Significant predictors are marked with a *.

DISCUSSION

The goal of the present study was to investigate cognitive performance and emotional reactions during infant crying in a childless adult sample. Results showed that participants made the most mistakes in a working memory task when listening to infant crying, compared to other disturbing noises. Furthermore, participants showed significantly more negative emotions when listening to infant crying as compared to listening to silence, traffic, or an excavator noise. In addition, participants showed significantly less positive emotions when listening to infant crying compared to silence and a dinner conversation. Negative emotions, but not positive emotions, significantly explained the number of correct trials in all noise conditions, except the Infant crying condition. Personality was not related to individual differences in cognitive performance or emotional reactions while listening to infant crying. However, participants high on empathy had fewer positive emotions over all conditions than participants low on empathy.

In sum, young childless adults showed cognitive disturbance and negative emotional reactions in response to infant crying. This was assumed by previous research to reflect attention and promote prompt reactions and caregiving efforts from parents. The current results suggest that, besides psychological and physiological changes during the transition to parenthood (e.g., Rilling, 2013; Seifritz et al., 2003; Swain et al., 2011), other mechanisms might be relevant for reactions to infant crying. It is possible that humans, regardless of parenting status, have the tendency to care for infants in distress. This notion is further supported by our findings that personality characteristics were not related to cognitive performance or negative emotional reactions when listening to infant crying. However, this is purely speculative and needs to be confirmed empirically.

A history of living in extended family groups (Brewer & Caporael, 2013), which allowed for alloparenting (i.e. care for offspring by someone other than the parent; Burkhart, Hrdy, & van Schaik, 2009) may have shaped this human predisposition to react to infant crying. Evolutionary psychologists suggest that humans evolved as so-called cooperative breeders, which is a reproductive strategy to ensure better chance of survival for the infant and a longer life for the mother (Hrdy, 2007). The mother might have left her offspring for short periods in care of alloparents, which allowed her to conserve energy and reproduce again sooner (Hrdy, 2007). Being predisposed to react to infant crying, regardless of parental status, might have been evolved in order to increase the reproductive capabilities of the human species.

Interestingly, in all noise conditions of the working memory task, except the Infant crying condition, the more negative emotions the participant felt, the worse the cognitive performance. There is evidence in the literature that emotions and cognition influence each other. For example, various studies have found that mood influences judgments (e.g. Bodenhausen, Sheppard, & Kramer, 1994), but research also showed that emotions can be regulated by cognition, e.g. cognitive reappraisal (Gleitman, Reisberg, & Gross, 2007). Our results are in line with Pessoa (2009), who argued that many networks of brain areas, which are neither specifically affective nor cognitive, work together in order to support complex cognitive-emotional behaviors.
Earlier studies have argued that both negative emotions and cognitive disturbance in reaction to infant crying would lead to prompt and appropriate caregiving (Chang & Thompson, 2011; De Pisapia et al., 2013; Morsbach et al., 1986). Our hypothesis was that both these reactions would occur in our participants, possibly because they share the same underlying mechanism designed to prompt caregiving. This was supported by the fact that, in a previous study, listening to infant crying improved motor performance, which was also argued to reflect facilitation in caregiving (Parsons et al., 2012). However, our results showed that the negative emotions after a working memory task were not related to cognitive performance in this task when participants were subjected to infant crying during the task.

Thus, given that cognitive performance during infant crying was not explained by negative emotions, infant crying is suspected to affect working memory by mechanisms other than negative emotions. According to Baddeley (1992), working memory has three subcomponents: a central executive, which controls attention, a visuospatial sketch pad that manipulates visual images, and a phonological loop, which handles speech-based information. All three subcomponents are responsible for performance on the working memory test, and infant crying might have affected the first two. For example, infant crying might direct participant attention elsewhere (i.e., influence the central executive), and/or interrupt processing of the visual images from the working memory test (i.e., influence the visuospatial sketch pad). Future neuroimaging studies are needed to see how infant crying affects participant brains during a working memory task.

A secondary finding was that participants showed comparable negative emotions when listening to Infant crying as to Dinner conversation noise. However, participants made significantly less mistakes when listening to the Dinner conversation compared to Infant crying. An explanation for this finding might come from research on social exclusion. According to Baumeister, Brewer, Tice, and Twenge (2007), social exclusion is a threat to the human need to belong: the feeling to be part of a group, which serves protection, survival and reproduction. When people are rejected, they feel emotional distress and are less likely to function in intelligent ways, such as in IQ and reasoning tests, but they remain competent in automatic information-processing tasks, such as relatively simple memory tests. In our study, participants might have felt excluded from a social group, namely the people having the dinner conversation, which might have made them feel more negative. However, their cognitive performance might have been unaffected by the feeling of social exclusion.

This study has several strengths. First, we used an experimental design rather than naturalistic observation, which allowed for control of potential confounding variables. Second, we studied emotions and cognitive performances as reactions to infant crying simultaneously, in a relatively large sample (N=120) of childless adults. However, our study sample of university students might not be representative of the whole population of young childless adults. Related to this might be the fact that we had more female than male participants. This design was not ideal to test gender differences, so we decided to include gender as a control variable. Although the gender distribution was not statistically problematic for our analyses, future research should aim to recruit an equal number of male and female participants, as well as young adults from less educated backgrounds. Other limitations are that we included only one measure of cognitive performance, i.e. working memory, while cognition spans other areas such as attention and executive functions as well, and that the internal consistencies of the empathy and positive emotions scales were not very high. Finally, due to technical difficulties, the heart rate data we collected from all participants was not readable. We therefore have no results concerning physiological reactions to infant crying. Future studies should include a valid and reliable measure of heart rate, since physiology has been shown to play a role in different reactions to infant crying (Leerkes et al., 2015).

Especially interesting for future research would be to test the effects of infant crying on more ecologically valid cognitive functions, e.g. attention deployment tasks, that mimic the realities of a (multitasking) caregiver. Also, including samples of parents-to-be and new parents would aid in understanding possible strengthening effects or changes in emotions and cognition in relation to infant crying. It might be possible, for example, that humans have a tendency to react to infant crying, but that this tendency is strengthened by pregnancy.

To conclude, negative emotions and cognitive disturbances may be joint processes in adult responses to infant crying. Moreover, these responses to infant crying are apparently unrelated to personality and not limited to parents, suggesting a broadly present human emotional and cognitive response to infant crying, that may underlie a general human predisposition to care for infants in distress.
REFERENCES


Chapter 3

Prenatal Predictors of Postnatal Quality of Caregiving Behavior in Mothers and Fathers

ABSTRACT

Objective. The quality of parental caregiving has been shown to affect children's development from birth onwards. Therefore, it is important to detect parents at risk for low-quality caregiving as early as possible – preferably before birth. Design. Observations of expectant mothers’ and fathers’ behavior when exposed to infant crying were examined as predictors of the quality of caregiving towards their own infant six weeks postpartum. Eighty-eight expectant mothers and 57 of their male partners were tested during the third trimester of pregnancy. Parents were filmed individually while caring for a crying Simulator Infant for 15 minutes; the quality of their caregiving was rated on sensitivity and cooperation (Ainsworth et al., 1978). Also, cognitive interference on a working memory task and the ability to regulate physical force when exposed to infant crying were assessed. When their baby was six weeks old, parents were filmed and rated for sensitivity and cooperation during a 15-minute interaction with their own infant at home. Results. Prenatal quality of caregiving behavior towards a simulator infant significantly predicted postnatal quality of caregiving towards the own infant, in both mothers and fathers. Cognitive interference and the ability to regulate physical force did not predict postnatal quality of caregiving behavior. Conclusions. Expectant parents’ quality of caregiving behavior towards a crying simulator infant predicted both mothers’ and fathers’ postnatal quality of caregiving behavior. Future research is needed to determine whether the simulator infant may be a useful screening instrument and training tool for parenting skills in at risk groups of parents-to-be.

INTRODUCTION

There is abundant evidence showing the importance of high quality of parental caregiving behavior for child development (Vesely, Levine Brown, & Mahatmya, 2013). High quality caregiving can be characterized as care that positively contributes to children’s wellbeing and development (see e.g., Layzer & Goodson, 2006). In this study, we focused on two aspects of parental caregiving behavior that are generally considered key characteristics of high-quality caregiving for young infants, namely sensitivity and cooperation. Sensitivity refers to the extent to which a caregiver timely and adequately responds to the infant’s needs and signals (Ainsworth, Blehar, Waters, & Wall, 1978). Parental sensitivity has been shown to contribute to a broad range of child developmental outcomes (see Helmerhorst, Riksen-Walraven, Vermeer, Fukkink, & Tavecchio, 2014; Mesman & Emmen, 2013). For example, sensitivity is associated with secure infant-parent attachment (Bakermans-Kranenburg, van IJzendoorn, & Juffer, 2003; Birmingham, Bub, & Vaughn, 2017), better regulatory capacities (Sanchez, McCormack, & Howell, 2015), and lower stress levels later in life (Smeekens, Riksen-Walraven, & van Bakel, 2007). Cooperation, the second quality of caregiving behavior we focused upon in this study, refers to the extent to which caregivers adjust their behavior to the infant and do not interfere with the infant's ongoing activity. Parental cooperation has been shown to contribute to children's development beyond parental sensitivity (see Helmerhorst et al., 2014). Given that parental caregiving affects child development from birth onwards (Feldman, Eidelman, & Rotenberg, 2004), and is relatively stable over time (Dallaire & Weinraub, 2005; Else-Quest, Clark, & Tresch Owen, 2011; Hall et al., 2015), it is important to predict the quality of caregiving as early as possible, potentially already during pregnancy. Early detection of parents at risk for low-quality caregiving may open a window for timely interventions designed to strengthening parenting skills. Therefore, the aim of this study is to investigate possible prenatal predictors of the quality of parents’ caregiving in the first weeks after the birth of their infant.

Research has shown a variety of factors predicting low quality of caregiving: demographic factors, such as a very young age (Riva Crugnola, Ierardi, Gazzotti, & Albizzati, 2014) and lower socioeconomic status (McConnell, Breitkreuz, & Savage, 2011); the quality of the parental support system, such as low partner and social support (Bryanton, Gagnon, Hatem, & Johnston, 2009; Jackson, Preston, & Thomas, 2013); parental characteristics, such as ineffective coping styles (Gudmundson & Leerkes, 2012), and insecure attachment (Jones, Cassidy, & Shaver, 2015), and mental health problems, such as depression (Edwards & Hans, 2016; Ngai, Wai-Chi, & Ip, 2010). Previous research has thus used a large variety of measures to predict quality of caregiving behavior. However, to our knowledge, no research used observed prenatal parental behavior as a predictor...
of parental postnatal caregiving behavior.

The present study aimed to fill this gap by observing prenatal parental behaviors that might theoretically predict the quality of their postnatal caregiving. We chose to observe parents particularly while being exposed to infant crying, because there is evidence that infant crying is powerful in eliciting caregiving-related behaviors. Infant crying, which is the most important signal indicating infant’s needs, has been shown to trigger positive caregiving behavior (Zeifman, 2001), but also to elicit negative reactions resulting in low quality caregiving (Leerkes, Parade, & Gudmundson, 2011). Furthermore, research suggests that parental caregiving behavior observed when the infant is crying is a better predictor of later developmental outcomes than parental caregiving behavior observed when the infant is not crying (Leerkes, 2010). In the following paragraphs, we further elaborate on the three types of prenatal parental behaviors we observed as potential predictors of the quality of postnatal parental caregiving, namely (1) the quality of caregiving towards a crying simulator infant, (2) cognitive interference by infant crying, and (3) the ability to regulate physical force when exposed to infant crying.

First, we examined whether the quality of prenatal caregiving behavior while “caring for” a crying simulator infant would predict parental postnatal quality of caregiving behavior. It was shown that participants experience the infant simulator as fairly realistic and comparable to a real baby (Voorhuis et al., 2013). In students, caring for a crying simulator infant, as compared to a non-crying simulator infant, decreased confidence in parenting ability (de Cock, Henrichs, Rijk, & van Bakel, 2015). In mothers, the quality of caregiving behavior towards a simulator infant was strongly correlated with concurrent observations of caregiving quality towards their infant (Bakermans-Kranenburg, Alink, Biro, Voorhuis, & van Uzendoorn, 2015), and could predict the quality of the mother’s future caregiving behavior (Bridgett, Rutherford, & Mayes, 2015).

We expected the prenatal quality of caregiving for the simulator infant to predict the postnatal quality of caregiving, because caregiving quality is relatively stable over time from birth onwards (Dallaire & Weinraub, 2005; Else-Quest et al., 2011).

Next to the quality of caregiving towards the crying simulator infant, parental characteristics that are prerequisites for high caregiving quality can also be expected to predict postnatal quality of caregiving. One basic prerequisite for sensitivity is the extent to which a parent notices infant signals; parents have to be first aware of a baby’s crying, before they can respond with high quality care (Ainsworth et al., 1978). Crying is hypothesized to reach mothers who are not attending to the baby, even when they are mentally focused elsewhere (De Pisapia et al., 2013). For example, mothers were distracted in their performance on a concentration task when listening to a tape of infant crying, as compared to machine noises (Morsbach, McCulloch, & Clark, 1986). We investigated whether the extent to which infant crying interferes with expectant parents’ cognitive performance predicts their postnatal caregiving quality. We assumed that cognitive interference by infant crying would mean that the infant’s signal is detected, triggering caregiving behavior (Chang & Thompson, 2011). The effect on caregiving quality may, however, be either positive or negative. On the one hand, cognitive distraction by an infant’s crying may lead a parent to sensitively soothe the infant. On the other hand, a parent could become irritated by the crying and being distracted, triggering less positive behaviors.

Research has indeed shown that infant crying can lead to negative emotional reactions ranging from sad to irritated to extremely hostile (Leerkes et al., 2011), and the interpretation of the excessiveness of the crying is related to negative parental behavior towards the infant (Reijneveld, van der Wal, Brugman, Hira Sing, & Verlooove-Vanhorick, 2004). Furthermore, an EEG study (Dudek, Faress, Bornstein, & Haley, 2016) showed that infant crying can interfere with the brain’s capacity for parallel processing. The authors suggest that this interference could ultimately lead to attention depletion, hampering high quality caregiving. In sum, the relation between cognitive interference and postnatal caregiving quality might be both positive and negative. This led us to refrain from proposing a hypothesis and to examine the relation between prenatal cognitive interference by infant crying and postnatal quality of caregiving exploratorily.

Another prerequisite for high caregiving quality is the parent’s ability to regulate arousal resulting from infant crying (Leerkes et al., 2011). Upon hearing infant crying, parents initially react with physiological arousal, such as increases in heart rate and blood pressure (Parsons, Young, Parsons, Stein, & Kringelbach, 2012). High levels of arousal, and not being able to downplay initial arousal upon hearing infant crying, is related to low quality caregiving behavior postnatally (Ablow, Marks, Feldman, & Huffman, 2013). Moreover, parents with high risk of physical child abuse were less able to regulate their physical force when exposed to infant crying than low-risk parents (Crouch, Skowronsiki, Milner, & Harris, 2008). The current study does not focus on abusive or harsh parenting, but rather on parenting in a community sample. However, a lack of ability to regulate physical force may also be related to lower quality caregiving in a community sample. For example, parents scoring lower on Cooperation (a scale also used in the present study to measure parenting quality) tend to physically intrude on their infant, by restricting the infant’s movements and using direct physical force (Ainsworth et al., 1978). We examined whether expectant parents’ ability to regulate physical force when exposed to infant crying predicted their postnatal caregiving quality. We expected that parents who were less able to regulate physical force when exposed to infant crying prenatally would show lower postnatal quality caregiving behavior, than parents who were better able to regulate physical force.
On a final note, we examined whether the three prenatal behavioral predictors would predict postnatal quality of caregiving in fathers as well as mothers. During the past decades, fathers have become more frequently involved in child care (Cohen-Bendahan, Beijers, van Doornen, & de Weerth, 2015), and mothers and fathers have become more similar in terms of their roles, and the types of behaviors they engage in with their children (Fagan, Day, Lamb, & Cambrera, 2014). Furthermore, fathers’ parenting has also been shown to affect child development (Malmberg et al., 2016; Ramchandani et al., 2013; Ryan, Martin, & Brooks-Gunn, 2006; Tamis-LeMonda, Shannon, Cabrera, & Lamb, 2004). Therefore, early detection of an increased risk for low quality caregiving is as important in fathers as it is in mothers.

What does prior research suggest about factors predicting caregiving quality in mothers versus fathers? Early models of the determinants of parenting, such as Belsky’s (1984) seminal model, mostly included mothers as participants. Later research showed that the important determinants in those models may also be associated with fathers’ caregiving behavior (e.g., Berg Nordahl, Zambrana, & Forgatch, 2016; Hall et al., 2014; Pelchat, Bisson, Bois, & Saucier, 2003; Prinzie, Stams, Dekovic, Reijntjes, & Belsky, 2009). The three specific prenatal behavioral predictors examined in the present study (i.e. caregiving behavior quality, cognitive interference, and physical force regulation, all in a context of infant crying) have not been examined earlier in relation to postnatal caregiving, neither in mothers nor in fathers. For our first prenatal predictor of postnatal quality of caregiving, i.e., the prenatal quality of parental caregiving, we expected that prenatal quality of caregiving would predict postnatal quality in both fathers and mothers. This is based on a study observing parents with their infants during the first two years after their infant’s birth that found that all observed aspects of parental behavior (including those examined in the present study) showed significant stability for both mothers and fathers from the first week after birth onwards (Hall et al., 2015). For the other two prenatal predictors in our study there is no relevant research evidence, but given the general similarity in the most important determinants of maternal and paternal caregiving, we hypothesized that the prediction of postnatal caregiving quality from the three prenatal predictors would not be different for mothers and fathers.

In sum, we examined whether the quality of mothers’ and fathers’ caregiving behavior towards their infant (operationalized as sensitivity and cooperation) could be predicted prenatally from (1) the quality of their caregiving behavior when interacting with a crying simulator infant, (2) the extent to which infant crying interfered with their cognitive performance, and (3) their ability to regulate physical force when exposed to infant crying. We hypothesized that higher postnatal quality of caregiving behavior would be predicted by higher quality of caregiving behavior towards the simulator infant and by a greater ability to regulate physical force when exposed to infant crying. The predictive effect of cognitive interference when exposed to infant crying on postnatal quality of caregiving behavior was examined exploratorily. Finally, we expected the prediction of postnatal caregiving quality from the three prenatal predictors not to differ between mothers and fathers.

METHOD

Participants

Participants are part of the BINGO (Dutch acronym for Biological Influences on Baby’s Health and Development) study; a longitudinal study examining prenatal predictors of parental caregiving behavior and infant health. This study was approved by the ethical committee of the Faculty of Social Sciences of the Radboud University [ECSW2014-1003-189]. Participants signed up via folders that were handed out in midwife practices and pregnancy courses around the region Arnhem-Nijmegen (the Netherlands). To facilitate a higher number of participants, mothers were allowed to participate alone and fathers were encouraged, but not required, to participate.

Initial prenatal inclusion criteria were: absence of drug use, no excessive alcohol use, sufficient mastery of the Dutch language, and a healthy pregnancy. In total, 88 expectant mothers and 57 of their partners met the inclusion criteria and signed informed consent. Mothers participated alone when the father had no interest (n=7), had no time (n=19), was a donor (n=2), was known at the university (n=1) or refused without reason (n=2). Independent samples t-test, t=-3.13, p=.002, revealed that mothers who participated alone, M=4.63, SD=2.00, scored higher on prenatal quality of caregiving towards the simulator infant compared to mothers who participated with their partner, M=3.35, SD=1.63. For all other study variables there were no differences between mothers who participated alone and mothers who participated with their partner.

Postnatal exclusion criteria were: complications during pregnancy (after initial contact), prematurity (gestational age ≤ 35 weeks), birth weight < 2500 grams, 5-minute Apgar score < 7, and child anomalies. One infant was born prematurely and one infant with brain damage; these two families were excluded from further analyses. The final sample thus consisted of 86 mothers and 56 of their partners. Seven families stopped participation after birth due to personal reasons. For more information about the participants, see Table 1. Participants that stayed in the study showed higher prenatal quality of caregiving, M=3.73, SD=1.82, than participants who dropped out, M=3.00, SD=1.02; t=2.14, p=.045. Infants (41 boys, 38 girls) were born full-term, M\textsuperscript{gestational age}=39.77, SD=1.52, with an average birth weight of 3531.07 grams, SD=428.43.
### Table 1

<table>
<thead>
<tr>
<th>Participant characteristics</th>
<th>Mothers</th>
<th>Fathers</th>
<th>t-test for the difference between mothers and fathers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participant characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age*</td>
<td>31.62 (86)</td>
<td>3.78 (25.13 - 40.82)</td>
<td>32.93 (56)</td>
</tr>
<tr>
<td>Place of birth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>n= 83</td>
<td>n= 51</td>
<td></td>
</tr>
<tr>
<td>Elsewhere</td>
<td>n= 3</td>
<td>n= 6</td>
<td></td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University education</td>
<td>n= 41</td>
<td>n= 27</td>
<td></td>
</tr>
<tr>
<td>Professional higher education</td>
<td>n= 31</td>
<td>n= 16</td>
<td></td>
</tr>
<tr>
<td>Vocational education</td>
<td>n= 14</td>
<td>n= 13</td>
<td></td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>n= 83</td>
<td>n= 55</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>n= 3</td>
<td>n= 1</td>
<td></td>
</tr>
<tr>
<td>Working hours*</td>
<td>&lt;30 hours</td>
<td>n= 60</td>
<td>n= 55</td>
</tr>
<tr>
<td>30-38 hours</td>
<td>n= 23</td>
<td>n= 0</td>
<td></td>
</tr>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity*</td>
<td>First-time</td>
<td>n= 73</td>
<td>n= 44</td>
</tr>
<tr>
<td>Second-time</td>
<td>n= 13</td>
<td>n= 12</td>
<td></td>
</tr>
<tr>
<td>Predictor variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive interference</td>
<td>(-.03 (85))</td>
<td>.04 (49)</td>
<td>.07 (-.21 - .12)</td>
</tr>
<tr>
<td>Physical force†</td>
<td>2.97 (.22)</td>
<td>2.92 (79)</td>
<td>2.60 (47)</td>
</tr>
<tr>
<td>Prenatal quality of caregiving</td>
<td>.88 (.22)</td>
<td>3.84 (80)</td>
<td>3.39 (45)</td>
</tr>
<tr>
<td>Outcome variable</td>
<td>- .13 (.22)</td>
<td>5.45 (77)</td>
<td>4.41 (47)</td>
</tr>
</tbody>
</table>

Note. -tests for the difference between mothers and fathers: \( t_p < .10 \), \( *= p < .05 \)

Procedure

After giving consent, participant filled in a questionnaire concerning their demographic information. Participants then visited the laboratory during the third trimester of pregnancy, \( M=34.54 \) weeks, \( SD=6.02 \) weeks, during the late afternoon or in the evening (after 15:00). Participants were blind to the exact goals of the tasks. If the mother and father both participated, they were tested separately on all tasks, and with the three tasks always in the same order: cognitive interference, physical force, and infant simulator. The first task was the working memory task to measure cognitive interference by infant crying. The N Back task (Kircher, 1958) was used, with 50 trials in total; first without sound, and then four times while exposed to the sound of a crying infant, an excavator, a dinner conversation, and traffic noises (see Hechler, Beijers, & de Weerth, 2015, for more details). The order of the sound conditions was counterbalanced.

The second task comprised the hand dynamometer task (Crouch et al., 2008) to assess participants’ ability to regulate physical force when exposed to infant crying. Participants were exposed to the sound of an excessively crying infant while being prompted three times to squeeze the handgrip, with an interval of 20 seconds. On each of these three trials, participants had to squeeze as hard as they could for three seconds, and then squeeze with half of their strength for three seconds. Before starting, participants practiced in silence until they were able to squeeze with half their strength on at least five consecutive trials.

Lastly, in the third task a simulator infant was used to elicit prenatal caregiving behavior (RealCare Baby; Realityworks, Eau Claire, WI, USA). The simulator infant resembles a real infant between the age of 0 and 3 months in appearance, weight, and size as well as in expressed needs. Like a real infant, the simulator starts fussing to express a need. Fussing eventually turns into crying if the need is not met. Participants were introduced to the simulator infant and instructed to imagine that the simulator infant was their own infant and that they were at home. The experimenter explained that the simulator infant reacts like a real infant and demonstrated the feeding function (giving the bottle when the infant started fussing). The simulator was then handed to the participant. Unbeknownst to the participant, the simulator only responded to a special chip worn by the experimenter, who left the room. The observation room included two cameras, a cot, a changing table, toys, a rocking chair, a feeding bottle, and a second diaper. Participants were subjected to three periods of fussing and unsoothable crying of an average 4.25 minutes each (Cochran, Dailey, Oddi, Bridgett, & Mayes, 2010). Prolonged fussing and unsoothable crying were chosen in order to trigger caregiving behavior of the expectant parents (Ainsworth et al., 1978). The interaction
between the participant and the simulator was video recorded. After 15 minutes, the experimenter entered the room and asked the participants to fill in two manipulation check questions on a 7-point scale, namely: (1) how difficult they found it to interact with the simulator infant as if it were real, and (2) how seriously they performed the task. Afterwards, participants were debriefed carefully and the experimenter explained that the simulator had not been responding to their soothing attempts due to our manipulation.

When the infant was 6 weeks old, M = 6.78 weeks, SD = 0.82, parents were visited at home during the late afternoon or in the evening (after 15:00). Mothers and fathers were asked to undress their infant, change the diaper, and dress the infant again, interacting with their infant as they normally would. Changing an infant at this age constitutes a mild physical stressor that may elicit crying and fussing (Jansen, Beijers, Riksen-Walraven, & de Weerth, 2010), and infants are usually more fussy and cry more in the evening than in the morning (Barr, Trent, & Cross, 2006). This procedure was chosen to (1) make the interaction comparable to the prenatal interaction with the simulator infant, and (2) to elicit caregiving behavior by infant crying. The 15-minute interaction was filmed as unobtrusively as possible by the experimenter. When both parents participated, the mother always interacted with the infant first. During the home visit, parents also filled in a questionnaire on depressive symptoms.

Measures

Cognitive interference. When performing the N Back task (Kirchner, 1985), participants were presented a computer screen with a 3 by 3 grid, in which every two seconds one of the cells turned red. On each trial, they were asked to indicate whether the cell that just turned red was the same cell that turned red two trials before. To check whether infant crying was related to fewer correct trials compared to the other sound conditions, a repeated-measures ANOVA with condition as the within-subject variable and infant crying condition as the reference category was performed. Participants had the fewest correct trials in the infant crying condition, as compared to the other sound conditions, F = 32.42, p < .001, see Table 2.

A score for cognitive interference by crying was computed by subtracting the number of correct trials during the infant crying sound condition from the number of correct trials during the no sound condition, and dividing the result by the number of correct trials in the no sound condition. Possible scores ranged from -1 to +1, with a score of 0 meaning no cognitive interference by infant crying, and +1 meaning maximal cognitive interference by infant crying. Negative values indicated more correct trials in the infant crying condition compared to the no sound condition.

### Table 2

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mothers (n = 85)</th>
<th>Fathers (n = 49)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Infant Crying</td>
<td>31.07</td>
<td>1.62</td>
</tr>
<tr>
<td>No Sound</td>
<td>32.25</td>
<td>2.60</td>
</tr>
<tr>
<td>Conversation</td>
<td>33.18</td>
<td>2.47</td>
</tr>
<tr>
<td>Traffic</td>
<td>32.12</td>
<td>2.28</td>
</tr>
<tr>
<td>Excavator</td>
<td>33.26</td>
<td>2.12</td>
</tr>
</tbody>
</table>

Note. Mothers’ df; error df = 1,84; fathers’ df; error df = 1,48; As Infant Crying condition is the reference category, corresponding F and p values are missing.

Regulating physical force. Following the procedure by Crouch et al. (2008), scores on the hand dynamometer task were computed per squeeze trial by dividing the half squeeze force by the full squeeze force. Scores equal to or below .50 indicated that participants squeezed with half or less than half their strength, whereas scores higher than .50 indicated that participants used more than half of their strength. Scores above .50 can be interpreted as a more hostile behavioral response. For each of the three trials, the scores were dichotomized; a score of 1 when more strength was used, and a score of 0 when half or less strength was used. Subsequently, the three scores were added up, yielding a total score for the ability to regulate physical force - operationalized as using more force than required -, ranging from 0 to 3 (Crouch et al., 2008).

Prenatal quality of caregiving. A check of the manipulation showed that the participants found it neither easy nor difficult to interact with the simulator infant as if it were real, M = 4.44, SD = 1.74, and reported taking the task rather seriously, M = 5.60, SD = 1.13. Paired samples t-test showed that there was no difference in scores between expectant mothers and fathers, t(48) = -1.18, p = .245; t(48) = 0.87, p = .398. The videos from the laboratory visit interaction with the simulator infant were rated for parental sensitivity and cooperation using 9-point rating scales (Ainsworth et al., 1978). These scales have been extensively validated in various cultures and are well applicable for rating the quality of parental behavior with very young infants in natural settings. Sensitivity (versus insensitivity) refers to the extent to which the parent timely and adequately responds to the infant’s needs and signals. Highly sensitive parents are aware of all, including subtle, signals from their infant, accurately interpret these signals, and react to these signals in a prompt and appropriate manner. In contrast, insensitive parents are often unaware of their infants’ signals, either by ignorance or failure to perceive subtle communications, may not understand their infants’ signals, may react inappropriately or late to these
signals, or not react at all (Ainsworth et al., 1978).

Parental cooperation refers to the extent to which caregivers adjust their behavior to the infant and do not interfere with the infant’s ongoing activity. Highly cooperative parents guide the infant’s activity, parent-child interactions are co-determined, and parental interventions occur when a natural break in the infant’s activity occurs. In contrast, interfering parents have no respect for their infant as an autonomous person with valid wishes and activities. Interfering parents often physically overwhelm the infant, by moving him or her around without apparent reason, restrict or restrain the infant’s movements, and use direct physical force (Ainsworth et al., 1978).

The rating scales range from 1=highly insensitive/interfering through 5=inconsistently sensitive/mildly interfering to 9=highly sensitive/cooperative. Trained observers, who were blind to the study goals, independently rated the interactions. About 30% of the videos were scored twice for reliability. Interrater agreement was good, Intraclass correlation=.92 and .88 for sensitivity and cooperation, respectively. Sensitivity and cooperation were highly correlated, r=.88, and therefore averaged as a measure for quality of caregiving. Note that participants scored more frequently below five (36 expectant fathers and 62 expectant mothers) than above five (9 expectant fathers and 18 expectant mothers). A Chi-square test showed that expectant mothers and fathers did not differ in whether they scored below or above five, χ²=0.11, p=.744. Prenatal quality of caregiving was used as a continuous variable.

Postnatal quality of caregiving. The videos from the home visit were rated for sensitivity and cooperation using the same 9-point scales (Ainsworth et al., 1978) that were used prenatally. About 30% of the videos were rated twice for reliability. Interrater agreement was good; Intraclass correlation=.82 for sensitivity and .75 for cooperation. Sensitivity and cooperation were highly correlated, r=.81, and therefore averaged as a measure for postnatal quality of caregiving. Note that participants scored more frequently above five (n=69) than below five (n=55). However, fathers scored more frequently below five (n=29) than above five (n=18), whereas mothers scored more frequently above five (n=51) than below five (n=26). Pearson Chi-square test showed that this difference between mothers and fathers was significant, χ²=9.23, p=.002. Postnatal quality of caregiving was used as a continuous measure.

Control variables. Next to controlling for infant sex, we also included parity as a control variable. Our sample consisted of 117 first-time parents and 25 second-time parents. Independent-samples t-test for all study variables showed a significant difference between first-time and second-time parents in postnatal quality of caregiving behavior, M₁first-time=5.23, SD₁=1.83, M₂second-time=4.35, SD₂=1.59; t=2.16, p=.033.

We also considered including depressive symptoms as a potential control variable, as postnatal depression has high incidence rates, with a prevalence estimated between 5.5 and 33.1 percent (Gaillard, Le Strat, Mandelbrot, Keita, & Dubertret, 2014), and depression has been associated with insensitive parenting behavior (e.g., Waxler, Thelen, & Musik, 2011). Depression was measured with the Edinburgh Postpartum Depression Scale (EPDS; Cox, Holden, & Sagovsky, 1987), a 10-item screening tool for postnatal depression. Given that the level of depressive symptoms was low in the present community sample, M=4.58, SD=4.39, and depression was unrelated to postnatal quality of caregiving, r=.50, p=.593, we did not include depression as a control variable in our analyses.

Missing values

During pregnancy, five mothers and five fathers were videotaped without sound in the laboratory due to technical issues. Six mothers were not able to perform the handgrip dynamometer task due to pregnancy-induced swelling of their hands. For one father, the N Back data were lost due to technical problems. Three fathers performed the handgrip dynamometer task incorrectly and corresponding data were excluded.

After birth, videos of the home visit were lost for one couple due to technical difficulties. Five fathers and one mother did not complete the interaction because the infant was too upset. One couple and five fathers only participated in the postnatal part of this study. The Missing Completely At Random (MCAR) test indicated that these missing values were completely random, χ²=39.75, p=.267.

Statistical analyses

Quality of prenatal caregiving behavior, prenatal cognitive interference by crying, and prenatal regulation of physical force when exposed to infant crying were used as predictors of postnatal quality of caregiving. To deal with the nested nature of the data (mothers and fathers in couples), we used multilevel (hierarchical) linear modelling (MLM), also known as mixed model analysis. Also, MLM is robust for missing data and is unaffected by unequal sample sizes (Tabachnik, & Fidell, 2007). Therefore, there was no need to control for the fact that more mothers than fathers participated, and we could run the analyses on the full data set, including data from participants that stopped participation after birth (i.e., N=142; 86 mothers and 56 fathers).

MLM is conveyed as a set of regression equations. First, the intercepts-only model (a model without predictors) was run in order to check whether a multilevel model was required, by means of the intraclass correlation. The intraclass correlation was 0.27, thus MLM was appropriate. Second, following Tabachnik and Fidell (2007), a build-up strategy was used. Variables were added one by one to the intercept-only model. After each addition, the -2 log likelihood ratio scale after generalized least square estimation was examined. The -2 log likelihood is a determinant of model fit. If model fitness increases, the added variable is kept. If model fitness decreases, the added variable is cut from the model.
The variables were tested in a certain order. First, gender (mother or father) was included as a fixed factor, and then as a random factor. Thereafter, the control variables were added, starting with parity and then infant sex. Then the prenatal predictors were added, i.e., cognitive interference, physical force, and prenatal quality of caregiving. Finally, interaction terms between parental gender and each of the prenatal predictors were added to check for differences between mothers and fathers. The final model is presented in the results. To check the residuals, this model was checked in regression analysis; the residuals showed normality. No outliers were detected. All analyses were done using SPSS 18.0.0.

RESULTS

Descriptive statistics

Table 1 shows the means and standard deviations for participant characteristics and study variables separately for mothers and fathers. Paired samples t-tests showed that mothers, on average, displayed higher postnatal quality of caregiving than their partners, t=2.30, p=.026. Table 3 depicts the correlations between the study variables. Prenatal and postnatal quality of caregiving were positively correlated for the group as a whole, r=.25, p<.01. In mothers, prenatal and postnatal quality of caregiving were positively correlated, r=.23, p=.05, whereas in fathers it was not, r=.22, p=.182. Furthermore, there was a significant and positive correlation, r=.25, p=.033, between postnatal quality of caregiving and cognitive interference in mothers. There were no other significant correlations.

Predicting postnatal quality of caregiving

The best fitting multilevel model is presented in Table 4. The model fit improved from 494.77 (intercept only model) to 381.05 (final model). The control variables (i.e., parity and infant sex) did not significantly improve model fit. All prenatal predictors significantly improved model fit.

Table 4 shows that the quality of postnatal caregiving was significantly higher for mothers than fathers, t=2.75, p=.008. As for the prenatal predictors in this model, higher prenatal quality caregiving significantly predicted higher postnatal quality caregiving, t=2.22, p=.028. The other two prenatal predictors –cognitive interference and physical force regulation– did not predict the quality of postnatal caregiving. As expected, the prediction of postnatal caregiving was not different for mothers and fathers, as evident from the non-significant effects of the three prenatal predictors x parental gender interaction terms (not included in the final model in Table 4 as they did not significantly improve model fit).

| &lt;ref id="table"&gt;Table 3&lt;/ref&gt; Correlations between the Study Variables |
|---|---|---|---|
| | 1. | 2. | 3. | 4. |
| **Total sample** | | | | |
| 1. Cognitive interference | - | - | - | - |
| 2. Physical force | -.10 | - | - | - |
| 3. Prenatal quality of caregiving | -.15† | .17† | - | - |
| 4. Postnatal quality of caregiving | .11 | .17† | .25** | - |
| **Mothers** | | | | |
| 1. Cognitive interference | - | - | - | - |
| 2. Physical force | -.06 | - | - | - |
| 3. Prenatal quality of caregiving | -.09 | .13 | - | - |
| 4. Postnatal quality of caregiving | .25** | .12 | .23* | - |
| **Fathers** | | | | |
| 1. Cognitive interference | - | - | - | - |
| 2. Physical force | -.11 | - | - | - |
| 3. Prenatal quality of caregiving | -.28* | .19 | - | - |
| 4. Postnatal Quality of Caregiving | -.12 | .06 | .22 | - |

Table 4 Estimates for the Best Fitting Model Predicting Postnatal Quality of Caregiving

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Estimate</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.97</td>
<td>.888</td>
<td>.028*</td>
</tr>
<tr>
<td>Gender (1=father, 2=mother)</td>
<td>.91</td>
<td>.331</td>
<td>.008*</td>
</tr>
<tr>
<td>Cognitive interference</td>
<td>4.37</td>
<td>2.440</td>
<td>.076</td>
</tr>
<tr>
<td>Physical force</td>
<td>2.21</td>
<td>.293</td>
<td>.476</td>
</tr>
<tr>
<td>Prenatal quality of caregiving</td>
<td>20</td>
<td>.088</td>
<td>.028*</td>
</tr>
</tbody>
</table>

Deviance: 381.048

SE=Standard Error, *p<.05

Note. This final model does not depict infant sex, parity, and the parent gender X predictor interactions, as these variables did not significantly improve model fit, and were therefore discarded from the model.
DISCUSSION

The present study examined whether the quality of parents’ caregiving towards their infant could be predicted from their prenatal behavior when exposed to infant crying. Results showed that, as expected, higher quality caregiving behavior towards a crying simulator infant during pregnancy predicted higher quality postnatal caregiving. Contrary to our expectation, the ability to regulate physical force while exposed to infant crying did not predict postnatal quality of caregiving, just as cognitive interference by infant crying on a working memory task, which we examined exploratory. As expected, the results of the analyses predicting the postnatal quality of caregiving were similar for mothers and fathers, whereas mothers showed higher postnatal quality of caregiving than fathers.

To our knowledge, this is the first study to show that the observed quality of parental caregiving behavior towards an infant can already be predicted during pregnancy through observations of parental behavior towards a crying simulator infant. This quite remarkable continuity in behavior from a simulator to the own infant, and from pregnancy to the postnatal period, supports the notion that parental caregiving behavior is rather stable (Dallaire & Weinraub, 2005; Else-Quest et al., 2011; Hall et al., 2015), and additionally shows that quality of postnatal caregiving can already be estimated in pregnancy, and in both sexes. Mothers showed higher postnatal quality caregiving than fathers, but there was continuity in pre- to postnatal quality of caregiving in both mothers and fathers.

Our finding that mothers displayed higher quality postnatal caregiving than their male partners does not contradict the results of prior research. A review by Lewis and Lamb (2003) showed that findings regarding the quality of mothers’ versus fathers’ parenting behavior are generally inconclusive, with mothers providing similar or higher quality. More recent studies confirm this conclusion, with some studies showing no difference between mothers and fathers in the quality of their parenting behavior (Feldman & Klein, 2003; Tamis-LeMonda et al., 2004), but others showing higher quality parenting in mothers than fathers (Blandon & Volling, 2008; Forbes, Cohn, Allen, & Lewinsohn, 2004; Kochanska & Aksan, 2004; Kwon, Jeon, Lewsader, & Elicker, 2012; Neri et al., 2017; Ryan et al., 2006; Volling, Blandon, & Gorvine, 2006). Thus, our study is in line with a series of recent studies showing that, despite fathers’ increasing involvement in child care in the past decades (Fagan et al., 2014), mothers still demonstrate higher quality interactions with their young children than fathers. What our study contributes to this series is that it is one of the very few comparing fathers’ and mothers’ behavior toward very young infants in an everyday care situation (undressing, diapering, dressing again), whereas most other studies observed parents interacting with older infants and toddlers, mostly in play situations.

It might be argued that fathers may be more interested in play than in care activities (Lewis & Lamb, 2003) and therefore have less experience with their children in care activities, which might explain their lower quality caregiving behavior in the present study. It should be noted, however, that we did not observe specific caregiving skills in this study, but two broad qualities of parental behavior (i.e., Sensitivity and Cooperation). These broad qualities have been shown to be stable in both mothers and fathers across the first two years – with observations in natural care situations in early infancy and play situations in later infancy (Hall et al., 2015), which supports the validity of our measure of postnatal paternal caregiving. Therefore, the present study may be a valuable addition to the growing but still limited database of studies comparing fathers’ and mothers’ parenting behaviors toward their infant. This is an important topic to study, not only because fathers and mothers are becoming more similar in terms of their roles and the amount of time they spend with their children (Fagan et al., 2014), but also because fathers’ parenting has been shown to affect child development, including attachment, from birth onwards (Barker, Elles, & Ramchandani, 2017; Bogels & Phares, 2008; Brand & Klimes-Dougan, 2010).

Our working memory task showed that expectant parents performed worse when exposed to infant crying as compared to silence, indicating cognitive interference by infant crying. These results are similar to what we found previously in a student sample (Hechler et al., 2015). In the present study, cognitive interference by infant crying failed to predict postnatal caregiving quality. Interestingly, in mothers the cognitive interference by infant crying was positively correlated with higher postnatal caregiving quality. Fathers on the other hand, showed a non-significant correlation between both. This may indicate that cognitive distraction by an infant’s crying may lead a mother to attend to and sensitively soothe the infant, whereas this is not the case in the father. Future research, with a higher number of fathers (i.e., for more statistical power), should investigate possible explanations of this result. Other explanations for the lack of association between cognitive interference and postnatal caregiving quality may lie in the nature of the task. For example, our task used different interfering noises next to crying, and this may have produced a less clear picture of how crying interferes with the participant’s cognitive functioning, decreasing the predictive value of this variable with respect to postnatal caregiving behavior. Also, our task used continuous infant crying for 90 seconds as a distracting noise. This may not have been ideal to distinguish between future high or low quality of caregiving, as such a salient sound may have distracted all participants in a similar manner. Based on Ainsworth et al. (1978), the most sensitive parent is alert to an infant’s most subtle, minimal, understated cues whereas the less sensitive parent seems to perceive only the most blatant and obvious communications. The use of more subtle, shorter sounds of crying or fussing may be better for
differentiating cognitive interference between parents-to-be and for predicting future quality of caregiving behavior.

Another unexpected finding is that the use of more physical force than required, when exposed to infant crying, failed to predict lower quality of postnatal caregiving behavior. The study of Crouch et al. (2008) found that parents with high risk of physical child abuse, measured by a self-report questionnaire, used more physical force in response to infant crying than parents with low risk of child abuse. Based on this, we assumed that the ability to regulate physical force when exposed to infant crying would not only differentiate non-abusive versus abusive parents, but also, more in general, parents showing higher-quality versus lower-quality caregiving (as reflected in our broader ratings of sensitivity vs. non-sensitivity and cooperation vs. intrusiveness). Harsh caregiving, if observed, would indeed have resulted in low ratings on our sensitivity and cooperation scales, but the 15-minute observation period was probably too short and not stressful enough for any harsh (and thus also very insensitive and very intrusive) parental behaviors to occur, given that such behaviors occur only very infrequently (Bradley & Lindsay, 1987). That regulating physical force did not predict the quality of caregiving in the present study may also be due to the nature of our sample, which was relatively highly educated. It is possible that a lack of ability to regulate physical force does predict low-quality caregiving, but only for parents in the very low caregiving quality range, that was underrepresented in our relatively low-risk sample. Future research with samples that include more high-risk parents can shed more light on this issue. A final possible explanation for the null finding with regard to regulation of physical force as a predictor is that the handgrip dynamometer may be an unsuitable instrument for use during pregnancy, as at least in women, pregnancy is associated with having less physical strength (Mbada, Adeyemi, Omosebi, Olowokere, & Faremi, 2015). Possibly, the handgrip measure has more predictive value for quality of care in fathers than mothers, in low-educated parents as compared to highly educated parents, and in high risk populations as compared to community samples.

There are some limitations to note. Our group of fathers is somewhat small, as it was more difficult to recruit fathers than mothers. Another limitation is that our sample was mostly highly educated which limits the generalizability of our findings. Future research should target parents with low socioeconomic status, as these parents might be more at risk to show low-quality parenting behavior (Euser, van IJzendoorn, Prinzie, & Bakermans-Kranenburg, 2011). Despite this limitation, our sample showed variation in quality of caregiving behavior, including parents scoring on the lower end. As we showed that postnatal quality of caregiving behavior can be predicted from prenatal caregiving observations, the question remains whether prenatal caregiving behavior is a unique predictor above and beyond the prenatal predictors found in previous studies, including parental attachment status and depression (Edwards & Hans, 2016; Jones et al., 2015; Ngai et al., 2010).

**Implications for application and practice**

This study shows that the quality of caregiving of mothers and fathers toward their young infants can be predicted by observing the quality of their caregiving towards a simulator infant already during pregnancy. This is a promising finding that could open interesting perspectives for future research and clinical applications in the long term. It has been suggested, for example, that an infant simulator could be used to test psychological and physiological correlates of parenting behavior, disentangle effects of maternal, paternal, and infant characteristics on parental sensitivity, and test the efficiency of specific interventions to boost parenting quality (Bakermans-Kranenburg et al., 2015). That would indeed be a very valuable application, but more scientific work needs to be done first before this may be realized. For example, the present study should be replicated in additional samples with more diversity in demographic background and parental characteristics. Additionally, the feasibility of using an infant simulator with different groups of persons at risk could be tested, as well as the validity of the procedure when used with these groups of persons. All in all, the results of the present study ask for further research that may pave the way to valuable practical applications.
REFERENCES


Chapter 4

Prenatal and Postnatal Cortisol and Testosterone are Related to Parental Caregiving Quality in Fathers, but not in Mothers

ABSTRACT

Testosterone and cortisol have both been implicated in human parenting behavior. We investigated the relations between observed quality of caregiving during parent-child interactions and pre- and postnatal testosterone and cortisol levels, in both mothers (N=88) and fathers (N=57). Testosterone and cortisol were measured before and after interaction with an infant simulator (prenatal) and with their own child (postnatal) to index basal levels as well as steroid reactivity to the interaction. Our findings are that in fathers, interactions between cortisol and testosterone are related to quality of caregiving both pre- and postnatally. Prenatally there was a stronger negative relation between T and quality of caregiving in fathers with lower cortisol levels, and postnatally there was a stronger negative relation between cortisol and quality of caregiving in fathers high in testosterone levels. Furthermore, prenatal cortisol levels were related to paternal quality of caregiving during interaction with their own child. In mothers, no associations between quality of caregiving and our endocrine measures were observed. We interpret our findings in the context of hyperreactive physiological responses observed in parents at risk for insensitive caregiving, and in light of the dual-hormone hypothesis. The current findings contribute to the growing literature on the endocrine antecedents of human caregiving behavior.

key words: caregiving quality; hormones; infant; social interaction; parenting

INTRODUCTION

The quality of parental caregiving is a critical factor in a child's cognitive and social-emotional development, with insensitive caregiving practices increasing the risk for developing various types of psychopathology (Gilbert et al., 2009; Keyes et al., 2012; Morris, Lui, & Steinberg, 2003; Pechtel & Pizzagalli, 2011). It is therefore of great importance to understand the underlying factors that bring forth variation in parents' caregiving quality, including endocrine factors (Bos, 2017; Feldman, 2015; Rilling, 2013). In particular, recent studies have shown involvement of the steroid hormones testosterone (T) and cortisol (CORT) in parenting behavior (Bos, 2017). Despite increased attention, our current knowledge is still scarce, in particular with regard to fathers. At the same time, human fathers contribute significantly to parental care across cultures and also have a strong impact on the outcomes of their child's development and wellbeing (Cabrera, Volling, & Barr, 2018; Gray & Anderson, 2010). Including both mothers and fathers, the current study will investigate the associations between T and CORT and quality of caregiving behavior in the prenatal and postnatal periods, and whether prenatal T and CORT can predict postnatal quality of caregiving behavior.

Most studies relating T and CORT to caregiving behavior investigated mothers in the postnatal period, and focused on CORT (e.g., Finegood, Blair, Granger, Hibel, & Mills-Koonce, 2016; Gonzalez, Jenkins, Steiner, & Fleming, 2012; Mills-Koonce et al., 2009). Early postnatal levels of CORT in human mothers have shown positive relations with affectionate infant-directed behavior (Fleming, Steiner, & Anderson, 1987), responsiveness to and attractiveness of baby odor (Fleming, Steiner, & Corter, 1997), and sympathy towards infant crying (Stallings, Fleming, Corter, Worthman, & Steiner, 2001). These apparent positive effects of CORT on maternal caregiving might be restricted to the first days after parturition, since a different pattern emerged when CORT was sampled at later time points. Lower maternal sensitivity was associated with higher CORT baseline at 2-6 months postnatally (Gonzalez et al., 2012; Mills-Koonce et al., 2009) and with reduced CORT downregulation when mothers interacted with their 3-month-old infant (Thompson & Trevathan, 2008). Furthermore, an extensive study by Finegood et al. (2016) showed a negative relation between maternal sensitivity and CORT sampled at 7, 15, and 24 months postnatally. For T, the relation with maternal behavior is much more unclear. One of the few existing studies reported a positive relation between maternal sensitivity and basal T levels, but an opposite relation with diurnal variability in T (Endendijk et al., 2016).

The few studies in fathers have mostly focused on relations between caregiving behavior and T. Basal levels of T generally decline during fatherhood (Gettler, McDade, Feranil, & Kuzawa, 2011). Furthermore, lower paternal T levels were related to more self-
reported parental investment (Mascaro, Hackett, & Rilling, 2013), more responsiveness to infant cues (Storey, Walsh, Quinton, & Wynne-Edwards, 2000), and stronger sympathy in response to infant crying (Fleming, Corter, Stallings, & Steiner, 2002). The studies that measured basal T levels in relation to actual observed paternal behavior showed that higher T was associated with less interactive and social touch when fathers interacted with their 1-6 month-old infants (Gordon, Pratt, Bergunde, Zagoory-Sharon, & Feldman, 2017) and with less interactive behavior when fathers interacted with their 3-to-8-month-old infants (Weisman, Zagoory-Sharon, & Feldman, 2014). No associations were observed between basal T levels and observed parenting when fathers interacted with their 1 year-old (Kuo et al., 2016), and with their 3-to-5-year-old children (Endendijk et al., 2016). However, in the study by Kuo et al. (2016) greater declines in T when fathers observed their infants were related to more sensitive caregiving. With regards to CORT, increased levels have been observed in fathers in response to infant crying (Fleming et al., 2002), and decreased levels when fathers interact with their own toddler (Storey, Noseworthy, Delahunty, Halyford, & McKay, 2011). Furthermore, a study that included fathers of 6-year-old children found CORT responses to parental conflict to be related to more frequent use of psychological control towards the child (Sturge-Apple, Davies, Cicchetti, & Cummings, 2009). The relation with CORT and actual early postnatal paternal behavior is currently unknown, as we are not aware of studies that investigated associations between CORT and quality of paternal caregiving during infancy.

Overall, most parenting studies have focused on the postnatal period, and hardly any data exists on relations between endocrine factors and quality of caregiving during the prenatal period. However, parental caregiving can affect infant development from birth onwards (Feldman, Eidelman, & Rotenberg, 2004). Early, preferably prenatal, detection of parents at risk for low-quality caregiving is therefore of great relevance. Abusive parents, or those at risk for abuse, generally show physiological hyperreactivity to signals eliciting stress, such as persistent infant crying (McCanne & Hagstrom, 1996). In nulliparous adults, physiological endocrine responses to crying have also been related to intended harsh caregiving (Out, Bakermans-Kranenburg, van Pelt, & van Uzendoorn, 2012). Therefore, in the current study we address the question whether prenatal T and CORT responses to infant crying are related to actual quality of caregiving towards the own child. Instead of audio recordings of infant crying, we used an unsoothable crying infant simulator to assess prenatal responses to infant crying and parental caregiving behavior (e.g., Rutherford, Booth, Luyten, Bridgett, & Mayes, 2015; Rutherford et al., 2017; van Anders, Tolman, & Volling, 2012; Vooorthuis et al., 2013).

Also, most studies have focused on basal levels of either CORT or T, and generally have not included both steroid hormones, or endocrine responses to caregiving interactions in their design. Importantly, as predicted by the dual-hormone hypothesis (Metha & Prasad, 2015), effects of T on human social behavior have been shown to be dependent on individuals’ CORT levels. Specifically, the effects of T are generally more pronounced, or only observed, in individuals with low levels of CORT. So far, the dual hormone hypothesis has focused on social behavior such as risk taking, aggression, and dominance behavior (Mehta & Prasad, 2015). In the current study, we included measures of both T and CORT which allowed us to address the interactions between these hormones in relations to human caregiving behavior for the first time. Furthermore, we investigated both baseline levels and reactivity of T and CORT in relation to parental caregiving behavior.

Finally, several studies have investigated relations between endocrine factors and caregiving behavior, but most of these studies have used only self-report measures. In the current study, quality of caregiving will be indexed from observations of interactions with the infant simulator (prenatally) and the own infant (6 weeks postnatally) to increase objectivity. These interactions will be rated for parental sensitivity and cooperation, key characteristics of caregiving quality (Helmerhorst, Riksen-Walraven, Vermeer, Fukkink, & Tavecchio, 2014).

In both mothers and fathers, the following research questions will be investigated: (1) is the quality of caregiving, when interacting with the infant simulator, associated with prenatal T and CORT baseline and reactivity? (2) is the quality of caregiving, when interacting with the own infant, associated with postnatal T and CORT baseline and reactivity?; and (3) is the postnatal quality of caregiving, when interacting with the own infant, associated with prenatal T and CORT baseline and reactivity? Regarding question 2, we hypothesized that postnatal quality of caregiving when interacting with the own child is negatively related to postnatal CORT and T, since positive effects of these steroids were only observed in the first few postnatal days (Fleming et al., 1987; Fleming et al., 1997; Stallings et al., 2001). The novelty of questions 1 and 3, regarding prenatal measurements of endocrine factors and pre- and postnatal parental quality, prohibits strong predictions about directionality of an effect. However, based on the notion of physiological hyperreactivity in insensitive caregivers (McCanne & Hagstrom, 1996), we anticipate negative relations between endocrine responses and quality of caregiving.

METHODS

Participants

Participants are part of the BINGO (Dutch acronym for Biological Influences on Baby’s Health and Development) study, a longitudinal study examining prenatal predictors of parental caregiving behavior and infant health. This study attained approval by the
Participants were introduced to the simulator infant in an observation room. The observation room included two cameras, a cot, a changing table, toys, a rocking chair, a bottle, and a second diaper. Participants were instructed to imagine that the simulator infant was their own infant and that they were at home. The experimenter then demonstrated the feeding function (giving the bottle when the infant started fussing) while explaining that the simulator infant reacts like a real infant. The simulator was then handed to the participant, and the experimenter left the room. The simulator infant immediately started fussing. Unlike during the demonstration, the simulator did not react to the participant’s caregiving attempts, since -unknowingly to the participant- the simulator only responds to a special chip worn by the experimenter. Subsequently, participants were subjected to three cycles of around five minutes each of fussing and crying sounds.

The experimenter entered the room after 15 minutes and participants were asked to fill in two manipulation check questions on a 7-point scale: 1) how difficult they found it to interact with the simulator infant as if it were real, and 2) how seriously they performed the task. Subsequently, participants were carefully debriefed and the experimenter explained that the simulator had not been responding to the participant’s soothing attempts due to the manipulation. Approximately 15 minutes (T2) and 35 minutes (T3) after the end of the interaction with the simulator infant, participants provided saliva samples again. Saliva samples were immediately stored at -20°C.

When the infant was 6 weeks old (M=6.77 weeks, SD=0.82), parents were visited at home. This infant age was chosen because infant crying increases from birth onwards and reaches a peak around 6 weeks of age, also known as the crying peak (Barr, Trent, & Cross, 2006), and crying is known to trigger caregiving behavior (Zeifman, 2001). All visits took place during the late afternoon (after 15:00) or in the evening (M=17:40, SD=01:59). During the home visit, parents were first asked to fill in some questionnaires, and then perform a working memory task and handgrip dynamometer task, not relevant for the current study. Afterwards, parents were asked to undress, change the diaper, and redress their 6-week-old infant, interacting with their infant as they would normally do. For ethical reasons and to make the postnatal parent-infant interactions comparable, they were only carried out when the infant was not overly distressed. The interaction was filmed as unobtrusively as possible by the experimenter and was 15 minutes long (in cases when the parent finished before the 15 minutes he/she was asked to continue interacting with the infant until 15 minutes were completed). Changing an infant at this age constitutes a mild physical stressor that may elicit crying and fussing (Jansen, Beijers, Riksen-Walraven, & de Weerth, 2010). When both parents participated, mothers and fathers separately interacted with their infant, and the mother always interacted with the infant first. Similar to the lab visit, three saliva samples were taken; approximately
45 minutes before the start of the interaction (T1; between the working memory and handgrip dynamometer task), at 15 minutes (T2), and 35 minutes (T3) after the end of the interaction. Immediately after the home visit, saliva samples were transported with a portable freezer and subsequently stored at -20°C. During both visits, the parents were blind to the exact nature of the tasks and the goals of the present study.

**Measures**

**Prenatal quality of caregiving.** The videos were rated for parental sensitivity and cooperation using 9-point rating scales (Ainsworth, Blehar, Waters, & Wall, 1978), ranging from 1=highly insensitive/interfering to 9=highly sensitive/cooperative. Sensitivity is generally considered a key aspect of high-quality caregiving that contributes to a broad range of child developmental outcomes (Helmerhorst et al., 2014). Parental cooperation (versus interference) is another key aspect of high-quality caregiving, which has been shown to contribute to children’s development beyond sensitivity (Helmerhorst et al., 2014). Trained observers, who were blind to the study goals, independently rated the interactions. About 30% of the videos were scored twice for reliability. Interrater agreement was good (ICC=.92 and .88 for sensitivity and cooperation, respectively). Sensitivity and cooperation were highly correlated (r=.88), and therefore averaged as a measure for quality of caregiving.

**Postnatal quality of caregiving.** The videos were rated for sensitivity and cooperation using the same 9-point scales (Ainsworth et al., 1978) that were used prenatally. About 30% of the videos were rated twice for reliability. Interrater agreement was good (ICC=.82 and .75 for sensitivity and cooperation, respectively). Sensitivity and cooperation were highly correlated (r=.81), and therefore averaged as a measure for postnatal quality of caregiving.

**Cortisol.** Saliva samples were analyzed at the University Medical Center of Utrecht University, the Netherlands. Saliva was thawed and assayed. CORT in saliva was measured without extraction using an house competitive radio-immunoassay employing a polyclonal anticoortisol-antibody (K7348). [1,2- 3H(N)]- Hydrocortisone (PerkinElmer NET396250UC) was used as a tracer. The lower limit of detection was 1.0 nmol/l, and inter-assay variation was <7% at 3.3-30 nmol/l (n=80). Intra-assay variation was <4% (n=10).

**Testosterone.** After determination of CORT, the saliva samples were sent to Nagasaki, Japan, and were analyzed in the Department of Neurobiology & Behavior of Nagasaki University. The concentration of salivary T in each sample was assayed by enzyme immunoassay (EIA) using a commercially available kit (Salimetrics Europe Ltd., Suffolk, UK). The sample was first thawed, centrifuged at 1500 × g for 15min, and the aqueous layer was aliquoted for assay. The cumulative intra-assay CV was <5% in samples assayed in the Nagasaki University lab. The assay kit has an analytical sensitivity of <1.0 pg/ml. We checked that the optical density of 1.0 pg/ml concentration could be reliably distinguished from a concentration of zero. The information about the recovery and specificity of the kit can be found in the EIA kit online manual.

**Control variables.** The following variables were included as control variables, as they have been demonstrated important in the relation between CORT, T and parental behavior (Bos, 2017; Saltzman & Maestripieri, 2011; Storey & Ziegler, 2016): parity, educational level, and parental age. Moreover, in the analyses concerning prenatal quality of caregiving towards the simulator infant, we also controlled for the reported difficulty and seriousness in interacting with the simulator infant. Control variables that did not improve the model significantly, were removed from the analyses.

**Statistical analyses**

Mothers and fathers were analyzed separately. Because of the longitudinal design (CORT and T were examined three times during pregnancy and three times during the postnatal period), multilevel (hierarchical) linear modeling (MLM), also known as a mixed model analysis, was used. This way it was possible to investigate both T and CORT baseline and reactivity (by including time and time quadratic), and their associations with quality of caregiving. Moreover, mixed model analyses are robust for missing data and are unaffected by unequal sample sizes (Tabachnik & Fidell, 2007). The parent was the level 2 identifier, and the outcome and predictors were the level 1 variables.

MLM is conveyed as a set of regression equations. First, the intercepts-only model (a model without predictors) is run to check whether a multilevel model is required, by means of the intraclass correlation. The intraclass correlations for the mother multilevel analyses were 0.52 for prenatal CORT, 0.65 for postnatal CORT, 0.55 for prenatal T, and 0.54 for postnatal T. The intraclass correlations for the father multilevel analyses were 0.76 for prenatal CORT, 0.65 for postnatal CORT, 0.69 for prenatal T, and 0.61 for postnatal T. Thus, multilevel analyses were appropriate.

Second, following Tabachnik and Fidell (2007), a build-up strategy was used. To the intercept-only model, variables were added one at a time. After each addition, the -2 log likelihood ratio scale after generalized least square estimation was examined. The -2 log likelihood is a determinant of model fit. If model fit increases, the added variable is kept. Time and quadratic time were entered first to check which time model proved a better fit and to investigate T and CORT reactivity to the interaction tasks. Thereafter, the control variables were added one by one, followed by the quality of caregiving predictors. Then, interaction terms between quality of caregiving and time, and between quality of caregiving and quadratic time, were added to investigate, respectively, whether T and CORT reactivity to the interaction tasks were predicted by quality of caregiving. Finally,
interaction terms between quality of caregiving and T or CORT, and between quality of caregiving, T or CORT and time were added. This way, the dual-hormone hypothesis was investigated by testing whether the relation between one hormone and quality of caregiving was dependent on the level of the other hormone.

To answer question 1 (i.e., whether the quality of caregiving, when interacting with the infant simulator, is associated with prenatal CORT and T baseline and reactivity), two multilevel models were built: (1) prenatal quality of caregiving (controlled for parity, educational level, age, difficultness, seriousness, and prenatal T) predicting prenatal CORT levels, and (2) prenatal quality of caregiving (controlled for parity, educational level, age, difficultness, seriousness, and prenatal CORT) predicting prenatal T levels.

To answer question 2 (i.e., whether the quality of caregiving, when interacting with the own infant, is associated with postnatal T and CORT baseline and reactivity), two multilevel models were built: (1) postnatal quality of caregiving (controlled for parity, educational level, age, and postnatal T) predicting postnatal CORT levels, and (2) postnatal quality of caregiving (controlled for parity, educational level, age, and prenatal CORT) predicting postnatal T levels. Lastly, to answer question 3 (i.e., whether the quality of caregiving when interacting with the own infant is associated with prenatal T and CORT), postnatal quality of caregiving behavior was added as a predictor to the multilevel models predicting prenatal T and CORT. Postnatal T and CORT levels were included in these last multilevel analyses as well, to investigate whether prenatal T and CORT levels were uniquely related to postnatal quality of caregiving.

The best fitting models are presented in the Results. A check of the VIF and Durbin Watson values indicated normality (see Table 3 and 4) and no outliers were detected.

RESULTS

Preliminary Analyses

Missing values. During pregnancy, five mothers and five fathers were video recorded without sound in the lab due to technical issues, and these videos were not rated for quality of caregiving behavior. After birth, videos of the interaction with the 6-week-old infant were lost for one couple due to technical difficulties. Five fathers and one mother did not complete the interaction because their infant was too upset. The saliva samples of mothers that had used antibiotics during pregnancy (N=2) and after birth (N=2) were excluded from hormonal analyses. As multilevel analyses are robust for missing values (Tabachnik & Fidell, 2007), missing values were not imputed.

Manipulation check. Participants found it neither easy nor difficult to interact with the simulator infant as if it were real (Difficulty; M=4.44, SD=1.74) and reported taking the task rather seriously (Seriousness; M=5.60, SD=1.13).

Descriptive statistics. Table 1 shows the means and standard deviations for the control variables and pre- and postnatal quality of caregiving separately for mothers and fathers. Paired samples t-tests showed that mothers, on average, displayed higher postnatal quality of caregiving than their partners (t=3.10, p=.002).

Table 1
Descriptive Statistics, Separately for Mothers and Fathers

<table>
<thead>
<tr>
<th>Control Variables</th>
<th>Mothers</th>
<th>Fathers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University</td>
<td>41 (48)</td>
<td>16 (29)</td>
</tr>
<tr>
<td>College</td>
<td>30 (35)</td>
<td>3 (5)</td>
</tr>
<tr>
<td>High school</td>
<td>1 (1)</td>
<td>10 (18)</td>
</tr>
<tr>
<td>Job training</td>
<td>13 (16)</td>
<td>13 (15)</td>
</tr>
<tr>
<td>Parity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-child</td>
<td>73 (85)</td>
<td>44 (79)</td>
</tr>
<tr>
<td>Second-child</td>
<td>13 (15)</td>
<td>12 (21)</td>
</tr>
<tr>
<td>Age</td>
<td>31.62 (3.79)</td>
<td>32.93 (4.16)</td>
</tr>
<tr>
<td>Difficulty&lt;sup&gt;1&lt;/sup&gt;</td>
<td>4.35 (1.66)</td>
<td>4.60 (1.86)</td>
</tr>
<tr>
<td>Seriousness&lt;sup&gt;1&lt;/sup&gt;</td>
<td>5.70 (1.11)</td>
<td>5.64 (1.15)</td>
</tr>
<tr>
<td>Quality of caregiving</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prenatal</td>
<td>3.84 (1.87)</td>
<td>3.39 (1.56)</td>
</tr>
<tr>
<td>Postnatal&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5.45 (1.74)</td>
<td>4.41 (1.76)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Difference between mothers and fathers is p=.026, as indicated by paired sample t-test.
<sup>b</sup>How difficult expectant parents found it to interact with the simulator infant as if it were real.
<sup>1</sup>How seriously expectant parents performed the interaction with the simulator infant.

Prenatal and postnatal T and CORT responses to the interactions

Figure 1 and 2 show T and CORT baseline and reactivity of mothers and fathers to the prenatal interaction task with the simulator infant and the postnatal interaction task with the 6-week-old infant. To test prenatal and postnatal parental T and CORT reactivity to the interaction tasks, multilevel time-only models of T and CORT were investigated (see Table 2). For mothers, there was a significant positive effect of time on both prenatal T (p=.014) and CORT (p=.027), meaning that maternal T and CORT levels increased in response to the interaction with the simulator infant. After birth, there was a significant negative effect of time and quadratic time on CORT but not T. In response to the
interaction with the own infant, maternal CORT, but not T, levels decreased significantly.

In fathers, there was a significant negative effect of time (\(p=.026\)) and a positive effect of time quadratic (\(p=.029\)) on prenatal T. In reaction to the interaction with the simulator infant, paternal T levels first increased and subsequently decreased. There was a significant positive effect of time (\(p=.003\)) and a negative effect of time quadratic (\(p=.002\)) on postnatal T. In response to the interaction with the own infant, paternal T levels decreased. There was a significant negative effect of time on prenatal (\(p=.012\)) and postnatal (\(p<.001\)) CORT in fathers. Fathers’ pre- and postnatal CORT levels decreased in response to the interaction with the simulator infant and their own infant.

Figure 1. Maternal and paternal cortisol responses to the interaction with the Simulator Infant and the own infant. Error bars represent standard error of the mean.

Figure 2. Maternal and paternal testosterone responses to the interaction with the Simulator Infant and the own infant. Error bars represent standard error of the mean.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>Mothers</th>
<th>Fathers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td><strong>PRENATAL</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cortisol</td>
<td>Intercept</td>
<td>17.94</td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>Time Quadratic</td>
<td>--</td>
</tr>
<tr>
<td>Deviance</td>
<td></td>
<td>1258.398</td>
</tr>
<tr>
<td>Testosterone</td>
<td>Intercept</td>
<td>4.58</td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Time Quadratic</td>
<td>--</td>
</tr>
<tr>
<td>Deviance</td>
<td></td>
<td>1851.911</td>
</tr>
<tr>
<td><strong>POSTNATAL</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cortisol</td>
<td>Intercept</td>
<td>35.54</td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>Time Quadratic</td>
<td>--</td>
</tr>
<tr>
<td>Deviance</td>
<td></td>
<td>846.190</td>
</tr>
<tr>
<td>Testosterone</td>
<td>Intercept</td>
<td>26.90</td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>Time Quadratic</td>
<td>--</td>
</tr>
<tr>
<td>Deviance</td>
<td></td>
<td>1352.739</td>
</tr>
</tbody>
</table>

Main Multilevel Analyses

Is prenatal quality of caregiving behavior associated with prenatal T and CORT baseline and reactivity? Mothers. The best fitting multilevel models for maternal prenatal T and CORT are presented in Table 3. For prenatal T, model fit improved from 1881.07 (intercept only model) to 1698.16 (final model). The control variables educational level and prenatal CORT improved model fit. There was a significant positive effect of prenatal CORT (\(p<.001\)). Higher CORT levels were related to higher T levels. There was no significant effect of prenatal quality of caregiving, the other control variables (i.e., parity, age, difficulty, and seriousness), or any of the interaction terms.

The model fit for prenatal CORT improved from 1279.09 (intercept only model) to 921.29 (final model). Of the control variables, parity, educational level, and prenatal T, and the interaction between parity and prenatal quality of caregiving, significantly improved model fit, whereas the other control variables (i.e., age, difficulty, and seriousness) did not.
Also, the interaction between prenatal quality of caregiving and T, and the interaction between prenatal quality of caregiving, time, and T, did not improve model fit. There was a significant positive association between prenatal CORT and prenatal T. Higher T levels were related to higher CORT levels ($p<.001$). However, parity, educational level, and the interaction between parity and prenatal quality of caregiving, were not significantly related to prenatal CORT. There were no further significant effects on prenatal CORT of prenatal quality of caregiving, or any of the interaction terms.

**Fathers.** The best fitting multilevel model for paternal prenatal T and CORT are presented in Table 4. For prenatal T, model fit improved from 1012.68 (intercept only model) to 949.71 (final model). Of the control variables, prenatal CORT and seriousness improved model fit. There was a significant positive relation between prenatal CORT and prenatal T ($p<.001$). Higher CORT levels were related to higher T levels. Additionally, there was a significant negative effect of seriousness on prenatal T ($p=.022$). Taking the interaction with the simulator more seriously was related to lower T levels. Finally, there was a significant negative effect of the interaction between sensitivity and CORT on prenatal T ($p=.035$). To qualify this interaction, a median split was performed creating a low and high CORT group, for which we plotted the relation between T and quality of caregiving (see Figure 3).

The figure shows that for the low CORT group, T is more negatively associated with quality of caregiving than for the high CORT group. Although the interaction was significant, the different slopes for the CORT groups are not significantly different from zero ($p=.335$ and $p=.252$, for the low and high CORT group, respectively). There were no significant effects on prenatal T by prenatal quality of caregiving, the other control variables (i.e., educational level, parity, age, and difficulty), or any of the interaction terms.

For prenatal CORT, model fit improved from 700.48 (intercept only model) to 516.01 (final model). Of the control variables, prenatal T improved model fit, with a significant positive effect on prenatal CORT. Higher T levels were related to higher CORT levels ($p<.001$). There were no further significant effects on prenatal CORT of prenatal quality of caregiving, the other control variables (i.e., educational level, parity, age, difficulty, and seriousness), or any of the interaction terms.

**Mothers.** The best fitting multilevel models for postnatal T and CORT are presented in Table 3. For postnatal T, model fit improved from 1356.74 (intercept only model) to 1252.43 (final model). The control variables parity, educational level, and postnatal CORT improved model fit. There was a significant positive effect of postnatal CORT ($p<.001$). Higher CORT levels were related to higher T levels. There was no significant effect of postnatal quality of caregiving, the other control variable (i.e., age), or any of the interaction terms.

For postnatal CORT, model fit improved from 876.75 (intercept only model) to 635.20 (final model). The control variables educational level and postnatal T significantly improved model fit, whereas the other control variables (i.e., parity and age) did not. There was a significant positive effect of postnatal T ($p<.001$). Higher T levels were related to higher CORT levels. There was no significant effect of postnatal quality of caregiving, or any of the interaction terms.

**Fathers.** The best fitting multilevel models for postnatal T and CORT are presented in Table 4. For postnatal T, model fit improved from 1104.49 (intercept only model) to 1013.85 (final model). Of the control variables, postnatal CORT improved model fit, whereas the other control variables (i.e., educational level, parity, and age) did not. There was a significant positive association between postnatal CORT and postnatal T ($p<.001$). Higher CORT levels were related to higher T levels. There was no significant effect of postnatal quality of caregiving, or any of the interaction terms.

For postnatal CORT, model fit improved from 652.57 (intercept only model) to 445.16 (final model). Of the control variables, age and postnatal T improved model fit, whereas the other control variables (i.e., parity and educational level) did not. There was a significant positive effect of postnatal T on postnatal CORT ($p<.001$). Higher T
levels were related to higher CORT levels. There was a significant negative interaction of postnatal quality of caregiving and T on CORT ($p<.001$). To qualify this interaction, a median split was performed creating a low and high T group, for which we plotted the relation between CORT and quality of caregiving (see Figure 4). The figure shows that for the high T group, CORT is more negatively associated with quality of caregiving than for the low T group. Both slopes are significantly different from zero ($p<.01$ and $p<.001$, for the low and high T group, respectively). There was no significant effect of postnatal quality of caregiving or any of the other interaction terms.

![Figure 4](image)

Is postnatal quality of caregiving related to prenatal T and CORT? Mothers. To investigate whether postnatal quality of caregiving is related to prenatal T and CORT, controlling for postnatal T and CORT levels, postnatal quality of caregiving behavior was added as a predictor to the models predicting prenatal T and CORT. The best fitting multilevel models for prenatal T and CORT are presented in Table 3. Maternal postnatal quality of caregiving behavior was unrelated to prenatal levels of T and CORT.

Fathers The best fitting multilevel model for prenatal T and CORT are presented in Table 4. Paternal postnatal quality of caregiving was unrelated to prenatal T levels. For prenatal CORT, model fit improved from 700.48 (intercept only model) to 476.92 (final model). There was a significant negative effect of postnatal quality of caregiving ($p=.05$) on prenatal CORT. Lower postnatal quality of caregiving was associated with higher prenatal CORT levels.
### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>Signif.</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cortisol</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.31</td>
<td>1.00</td>
</tr>
<tr>
<td>Time</td>
<td>0.041</td>
<td>0.46</td>
<td>0.88</td>
<td>0.39</td>
<td>1.00</td>
</tr>
<tr>
<td>Age</td>
<td>0.042</td>
<td>0.08</td>
<td>0.54</td>
<td>0.59</td>
<td>1.00</td>
</tr>
<tr>
<td>Educational level</td>
<td>0.031</td>
<td>0.08</td>
<td>0.40</td>
<td>0.68</td>
<td>1.00</td>
</tr>
<tr>
<td>Testosterone prenatal</td>
<td>0.023</td>
<td>0.03</td>
<td>0.67</td>
<td>0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>Testosterone postnatal</td>
<td>-0.000</td>
<td>0.00</td>
<td>-0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Quality of caregiving prenatal</td>
<td>-0.01</td>
<td>0.01</td>
<td>-1.00</td>
<td>0.31</td>
<td>1.00</td>
</tr>
<tr>
<td>Quality of caregiving postnatal</td>
<td>-0.01</td>
<td>0.01</td>
<td>-1.00</td>
<td>0.31</td>
<td>1.00</td>
</tr>
<tr>
<td>Interaction caregiving*CORT</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Deviance</td>
<td>103.851</td>
<td>1013.851</td>
<td>0.103</td>
<td>0.937</td>
<td></td>
</tr>
</tbody>
</table>

### Testosterone

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>Signif.</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.00</td>
<td>0.14</td>
<td>0.01</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>Time</td>
<td>-0.95</td>
<td>0.40</td>
<td>0.95</td>
<td>0.35</td>
<td>1.00</td>
</tr>
<tr>
<td>Time quadratic</td>
<td>-0.035</td>
<td>0.07</td>
<td>0.50</td>
<td>0.60</td>
<td>1.00</td>
</tr>
<tr>
<td>Testosterone prenatal</td>
<td>0.023</td>
<td>0.03</td>
<td>0.68</td>
<td>0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>Testosterone postnatal</td>
<td>-0.003</td>
<td>0.00</td>
<td>-3.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Quality of caregiving prenatal</td>
<td>0.036</td>
<td>0.03</td>
<td>1.00</td>
<td>0.31</td>
<td>1.00</td>
</tr>
<tr>
<td>Quality of caregiving postnatal</td>
<td>-0.01</td>
<td>0.01</td>
<td>-1.00</td>
<td>0.31</td>
<td>1.00</td>
</tr>
<tr>
<td>Interaction caregiving*CORT</td>
<td>0.015</td>
<td>0.01</td>
<td>1.10</td>
<td>0.29</td>
<td>1.00</td>
</tr>
<tr>
<td>Deviance</td>
<td>949.706</td>
<td>1013.851</td>
<td>0.937</td>
<td>0.35</td>
<td></td>
</tr>
</tbody>
</table>

### DISCUSSION

In the current study we aimed to answer three questions. 1: Is prenatal quality of caregiving behavior associated with prenatal testosterone (T) and cortisol (CORT) baseline and reactivity? 2: Is postnatal quality of caregiving associated with postnatal T and CORT baseline and reactivity? 3: Is postnatal quality of caregiving related to prenatal T and CORT? Furthermore, in light of the dual-hormone hypothesis (Mehta & Prasad, 2015) we also investigated interactions between T and CORT in relation to quality of caregiving. The results show that: (1) for both mothers and fathers, prenatal quality of caregiving behavior was not associated with prenatal T or CORT levels, although in fathers there was a significant interaction between quality of caregiving and CORT on prenatal T. This was driven by a stronger negative relation between T and quality of caregiving in fathers with lower CORT levels. (2) In the postnatal period, quality of caregiving was again unrelated to either postnatal T or CORT levels in both mothers and fathers, although in fathers there was now an interaction of postnatal quality of caregiving and T on CORT. This was caused by a stronger negative relation between CORT and quality of caregiving in fathers high in T levels. Finally, (3) lower quality of postnatal caregiving was associated with higher prenatal CORT levels in fathers, whereas no such association was observed for the mothers. There was no relation between postnatal quality of caregiving and prenatal T levels in fathers or mothers. Thus, in both the prenatal and postnatal period, T and CORT were associated with our index of quality of caregiving through an interaction between the two hormones, which was only observed in fathers. Also, we observed a significant relation between higher prenatal CORT and lower postnatal quality of caregiving in fathers, but not in mothers. To our knowledge, this is the first time that these relations have been observed in fathers, and the question remains what mechanisms underlie these relations.

One of the underlying mechanisms could involve the level of experienced stress during provision of care. Stress has previously been shown to be negatively associated with parenting quality (Deater-Deckard & Panneton, 2017), and abusive parents generally show hyperreactive stress-responses to infant distress (McCanne & Hagstrom, 1996). Also, one study showed that higher levels of α-amylase, a marker for activation of the sympathetic nervous system, were related to intended harsh caregiving in response to infant crying (Out et al., 2012). Therefore, fathers who experience more stress, as reflected in endocrine responses, during a prenatal caregiving interaction which includes infant crying, may be less sensitive and cooperative when interacting with their own child. This relation between prenatal CORT and postnatal quality of caregiving was unaffected by levels of T, unlike the observed relations for fathers in the pre- and postnatal condition separately. In the prenatal session, lower levels of CORT were related to a more negative
relation between T and quality of caregiving. Maybe the negative effect of T on quality of caregiving in fathers was only observed in those less stressed by the infant simulator. This finding seems in line with the dual-hormone hypothesis (Mehta & Prasad, 2015), which states that effects of T are generally observed in participants with relative low levels of CORT. However, postnatally we observed a relation opposite to the predictions of the dual-hormone hypothesis, as CORT was more negatively associated with quality of caregiving for fathers with higher levels of T. Although this is not in line with the dual-hormone hypothesis, such opposite effects have previously been observed in men (Welker, Lozoya, Campbell, Neumann, & Carré, 2014). Furthermore, this finding for the postnatal session is in line with the idea that hyperreactive physiological responses are negatively related to parental sensitivity (McCanche & Hagstrom, 1996). Contextual differences between the two conditions, i.e. caring for an unsoothable infant simulator versus interaction with the own child, might explain the opposite effects in these interactions.

In our study, the negative relations between T and quality of caregiving depended on CORT levels, and this direction of the effects is in line with previous work. In a study by Gordon et al. (2017), father’s T levels were negatively related to the quality of interactive behavior with their 3-to-8-month-old infants. Furthermore, paternal T levels were also negatively related to the quality of interactive and social touch behavior with 1-6-month-old infants (Gordon et al., 2017; Weisman et al., 2014). Since our sample of fathers was relatively small (n=57), we must be careful in interpreting these interactions. However, our findings do point out that when investigating endocrine antecedents of parenting, both T and CORT should be taken into account (Bos, 2017).

The above interpretations in terms of stress sensitivity are supported by findings that intranasal administration of oxytocin, a neuropeptide known to reduce CORT stress responses in a social context (Heinrichs, Baumgartner, Kirschbaum, & Ehler, 2003), facilitate positive social interaction and caregiving quality in fathers interacting with their own children (Naber, van Uzendoorn, Deschamps, van Engeland, & Bakermans-Kranenburg, 2010; Weisman et al., 2014). This proposed mechanism can however not explain why in the current data, postnatal parenting quality was related to prenatal cortisol, but not to postnatal (i.e., concurrent) cortisol. Possibly a third factor underlies both quality of caregiving and altered CORT levels. For example, experienced early life stress can affect both CORT levels and parenting quality (Bos, 2017). Early life stress, such as experienced insensitive caregiving can, depending on the severity and timing, lead to either increased or decreased basal CORT levels, as well as decreased quality of parental caregiving (Bailey, Mill, Oesterle, & Hawkins, 2009; Bos, 2017). Also, parental motivation could be an underlying third factor; fathers that are less motivated for infant caregiving in general are perhaps more stressed when interacting with the crying infant simulator, resulting in higher CORT levels during this interaction, and are also less sensitive when interacting with their own child. In future work, the incorporation of measures for parental motivation (e.g., Buckels et al., 2015) could help to reveal and disentangle these potential underlying factors. Ideally, these questions are addressed in longitudinal work, wherein such assessments can be collected during or even before pregnancy.

Although our findings in fathers are in line with reported negative associations between quality of caregiving in mothers and CORT levels (Finegood et al., 2016; Gonzalez et al., 2012; Mills-Koonce et al., 2009; Thompson, Gupta, Miller, Mills, & Orr, 2004), we did not observe a relation between our endocrine measures and parenting quality in mothers. Several factors might account for the fact that we did not observe such a relation. First, the children included in previous work were generally older (varying from 2 to 24 months) compared to those included in the current study in which the average age of the child was 6.77 weeks. Since positive relations between CORT and maternal sensitivity have been observed soon after delivery (Fleming et al., 1987; Fleming et al., 1997; Stallings et al., 2001), it could be that negative relations between maternal CORT and quality of caregiving appear later. Our study on 6-week-olds may therefore have been in a transition period, in which no clear associations between CORT and quality of maternal caregiving are found. In addition, most of the mothers in the current sample were breastfeeding (76 %), and this is known to reduce endocrine stress-responses (Heinrichs, Neuman, & Ehler, 2002), and to be positively related to maternal sensitivity (Tharner et al., 2012). Thus, breastfeeding might have served as a protective factor obscuring a relation between CORT and parenting quality at the postpartum assessment moment. Neither did we observe a relation between T levels and caregiving behavior in mothers. It might be that such a relation in mothers depends on other endocrine factors not taken into account in the current study. For example, in the study by Gordon et al. (2017), T was shown to affect maternal caregiving behavior, but only in interaction with oxytocin.

Although it was not the primary question for which the study was set up, the overall endocrine responses to the interaction with the infant simulator and to the interaction with the own child is also of interest, especially with respect to the use of the infant simulator to study natural caregiving behavior (Rutherford et al., 2015, 2017; van Anders et al., 2012; Voorthuis et al., 2013). Other studies that looked at endocrine responses to an infant simulator have found that, in young nulliparous women, T levels decrease during interaction with the infant simulator (Voorthuis, Bakermans-Kranenburg, & van Uzendoorn, 2017), whereas in our group of pregnant mothers both T and CORT levels increased during the interaction with the simulator. Whether this difference can be explained by the participant sample (pregnant versus nulliparous women) or by methodological differences (in the study by Voorthuis et al. (2017), the women practiced
two evenings with the simulator) is currently unknown. The same question holds for the data on males, since in our fathers, T levels first increased and subsequently decreased in reaction to the interaction with the simulator infant. This finding corroborates with previous work in which males were exposed to infant cry sounds and in which T levels showed a similar pattern (Fleming et al., 2002), although no infant simulator was used in that study. Studies performed so far with the infant simulator only investigated young nulliparous males, and these studies have failed to show overall increases in T during interaction with the simulator (van Anders, Tolman, & Jainagaraj, 2014; van Anders et al., 2012). Comparing males and females that are expecting a child with nulliparous controls in a similar experimental setup could give more insight into the origin of these disparate findings. Furthermore, compared to the prenatal measures, different responses were observed after birth, when T in fathers, and CORT in both mothers and fathers, declined when interacting with the own child. These differences can however be caused by the differences in experienced stress between caring for a crying simulator and caring for the (non-crying) own infant.

An additional interesting result is that we did observe a significant relation between subjective reports on how seriously the fathers took the interaction with the infant simulator, and paternal T levels during the interaction. Fathers that reported to have taken the interaction less seriously had higher T levels. Perhaps for fathers with higher levels of T, pretending actual care behavior with a doll while being observed is considered a threat to one’s status (Eisenegger, Haushofer, & Fehr, 2011), and is therefore taken less seriously. For example, fathers that report less parental investment and show less sensitivity to infant stimuli also have higher T levels (Mascaro et al., 2013). Such fathers might also feel more uncomfortable in a lab-setting acting out caregiving behavior. Alternatively, fathers with higher T might have more difficulty in empathically imagining the situation as real (van Honk et al., 2011).

Some limitations of the current study need to be addressed. First, although the use of the crying simulator is an innovative approach for studying actual parenting behavior, quality of caregiving assessed by using an unsoothable crying simulator is different from quality of caregiving when interacting with a non-crying own baby. Although this limitation cannot be methodologically solved, it is important to consider as endocrine responses to both situations can reflect different processes. Furthermore, the current findings need to be replicated in other samples as the sample size of the group of fathers was relatively small due to fewer fathers than mothers wanting to participate. Another limitation of the current sample is that it consists of a generally highly-educated sample, which limits the generalizability of the findings. An important question to be addressed in future studies is whether the relation between quality of caregiving and prenatal CORT observed in our sample is also observed in larger community samples.

In conclusion, the current study investigated how prenatal and postnatal endocrine factors are related to quality of caregiving in both mothers and fathers, and provided novel insights into how fathers’ prenatal cortisol concentrations are related to the quality of caregiving for their own infant after birth. Although we only addressed the role of T and CORT in this study, most work so far has focused on only a single endocrine factor (Bos, 2017). Studies that include more factors, such as the work of Gordon et al. (2017), or the current longitudinal study, can give more insight into how different endocrine factors bring forth variations in caregiving. Ultimately, better understanding of the antecedents of the quality of human parenting will allow us to identify profiles for parents at risk and will provide avenues for intervention.
REFERENCES


Prenatal and Postnatal Cortisol and Testosterone are Related to Parental Caregiving Quality in Fathers, but not in Mothers

Chapter 4


Chapter 5

Association Between Fecal Microbiota and Anxiety in Pregnant Women

ABSTRACT

Maternal prenatal psychosocial stress is associated with altered child emotional and behavioral development. One potential underlying mechanism is that prenatal psychosocial stress affects child outcomes via the mother’s, and in turn the child’s, intestinal microbiota. This study investigates the first step of this mechanism: the relation between psychosocial stress and fecal microbiota in pregnant mothers. Mothers (N=70) provided a late pregnancy stool sample and filled in questionnaires on general and pregnancy-specific stress and anxiety. Bacterial DNA was extracted and analysed by Illumina HiSeq sequencing of PCR-amplified 16S ribosomal RNA gene fragments. Associations between maternal general anxiety and microbial composition were found. No associations between the other measured psychosocial stress variables and the relative abundance of microbial groups were detected. This study shows associations between maternal pregnancy general anxiety and microbial composition, providing first evidence of a mechanism through which psychological symptoms in pregnancy may affect the offspring.

Keywords
Women; Intestinal microbiota; Prenatal psychosocial stress; General anxiety; Pregnancy; Mechanism

INTRODUCTION

Accumulating evidence indicates that maternal psychosocial stress during pregnancy may affect child emotional, behavioral and cognitive development, as well as physical health (Zijlmans, Beijers, Riksen-Walraven, & de Weerth, 2017; Graignic-Philippe, Davan, Chokron, Jacquet, & Tordjman, 2014; Beijers, Jansen, Riksen-Walraven, & de Weerth, 2010). Many of these hypothesized effects of prenatal stress appear to be adverse. The mechanisms underlying the relations between prenatal psychosocial stress and child outcomes are only partly understood (Beijers, Buitelaar & de Weerth, 2014). One possible underlying mechanism is that prenatal psychosocial stress affects the child via the mother’s, and in turn the infant’s intestinal microbiota (Beijers et al., 2014). The intestinal bacteria have a central position in human health and disease and are suggested to also play a role in the development of emotion regulation, behavior, and higher cognitive functions (de Weerth, 2017). To our knowledge, this study is the first to investigate an essential part of this potential mechanism, namely the relation between psychosocial stress and fecal microbiota in pregnant mothers.

The proposed mechanism is based on the knowledge that an infant’s intestines are virtually sterile at birth, and that the microorganisms important for the colonization originate mainly from the mother during vaginal delivery (de Weerth, Fuentes, & de Vos, 2013; Grönlund, Grzeskowiak, Isolauri, & Salminen, 2011). If the maternal microbiota is unbalanced, e.g. as a possible result of psychosocial stress, infant intestinal colonization might be altered, with possible consequences for child mental and physical development (Dimmitt et al., 2010; Bäckhed, 2011).

Results from rodent and primate models support the link between prenatal psychosocial stress and offspring intestinal microbiota (Bailey, Lubach, & Coe, 2004; Golubeva et al., 2015; Jašarević, Rodgers, & Bale, 2015). Additionally, one human study found that infant intestinal microbiota from mothers with high prenatal psychosocial stress was characterized by more Proteobacteria, and less Actinobacteria and lactobacilli (Zijlmans, Korpela, Riksen-Walraven, de Vos, & de Weerth, 2015).

Indications that psychosocial stress might be related to fecal microbiota during pregnancy come from a mouse study indicating that stress during pregnancy was associated with changes in the gut microbiota (Jašarević, Howard, Misić, Beiting, & Bale, 2017). For example, mice exposed to a social stressor showed decreased relative abundance of bacteria in the genus Bacteroides and increased relative abundance of bacteria in the genus Clostridium (Bailey et al., 2011). Finally, in a study on healthy non-pregnant female students, the concentration of beneficial lactic acid bacteria was lower during a stressful week (first week of exams) as compared to a low-stress week (beginning of semester) (Knowles, Nelson, & Palombo, 2008).
The current study examined associations between maternal psychosocial stress and intestinal microbiota composition in late pregnancy. Based on the findings by Zijlmans et al. (2015), we hypothesized that mothers with high psychosocial stress would have phylum-level microbial compositions characterized by more Proteobacteria, and less Actinobacteria, compared to mothers with low reported psychosocial stress. We additionally explored potential differences at genus-level, where we hypothesized to find lower levels of lactobacilli in mothers with high psychosocial stress.

**METHODS**

**Participants**

Participants were part of the BINGO (Dutch acronym for Biological Influences on Baby’s Health and Development) study, an ongoing longitudinal study investigating prenatal predictors of infant health and development. This study was approved by the ethical committee of the Faculty of Social Sciences of the Radboud University [ECSW2014-1003-189]. Participants signed up via the project’s website or folders that were handed out in midwife practices, pregnancy courses, and baby stores in the region Arnhem-Nijmegen (the Netherlands). Participants received a voucher with a value of 20€ and two small presents for the baby. Maternal exclusion criteria were: twin pregnancy, drug use, regular alcohol consumption, and insufficient knowledge of the Dutch language. A total of 87 expectant mothers enrolled for the study and signed the informed consent form. Of these, 73 were able to collect a stool sample. Three mothers took antibiotics at the time of collection and were therefore excluded. Subsequently, 70 healthy mothers participated in the part of the project reported here.

**Procedure**

After expectant mothers signed up for the project, they completed a demographics questionnaire and a questionnaire on general anxiety. The expectant mothers were then invited for a laboratory session, which took place during the third trimester of pregnancy (M prenatal = 33.9 weeks, SD prenatal = 2.3 weeks). During the laboratory visit, they completed additional self-report questionnaires, including the remaining questionnaires on prenatal psychosocial stress, and performed two computer tasks and an interaction task not relevant for the current study.

Prior to the lab visit, expectant mothers collected a stool sample using a sterile stool vial (80x16.5mm) with a spoon attached to the lid (Sarstedt inc.). The mothers were asked to fill one-third of the vial and to immediately store the vial in their home freezers (i.e., fresh frozen collection) until collected by the researcher. After collection, samples were stored at -80°C until analysis. Mothers were also asked to provide information on whether they were currently ill or had been ill the previous week, whether they had used antibiotics in the past three months, and whether they took food supplements during pregnancy.

**Measures**

**Maternal psychosocial stress.** Psychosocial stress can be defined as demanding conditions that exceed behavioral resources (Lazarus, 1966), and during pregnancy it includes maternal general and pregnancy-specific stress and anxiety (Beijers et al., 2014). Therefore, to measure prenatal psychosocial stress, expectant mothers were asked to fill in questionnaires related to general, as well as pregnancy-related stress and anxiety.

**General stress.** General stress was measured with the Alledaagse Problemen Lijst (Everyday Problem Checklist; EPL; Vingerhoets, Jeninga, & Mengers, 1989), a Dutch questionnaire that assesses the occurrence and intensity of daily hassles. This questionnaire contains 49 events, and participants have to check whether each event had occurred in the past two months, and if so, how much the event had bothered them on a 4-point Likert scale (1 = not at all, 4 = a lot). Subsequently, the mean intensity rating of daily hassles was calculated as the sum of how much the events bothered the participant divided by the frequency of the events. Hence, this variable could range from 0 to 4, with higher values indicating more experienced negativity as a result of daily hassles. Scale reliability (i.e., how closely related the set of items of the questionnaire are related as a group) in this sample was good, with Cronbach’s α, a measure of internal consistency, equal to 0.88.

**General anxiety.** To measure general anxiety, the state items from the State Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) were used. The STAI is the most widely researched and used questionnaire to measure general anxiety that has proven high internal consistency (Julian, 2011). Furthermore, it is relatively brief and easy to answer. The STAI questionnaire consists of 20 statements related to feelings of anxiety, yielding a score of how the participant feels at the present moment. Answers are given on a four-point Likert scale, ranging from 1 = not at all to 4 = a lot. Answers were summed up; hence scores could range from 0 to 80, with higher scores reflecting more general feelings of anxiety. Reliability of this scale in the current sample was good, with Cronbach’s α = 0.87.

**Pregnancy-related stress.** Pregnancy related stress was measured with the Pregnancy Experience Scale (PES; DiPietro, Ghera, Costigan, & Hawkins, 2004). This scale contains 43 pregnancy specific experiences. Participants are asked to rate the degree to which each experience constitutes both a hassle and an uplift during the whole pregnancy, both rated on a 4-point scale (0 = not at all, 3 = totally). Scores were derived by
calculating the ratio of hassles to uplifts, i.e., the sum of intensities of hassles divided by the sum of intensities of uplifts. Scores could thus range from 0 to 3, and higher scores indicate greater emotional valence towards pregnancy. Scale reliability in this sample was good, Cronbach’s α=0.89.

**Pregnancy-related anxiety.** Anxiety related to pregnancy was measured with two subscales of the Pregnancy specific Anxieties Questionnaire-Revised (PRAQ-R; Huizink, Mulder, Robles de Medina, Visser, & Buitelaar, 2004). These subscales measure ‘fear of giving birth’ (3 items), and ‘fear of bearing a handicapped child’ (4 items) during the whole pregnancy. Items could be answered on a scale from 1=not at all true to 5=totally true. For ‘fear of giving birth’ scores could range between 0 and 15; for ‘fear of bearing a handicapped child’ scores could range from 0 to 20. Higher scores indicate higher levels of pregnancy-related anxiety. Cronbach’s α scale reliability was 0.52 for ‘fear of giving birth’, and 0.85 for ‘fear of bearing a handicapped child’.

**Maternal microbiota.** Approximately 0.1-0.15g of fecal sample from each participant was used for DNA extraction. Total microbial DNA was extracted using the Maxwell® 16 Total RNA system (Promega) with Stool Transport and Recovery Buffer (STAR; Roche Diagnostics Corporation, Indianapolis, IN). Briefly, the fecal sample was homogenized with 0.25g of sterilized 0.1mm zirconia beads and three glass beads (2.5mm) in 350µL STAR buffer for 3×1min at 5.5ms using a Precellys 24 beadbeater (Bertin technologies, France). Samples were then incubated with shaking at 100rpm for 15min at 95°C and pelleted by 5min centrifugation at 4°C and 14000g. The supernatant was removed and the pellets were processed again as described above using 200µL of fresh STAR buffer. The supernatant was removed, pooled with the first supernatant, and 250µL were used for purification with Maxwell® 16 Tissue LEV Total RNA Purification Kit customized for fecal DNA extraction (AS1220) following the manufacturer’s instructions. DNA was eluted with 50µL of DNase and RNase free water (Qiagen, Hilden, Germany). DNA concentrations were measured spectrophotometrically with a NanoDrop® Technologies, Wilmington, DE, USA) and adjusted to 20ng/µL with DNAse and RNase free water. The V4 region of 16S ribosomal RNA (rRNA) gene was amplified with a unique barcoded primer 515F-n and 806R-n (10µM each/reaction; 40), 1xHF buffer (Finnzymes, Vantaa, Finland), 1µL dNTP Mix (10mM each, Roche Diagnostics GmbH, Mannheim, Germany), 1 U Phusion® Hot Start II High Fidelity DNA Polymerase (Finnzymes, Vantaa, Finland) and 36.5µL of DNase and RNase free water. The amplification program included a 30s initial denaturation step at 98°C, followed by 25 cycles of denaturation at 98°C for 10s, annealing at 56°C for 10s and elongation at 72°C for 10s, and a final extension at 72°C for 7min.

The PCR product presence and size (~290 bp) was confirmed with gel electrophoresis using the Lonza FlashGel® System (Lonza, Cologne, Germany). Seventy unique barcode tags were used in each library, and artificial control (Mock) communities were included. PCR products were purified with the HighPrep® PCR kit (MagBio Genomics, Alphen aan den Rijn, Netherlands), and DNA concentrations were measured with the Qubit® dsDNA BR Assay Kit (Life Technologies, Leusden, Netherlands). From each barcoded sample, 100ng was added to the amplicon pool that was then concentrated with the HighPrep® PCR kit to 20µL. The concentration was measured with the Qubit® dsDNA BR Assay Kit and adjusted to 100ng/µL final concentration. The pooled libraries were sent for adapter ligation and Illumina HiSeq sequencing to GATC-Biotech, Konstanz, Germany.

Data processing and analysis were carried out using the NG-Tax pipeline (Ramiro-Garcia et al., 2016). Alpha diversity analyses were carried out in QIIME with rarefaction cutoff of 3000 reads (Kuczynski et al., 2011). Principal components analysis (PCA) was performed in CANOCO 5 (Lepš & Šmilauer, 2014).

**Statistical analyses**

We performed Principal Coordinate Analysis (PCoA) based on weighted or unweighted unifrac distances in Qiime, and Principal Component Analysis (PCA) and Redundancy analysis (RDA) based on relative abundance distributions using Canoco 5 software to check for sample clustering at phylum level and then genus level in relation to the maternal stress variables.

Then, Partial least squares (PLS) regression (Wold, Sjöström, & Eriksson, 2001) was used to assess to what extent the maternal stress variables (‘response’) could be predicted based on microbial community composition data. Briefly, PLS regression is a generalization of multiple regression, which searches for a set of components that performs a simultaneous decomposition of the predictor matrix (X=microbiota) and of the response matrix (Y=stress index), with the constraint that these components explain as much as possible of the covariance between X and Y. The optimal number of PLS components was defined using double-cross validation (Szymanśka, Saccenti, Smilde, & Westerhuis, 2012). One component was found to be appropriate to model the data, and the R² parameter (variance explained, where R²=1 indicates perfect prediction ability of the model) was used to assess the quality of the final regression model. Model significance was assessed using a permutation test with 1000 permutations, with significance cut-offs of 0.01 and 0.05. Selection of the most important bacteria (i.e., bacteria that were detected in most samples, and drove the results) in the final model was performed using the Significance Multivariate Correlation criterion (Tran, Afanador, Buydens, & Blanchet, 2014). Data (X) were centered and scaled to unit variance; the
response Y was log transformed before analysis. PLS was performed using the Matlab MEDA-toolbox (Camacho, Pérez-Villegas, Rodríguez-Gómez, & Jiménez-Mañas, 2015).

Shannon and Chao1 scores from alpha diversity analyses were compared between high and low groups for the maternal stress variables that significantly predicted bacteria clustering using QIIME (Kuczynski et al., 2011).

**RESULTS**

Descriptive Statistics

Descriptive statistics for the study variables can be found in Table 1. Figure 1 shows the Pearson correlations between the psychosocial stress variables. General stress was positively correlated with general anxiety \((r=0.34, p=.005)\) and pregnancy-related stress \((r=0.34, p=.005)\). Furthermore, general anxiety was positively related with pregnancy-related stress \((r=0.40, p=.001)\) and fear of giving birth \((r=0.26, p=.027)\), and pregnancy-related stress was positively related with fear of giving birth \((r=0.35, p=.003)\). The strength of these correlations was weak, indicating that despite these associations, the variables generally tap into different aspects of maternal psychosocial stress.

Fecal samples from 70 mothers were analyzed for microbial composition using Illumina HiSeq sequencing of barcoded 16S rRNA gene amplicons. The total number of resulting sequencing reads was 10,201,505 and ranged from 6,447 to 632,101 reads per sample with an average number of reads per sample of 139,747 \((SD=130,752, SE=15,303)\). A total of 113 genus level taxa were identified, of which 76 were present in more than 95% of all samples. The average relative abundance of these taxa is summarized in Table 2.

### Table 1

<table>
<thead>
<tr>
<th>M (SD)</th>
<th>Range</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>31.61(3.66)</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>25.36-40.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Educational background</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>21 (29%)</td>
<td>70</td>
</tr>
<tr>
<td>University</td>
<td>39 (54%)</td>
<td></td>
</tr>
<tr>
<td>Gestational age at collection</td>
<td>33.60(2.38)</td>
<td>62*</td>
</tr>
<tr>
<td></td>
<td>28.00-38.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prenatal psychosocial stress</td>
<td></td>
</tr>
<tr>
<td>General stress (EPL)</td>
<td>2.22(0.42)</td>
<td>68*</td>
</tr>
<tr>
<td></td>
<td>1.36-3.14</td>
<td></td>
</tr>
<tr>
<td>General anxiety (STAI)</td>
<td>29.97(6.06)</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>20-45</td>
<td></td>
</tr>
<tr>
<td>Pregnancy-related stress (PES)</td>
<td>0.29(0.21)</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>0-0.87</td>
<td></td>
</tr>
<tr>
<td>Fear of giving birth (PRAQ-R)</td>
<td>5.76(2.11)</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>3-11</td>
<td></td>
</tr>
<tr>
<td>Fear of bearing a handicapped child (PRAQ-R)</td>
<td>8.91(2.94)</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>4-20</td>
<td></td>
</tr>
</tbody>
</table>

*Some participants did not fill in all questionnaires, hence N<70.

**Figure 1.** Correlations between the psychosocial stress variables (\(^*p<.05, \)^\(^{**}p<.001\)). Note that Fear of Bearing a Handicapped Child and Fear of Giving Birth are considered pregnancy related anxieties.
<table>
<thead>
<tr>
<th>Genus Taxon</th>
<th>Average RA (%</th>
<th>SE</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blautia</td>
<td>13.38</td>
<td>1.6E-02</td>
<td>100.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>8.65</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.063</td>
<td>7.6E-05</td>
<td>8.65</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>7.12</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.055</td>
<td>6.7E-05</td>
<td>11.4</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>6.61</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.053</td>
<td>6.5E-05</td>
<td>18.6</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>6.06</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.051</td>
<td>6.3E-05</td>
<td>14.3</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>5.43</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.049</td>
<td>6.2E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>4.93</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.047</td>
<td>6.1E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>4.32</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.045</td>
<td>6.0E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>3.82</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.043</td>
<td>5.9E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>3.31</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.041</td>
<td>5.8E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>2.80</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.039</td>
<td>5.7E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>2.30</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.037</td>
<td>5.6E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>1.80</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.035</td>
<td>5.5E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>1.30</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.033</td>
<td>5.4E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>0.80</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.031</td>
<td>5.3E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>0.30</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.029</td>
<td>5.2E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>0.00</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.00</td>
<td>5.1E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>0.00</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.00</td>
<td>5.0E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>0.00</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.00</td>
<td>4.9E-05</td>
<td>10.0</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>0.00</td>
<td>8.5E-03</td>
<td>100.0</td>
</tr>
<tr>
<td>Methanosphaera</td>
<td>0.00</td>
<td>4.8E-05</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Note: When a taxonomic assignment could not be made at genus level, the lowest classifiable taxonomy assignment (family-f, class-c, order-o, kingdom-k) was used, and the unidentified genus with that level was denoted as “g_g”.
Associations between Maternal Prenatal Psychosocial Stress and Microbiota

Partial least squares (PLS) modeling was performed to predict the psychosocial stress variables from microbiota relative abundance profiles at phylum level. No significant associations were found between any of the maternal psychosocial stress variables and the microbiota at phylum level.

PLS modeling was repeated at genus level. These statistical analyses showed that general anxiety (STAI) was significantly associated with microbial relative abundance profiles in mothers ($R^2=0.65$, $p=0.008$). There were no associations between microbial relative abundance profiles at genus level and the other psychosocial stress variables.

The Significance Multivariate Correlation criterion was used to select the taxa contributing to the predictive model of general anxiety. *Parasutterella* was found to significantly contribute to the model at the 0.01 significance level. Ten genus level groups were significant at the 0.05 level, and the average relative abundance of these microbial taxa was calculated for the mothers with high (above the median of 30) and low (below 30) general anxiety (see Table 3). The fecal microbiota of mothers with lower prenatal anxiety was characterized by higher relative abundance of *Oscillospira*, *Eubacterium*, and *Megamonas*. The fecal microbiota of mothers with higher prenatal anxiety was characterized by higher relative abundance of *Oxalobacter*, *Rothia*, *Acetitomaculum*, *Acidaminococcus*, and *Staphylococcus*.

Two tailed, unpaired t-tests showed no differences in microbial richness (i.e., how many genera there are in each mother) and diversity (i.e., number and relative abundance distribution of genera) as respectively estimated with the Chao1 richness and Shannon diversity scores between mothers in the high and low general anxiety groups ($p=0.92$ and $p=0.88$, respectively). PCoA based on weighted and unweighted unifrac distances as well as PCA and RDA analyses based on relative abundance data were used to see whether the overall microbiota profiles of mothers with low and high general anxiety were similar or not. None of these analyses showed separation of the data in relation to low/high general anxiety. Thus, the microbiota of mothers with low and high general anxiety showed no specific patterns overall.

### Table 3 — Microbial Taxa Significantly Associated at p<.05-level with High and Low General Anxiety Groups

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Class</th>
<th>Order</th>
<th>Family</th>
<th>Genus</th>
<th>Mean Relative Abundance</th>
<th>SE</th>
<th>Prevalence</th>
<th>Mean Relative Abundance</th>
<th>SE</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firmicutes</td>
<td>Clostridia</td>
<td>Lachnospirales</td>
<td>Lachnospiraceae</td>
<td>Acetitomaculum</td>
<td>0.00004</td>
<td>0.0002</td>
<td>1</td>
<td>0.00068</td>
<td>0.0004</td>
<td>2</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Clostridia</td>
<td>Veillonellales</td>
<td>Veillonellaeeae</td>
<td>Acidaminococcus</td>
<td>0.00003</td>
<td>0.0002</td>
<td>1</td>
<td>0.00049</td>
<td>0.0022</td>
<td>6</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Clostridia</td>
<td>Eubacteriales</td>
<td>Eubacterialesae</td>
<td>Eubacterium</td>
<td>0.00191</td>
<td>0.0172</td>
<td>32</td>
<td>0.02560</td>
<td>0.0361</td>
<td>34</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Clostridia</td>
<td>Peptococcaceae</td>
<td>Peptococcaceae</td>
<td>Incertae_Sedis</td>
<td>0.00014</td>
<td>0.0007</td>
<td>6</td>
<td>0.00071</td>
<td>0.0008</td>
<td>7</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Clostridia</td>
<td>Veillonellales</td>
<td>Veillonellaeeae</td>
<td>Megamonas</td>
<td>0.00046</td>
<td>0.0022</td>
<td>2</td>
<td>n.d.</td>
<td>n.d.</td>
<td>0</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Clostridiales</td>
<td>Oscillospirales</td>
<td>Oscillospira</td>
<td>n.d.</td>
<td>n.d.</td>
<td>0</td>
<td>0.00007</td>
<td>0.0003</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Clostridiales</td>
<td>Ruminococcaceae</td>
<td>Ruminococcaceae</td>
<td>n.d.</td>
<td>n.d.</td>
<td>0</td>
<td>0.00004</td>
<td>0.0003</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Bacillales</td>
<td>Staphylococcaceae</td>
<td>Staphylococcus</td>
<td>n.d.</td>
<td>n.d.</td>
<td>0</td>
<td>0.00004</td>
<td>0.0004</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Bacillales</td>
<td>Staphylococcaceae</td>
<td>Staphylococcus</td>
<td>n.d.</td>
<td>n.d.</td>
<td>0</td>
<td>0.00004</td>
<td>0.0001</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Bacillales</td>
<td>Staphylococcaceae</td>
<td>Staphylococcus</td>
<td>n.d.</td>
<td>n.d.</td>
<td>0</td>
<td>0.00004</td>
<td>0.0001</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Note: RA=Relative Abundance, SE=Standard Error, n.d.=not detected, Prevalence=number of samples where genus was found. Significant at p<0.05 level.
DISCUSSION

The current study investigated an essential step of the proposed mechanism behind the links between maternal pregnancy psychosocial stress and child outcomes: the relation between psychosocial stress and fecal microbiota in pregnant mothers. Based on previous findings (Zijlmans et al., 2015), mothers with high self-reported psychosocial stress during pregnancy were hypothesized to have fecal microbial profiles characterized by a higher relative abundance of Proteobacteria and lower relative abundance of Actinobacteria and lactobacilli, as compared to mothers with low reported psychosocial stress. This hypothesis was not confirmed as in our study sample there were no differences in microbial profiles at phylum level based on psychosocial stress.

Our additional exploratory analyses, however, revealed that several microbial taxa at genus level significantly contributed to a PLS model for the prediction of general anxiety during pregnancy. These taxa, however, did not include the genus Lactobacillus, which had previously been found to be reduced in relative abundance in mothers with high prenatal psychosocial stress (Zijlmans et al., 2015). Finally, we did not find significant associations between general stress and pregnancy-specific stress and anxiety and maternal fecal microbiota composition at genus level.

We assumed that prenatal psychosocial stress might affect the child’s development via the mother’s, and in turn the infant’s, intestinal microbiota (Beijers et al., 2014), as the largest bacterial colonization of the infants’ intestines occurs after transfer of maternal-origin bacteria during vaginal delivery (de Weerth et al., 2013; Grönlund et al., 2011). Furthermore, infants’ intestinal bacteria have been shown to differ based on the level of maternal psychosocial stress (Zijlmans et al., 2015). Our results showed no relation between maternal psychosocial stress and the microbial groups previously found to differ in infants, i.e. Proteobacteria, Actinobacteria and lactobacilli, which seems to disprove our assumption. However, given that we did not investigate the infants’ intestinal microbiota, we cannot discard the hypothesis that maternal psychosocial stress affects the child via intestinal microbiota. Furthermore, it should be noted that we did find associations between general anxiety and maternal microbiota at genus level.

Results showed that the fecal microbiota of mothers with lower prenatal anxiety was characterized by higher relative abundance of the genus Oscillospira and other genera from the phylum of Firmicutes (Eubacterium and Megamonas). In contrast, the fecal microbiota of mothers with higher prenatal anxiety was characterized by higher relative abundance of Oxalobacter (belonging to the phylum of Proteobacteria), Rothia (belonging to the phylum of Actinobacteria), and genera from the phylum of Firmicutes, including Acetitomaculum, Acidaminococcus, Staphylococcus, and unidentified genus-level taxa within the families Peptococcaceae and Peptostreptococcaceae. Of these taxa, particularly the two unidentified genera within the families Peptostreptococcaceae and Peptococcaceae were detected in the samples of almost every mother, whereas others were only found in a few samples, explaining their low average relative abundance.

Peptostreptococcaceae are a family of bacteria from the class Clostridia. They appear to be over-represented in colorectal cancer patients (Ahn et al., 2013). Peptostreptococcaceae were also related to poor cognition and neuro-inflammation in cirrhosis patients with brain dysfunction (Ahlulwalia et al., 2016). Animal research shows that Peptostreptococcaceae might also be related to stress. However, these studies suggested that high levels of stress are associated with a decrease in the relative abundance of Peptostreptococcaceae. For example, after treatment with antibiotics, the relative abundance of Peptostreptococcaceae significantly decreased in chronically stressed rats, compared to control animals (Xu et al., 2014). Furthermore, members of the Peptostreptococcaceae seem to be strongly affected by diet. For example, in adult pigs, the relative abundance of Peptostreptococcaceae increased when dietary protein intake decreased from 16% to 13% (Fan, Liu, Song, Chen, & Ma, 2017). Also, mice fed with either a low-fat diet or high-fat diet showed reduced Peptostreptococcaceae when calories were restricted (Zhang et al., 2013). Peptostreptococcaceae were also found to be negatively related to life-span, regardless of fat intake (Zhang et al., 2013). Since we did not measure dietary habits, we do not know whether the difference in relative abundance in Peptostreptococcaceae between the low and high anxiety group is based on differences in food intake, or indeed due to differences in anxiety levels.

Peptococcaceae are also from the class Clostridia. Peptococcaceae have complex nutritional requirements, they may or may not ferment carbohydrates, and they are found in normal and pathologic female urogenital tracts (Farlex Partner Medical Dictionary, 2012). Hence, vaginally delivered infants may be exposed to them during birth. Peptococcaceae may or may not be pathogenic (Farlex Medical Dictionary, 2009). Changes in urogenital bacteria in mothers might therefore expose a neonate’s gut to potentially pathogenic bacteria. However, human studies on Peptococcaceae are rare. Animal studies revealed a possible relationship of Peptococcaceae with stress and diet. In rats reared under severe crowding stress, Peptococcaceae were increased compared to control rats (Suzuki, Kodama, & Mitsuoka, 1989). Furthermore, high fat diet and induced stress were related to increases in Peptococcaceae in female rats (Bridgewater et al., 2017).

As said previously, we also detected differences between mothers with low and high general anxiety in genus level bacteria besides Peptococcaceae and Peptostreptococcaceae. Megamonas, Eubacterium, and Oscillospira were detected in either one or two samples from mothers with low general anxiety, and in none of the samples from mothers with high general anxiety. Rothia and Oxalobacter were each detected in one sample from mothers with high general anxiety, and in none of the samples from mothers with low
general anxiety. *Staphylococcus* and *Acetitomaculum* were detected in, respectively, one and two samples from mothers with high general anxiety, and each in one of the samples from mothers with low general anxiety. As these bacteria were detected in only a small number of samples, we will refrain from (over-) interpreting these results. Confirmation of the findings in larger study populations is needed before an in-depth discussion of the results is warranted. Additionally, it would be important to investigate whether these bacteria are also related to anxiety in a non-pregnant population.

Our results did not provide evidence of an association between general stress or pregnancy-specific stress and anxiety and microbial abundances. It is difficult to explain why in our sample maternal gut bacteria were specifically associated to self-reports of general anxiety and not to self-reports of general stress or pregnancy-specific stress and anxiety. A possible explanation may lie in the nature of the questionnaires used. The anxiety questionnaire used in our study requires reporting on current feelings, while the other questionnaires ask about feelings over a given period of time. A questionnaire on momentary emotions may be more reliable and closely linked to reality than questionnaires that require the participant to ‘summarise’ emotions over a longer period of time and that are automatically subject to problems of recall (Solhan, Trull, Jahng, & Wood, 2009).

As we only found an association between general anxiety and maternal intestinal microbiota, whereas we did not find an association between psychosocial stress and intestinal microbiota previously found to be different in infants from mothers with psychosocial stress, it is essential to discuss other potential links between prenatal stress and infant microbiota. One of these links might be related to cortisol. Cortisol concentrations in plasma increase when humans are confronted with stress and the HPA-axis is activated (Beijers et al., 2014). Maternal cortisol is known to cross the placenta and to increase cortisol concentrations in the fetus (Duthie & Reynolds, 2013). In turn, these heightened cortisol levels can affect the developing HPA axis of the fetus, resulting in increased basal cortisol concentrations and cortisol reactivity in the infant after birth (Tollenaar, Beijers, Jansen, Riksen-Walraven, & de Weerth, 2011). Cortisol, in turn, can change the permeability of the gut and affect the immune cells in the gut, affect gut motility and secretion, and produce increases in bile acid, all of which can potentially influence the infant intestinal microbiota (De Palma et al., 2015; Zijlmans et al., 2015; Cryan & Dinan, 2012).

Another physiological route for maternal prenatal psychosocial stress to affect offspring microbiota might take place in the postpartum period. If maternal prenatal stress continues after birth as high postnatal psychosocial stress, it might affect breast milk composition, including breast milk cortisol concentrations. Maternal plasma cortisol is transferred to maternal breast milk (Hamosh, 2001). This cortisol from milk arrives in the infants’ intestines (Hinde et al., 2014), where it binds to cortisol receptors, influencing the maturation of the gastrointestinal tract (Hinde et al., 2014), which may in turn affect which bacterial species establish themselves in the gut.

As far as we know, this is the first study to look at psychosocial stress and microbiota composition during the third trimester of pregnancy. A positive feature of the study is that several distinct aspects of psychosocial stress (i.e. general stress and anxiety, and pregnancy-specific stress and anxiety) were distinguished. Nonetheless, the study also has limitations. First, the sample consisted of mothers from a highly educated background. The current findings might not be generalizable to the whole population. Second, information on food intake was not included. As for example *Peptostreptococcaceae* and *Peptococcaceae*- which were more abundant in the high anxiety group- have previously been associated with diet, including maternal diet information in future studies will help obtain a more comprehensive understanding of observed differences in microbial community composition. In a large enough sample, complex models of associations between maternal psychosocial stress and diet during pregnancy, and intestinal microbiota can be investigated.

This study investigated the first step of a mechanism potentially underlying links between maternal prenatal psychosocial stress and infant outcomes, namely the relation between psychosocial stress and fecal microbiota in pregnant mothers. Contrary to our hypotheses, we did not find that mothers with high psychosocial stress had phylum-level microbial compositions characterized by more *Proteobacteria*, and less *Actinobacteria*, and lower levels of the genus *Lactobacillus*, compared to mothers with low reported psychosocial stress. However, we did find a significant association between late pregnancy general anxiety and the women’s fecal microbial composition at genus level. More specifically, the fecal microbiota of mothers with lower anxiety was characterized by higher relative abundances of the genera *Eubacterium* and *Oscillospira* compared to mothers with higher prenatal anxiety. These bacteria have been previously termed beneficial microbes. Additionally, mothers with higher prenatal anxiety had higher relative abundances of unidentified genera within the families *Peptostreptococcaceae* and *Peptococcaceae*. Previous studies have associated these bacterial groups to stress and poor health in rats and mice. Finally, we also found differences between mothers with low and high anxiety in bacteria that have not been associated with anxiety in earlier studies. The current study therefore offers insights into associations between maternal mental health and gut microbial composition during pregnancy and provides a starting point for future investigations in which maternal diet as well as infant microbiota and development should also be assessed.
REFERENCES


Kant, E., Saulyte, A., Neve, K., & Hansen, S. (2014). Moderate dietary protein restriction alters the composition of gut microbiota and improves ileal barrier function in adult pig model. Scientific Reports, 7, 43412. doi:10.1038/srep43412


Chapter 6

Are Cortisol Concentrations in Human Breast Milk Associated with Infant Crying?

ABSTRACT

The present longitudinal study is the first to investigate the association between human breast milk cortisol and infant crying over the first three months of life. Higher concentrations of breast milk cortisol were expected to be differentially associated with fussing and crying in boys and girls. At 2, 6, and 12 weeks of infant age, mothers (N=70) collected a morning sample of their milk and kept a 3-day diary to measure infant fussing and crying. Cortisol was extracted and quantified from milk samples. Results showed that breast milk cortisol concentrations increased from 2 weeks through 12 weeks of infant age. Milk cortisol was unrelated to the total duration, frequency, and bout length of infant fussing and crying for both boys and girls. Directions for future research aiming to extend our knowledge on the biology of milk cortisol in relation to infant behavior and development are discussed.

Key words: Humans; Breast milk cortisol; Infant crying; Breastfeeding; Lactational programming

INTRODUCTION

Research shows that breastfeeding carries many benefits for infant health and development. For example, a longer duration of breastfeeding has been related to fewer infant illnesses (Pettigrew et al., 2003; Thomas, 2014) and better cognitive development, even beyond infancy and childhood (Agostoni et al., 2009; Evenhouse & Reilley, 2005; Huang, Peters, Vaughn, & Witko, 2014). Less research has been devoted to the possible effects of biological constituents of breast milk on offspring phenotype. Breast milk contains, among others, water, protein, carbohydrates, immune factors, and several hormones (Miller et al., 2013). The present study focuses on the hormone cortisol, the primary glucocorticoid in humans (Glynn et al., 2007). Cortisol serves numerous functions, orchestrating aspects of the physiological stress response, immunity, and metabolism. For example, cortisol concentrations increase when humans are confronted with an acute psychological or biological stressor and the hypothalamic-pituitary-adrenal (HPA)-axis is activated (Dickerson & Kemeny, 2004). Furthermore, cortisol concentrations are also involved in the routine metabolism of proteins, carbohydrates, and fats (Hinde et al., 2014). Earlier research has shown that in animals, concentrations of glucocorticoids in maternal milk are related to offspring behavior (Hinde, 2013; Sullivan, Hinde, Mendoza, & Capitanio, 2011). The current study investigated whether breast milk cortisol is related to behavior, more specifically fussing and crying, in human infants.

During pregnancy, a mother’s physiology can impact the developing fetus (Beijers, Buitelaar, & de Weerth, 2014). Maternal physiology is thought to relay information about the environment to the fetus, and thereby prepare the infant for the time after birth. The fetus can thus start to develop a behavioral profile matching the environment it will likely live in, a process also referred to as fetal programming (Del Giudice, 2012; Frankenhuis & de Weerth, 2013). After birth and during lactation, mothers can still transfer physiological signals to the infant through the biological constituents of breast milk (Hinde et al., 2014), including cortisol. Cortisol concentrations are transferred from plasma to breast milk, as there is no mammary synthesis of cortisol (Hamosh, 2001). In line with the fetal programming hypothesis, it might be argued that, when the environment is stressful, mothers could physiologically signal this information to their infants via breast milk cortisol, and infants could subsequently adjust their behavioral phenotype to their (future) environment. That is, mothers with higher milk cortisol concentrations might be experiencing more stressors in their life, and milk cortisol would partly underlie the maternal organization of biobehavioral processes of the offspring (Dettmer et al., 2017). This hypothesis is also referred to as the lactational programming hypothesis (Hinde, 2013). The lactational programming hypothesis is the idea that biological constituents of breast milk influence metabolic and neurobiological development, which in turn
influence offspring phenotype and behavior (Hinde, 2013). This hypothesis is supported by studies showing that early infancy is a sensitive period in which the brain shows large plasticity and openness to environmental influences (Kolb & Gibb, 2011). Additionally, especially early in life, the infant’s intestinal tract has a high number of glucocorticoid receptors (Allen-Blevis, Sela, & Hinde, 2015). Hence, milk cortisol that reaches the infants’ intestines can easily cross the intestinal epithelial barrier (Hinde et al., 2014), as well as later on pass the blood-brain barrier (Glynn et al., 2007). In the brain, cortisol seems to specifically bind to receptors in the limbic region (Owen, Andrews, & Matthews, 2005), a brain area involved in, among others, the regulation of emotions and behavior (Grey, Davis, Sandman, & Glynn, 2013).

In stressful environments, infants are suggested to signal needs to the caregivers, as caregivers are less likely to be involved with and available for their children when confronted with stress (Baker, 2014). To ensure being fed and taken care of, young infants are expected to trigger caregivers’ attention by fussing and crying. Indeed, in some contexts, fussing and crying may be particularly appropriate and even life-saving for the infant (Barr, 2000). This hypothesis is in line with findings of de Vries (1984), who showed that in an environment with extreme drought and hunger, fussy infants survived more often than easy infants. Hence, and based on the lactational hypothesis (Hinde, 2013), higher breast milk cortisol concentrations are expected to be related to more fussing and crying behavior in infants. It is important to note, however, that even though more infant fussing and crying might be considered adaptive in stressful environments, excessive fussing and crying is energetically costly, and may be deviating energy from other developmental purposes, including physical growth and health (De Lauzon-Guillain et al., 2012). Also, excessive infant fussing and crying may in some cases be indicating a regulatory problem that will persist into childhood (Bilgin & Wolke, 2017). Moreover, excessive infant fussing and crying causes considerable concern, distress, and exhaustion in parents (de Weerth, Fuentes, & de Vos, 2013), and in extreme cases can lead to child abuse and neglect (Reijneveld, van der Wal, Brugman, Hira Sing, & Verloove-Vanhorick, 2004). Thus, even though more fussing and crying might be adaptive in stressful environments, excessive fussing and crying may be associated with negative outcomes for both the infant and the parents.

To date, no research has been published on the link between breast milk cortisol and infant fussing and crying. However, previous animal research indicates that a link exists between breast milk glucocorticoids and offspring behavior. Higher glucocorticoid concentrations in breast milk were related to reduced anxiety and improved learning in rat offspring (Catalani et al., 2002, 2000). In monkey offspring, higher cortisol concentrations were related to more frequent play and social behavior (Dettmer et al., 2017). Also, in rhesus monkeys, breast milk cortisol concentrations were positively related to more confident temperament, but only in male offspring (Sullivan et al., 2011). However, in a following, more extensive study in the same primate colony, breast milk cortisol concentrations were positively related to more nervous and less confident temperament in both male and female offspring (Hinde et al., 2014). In this study, for female offspring, this more nervous and less confident temperament was related to absolute concentrations of breast milk cortisol, whereas for male offspring, this type of temperament was related to an increase in cortisol concentrations from 1 to 3-4 months.

Only a few human studies have investigated the link between breast milk cortisol and infant behavior. In the first study on breastfeeding mother-infant pairs, higher cortisol concentrations in maternal plasma - an indicator of cortisol concentrations in breast milk - were related to more mother reported fearful temperament in 2-month-old infants (Glynn et al., 2007). A later study from the same research group on 3-month-old infants showed that higher cortisol concentrations in breast milk predicted more reported temperamental negative affectivity (including sadness and fear), but only in female infants (Grey et al., 2013). Finally, a recent study found that higher breast milk cortisol concentrations at 2.5 months postpartum were associated with higher induced fear reactivity in 8-month-old infant girls, but not boys (Nolvi et al., 2017).

In sum, only a few studies examined links between human breast milk cortisol and infant behavior and found that higher cortisol concentrations were related to more infant temperamental negativity (Glynn et al., 2007; Grey et al., 2013; Nolvi et al., 2017). The present study aims to longitudinally investigate associations between breast milk cortisol and infant fussing and crying. Note that the link between breast milk cortisol and infant fussing and crying behavior may be bi-directional in nature, as it is also possible that infant fussing and crying predicts more maternal stress, and subsequently, higher cortisol in breast milk. Infant crying has been found to impair maternal psychological well-being and decrease sleep quality (Barr et al., 2014; Brand, Furlano, Sidler, Schulz, & Holsboer-Trachsler, 2014), which in turn may be related to more cortisol in breast milk. In the current study, infant fussing and crying was measured at 2, 6, and 12 weeks of infant age, because these ages represent, respectively, the start, peak, and decrease of the normal developmental crying curve (Barr, Trent, & Cross, 2006). Advantages of the longitudinal design are the ability to examine the stability of breast milk cortisol over time and to investigate the associations of milk cortisol with infant fussing and crying over time. We expected higher concentrations of breast milk cortisol to be related to higher levels of infant fussing and crying. Furthermore, because of the indications that breast milk cortisol may predict offspring behavior differentially for male and female offspring (Dettmer et al., 2017; Grey et al., 2013; Nolvi et al., 2017), we explored sex differences in the link between breast milk cortisol and infant fussing and crying.
METHOD

Participants
Participants were part of the BINGO study (Dutch acronym for Biological Influences on Baby’s Health and Development), a longitudinal study aimed to identify prenatal and early postnatal predictors of infant health and development. This study was approved by the ethical committee of the Faculty of Social Sciences of the Radboud University [ECSW2014-1003-189]. Participants signed up via the project’s website, or folders that were handed out in midwife practices, pregnancy courses, and baby stores in the region Nijmegen-Arnhem (the Netherlands). Participants received a voucher for 20€ and two presents for the infant.

Initial exclusion criteria were: drug use, excessive alcohol use (i.e., alcohol dependency), insufficient knowledge of the Dutch language, and an unhealthy, complicated pregnancy. Nulliparity was not requested. Eighty-eight eligible expectant mothers signed the informed consent form and subsequently participated in this project. Postnatal exclusion criteria were: complications during pregnancy (after initial contact), gestational age at birth <37 weeks, birth weight <2500 grams, 5-minute Apgar score <7, and congenital malformations. Five infants were born between 35 and 37 weeks of pregnancy. As these infants were otherwise healthy and fit the inclusion criteria, they were not excluded. Statistical analyses with and without these five infants rendered comparable results. One infant was born prematurely (week 32), and one infant was born with brain damage; both were excluded from further participation. Five families stopped participation after birth due to personal reasons. There were no differences in educational level, maternal age, gestational age at birth or infant birth weight between participating and non-participating mothers. Of the 81 remaining mothers, 77 mothers started breastfeeding after birth and were therefore included in the current study. See Figure 1 for a flow chart of the sample.

Procedure
Mothers were asked to wash their hands, breasts, and nipples before collection, and, in the case of the breast pump, to first boil the parts that came in contact with the milk. Mothers noted the date, time of waking, time of collection, whether they were or had been ill and/or taken medication the preceding week, and if so, which medication. The mothers collected the milk before feeding the infant. Samples were stored in the mothers’ freezers and collected with a portable freezer after the last sample was taken (approximately when the infant was 13 weeks of age). At the Radboud University, all samples were stored in a freezer at -80 degrees Celsius, and subsequently sent by temperature-controlled shipment to the Utrecht University Medical Centre. About 4ml was needed for cortisol analysis.

Infant fussing and crying. Infant fussing and crying was assessed using the Baby Day Diary (Barr, Kramer, Boisjoly, McVey-White, & Pless, 1988) for three consecutive days at each time point. Each day, mothers reported the occurrence of the following infant behaviors: fussing, crying, unsoothable crying, sleeping, feeding, and being awake without crying. This was indicated with lines/symbols assigned to each behavior on time bars. Each 24-hr period was represented on a sheet of paper with four horizontal 6-hr time bars, subdivided into periods of five minutes (Barr et al., 1988). Mothers were asked to retrospectively fill in this diary every two to three hours. The measurement of fussing and crying in the Baby Day Diary is known to be valid, and to produce data comparable to actual audio recordings (n=90; Barr et al., 1988). The Baby Day Diary has been extensively used in other research (e.g., Kusaka, Ohgi, Shigemori, & Fujimoto, 2008; de Weerth, Fuentes, Puylaert, & de Vos, 2013; Radesky et al., 2013; Korja et al., 2014). The breast milk sample was collected on one of the three days the Baby Day Diary was filled in.

Measurements
Breast milk cortisol. Breast milk samples were analyzed in the University Medical Centre in Utrecht, the Netherlands. Samples were centrifuged with a force of 21,000g, and the fatty layer was discarded. Then, d4-cortisol was added as an internal standard, and cortisol was extracted with methyl tertiary butyl ether. Cortisol was quantified by Liquid chromatography-tandem mass spectrometry (LC-MS/MS). Calibrator solutions were prepared from Sigma-Aldrich cortisol preparations. The UHPLC-MS/MS system consisted of a Dionex Ultimate 3000 UHPLC system coupled with a TSQ Quantiva mass spectrometer (ThermoElectron Corp, West Palm Beach, FL). A Hypersil Gold 50x2.1mm 1.9µm column was used with a methanol/water gradient containing ammoniumformiate and formic acid for the UHPLC separation. Day-to-day imprecision for cortisol was 1.9% at 6 nmol/L (n=6), LLQ is 1.0 nmol/L for cortisol.

Infant fussing and crying. The diary data were prepared for analysis by noting the number of times fussing, crying, and unsoothable crying occurred (i.e. frequency), and
how long the behaviors lasted (i.e. bout length). No distinction was made between fussing, crying and unsoothable crying in the analyses, and the behavior is henceforth referred to as ‘crying’. This was done for each of the three days separately and then the mean daily frequency and bout length were calculated. The multiplication of the mean daily frequency and bout length rendered the mean total duration of crying. This lead to three outcome measures for crying: total duration (mean duration in minutes per 24 hr), frequency (mean number of episodes per 24 hr), and bout (mean bout length of each episode) (Fujiwara, Barr, Brant, & Barr, 2011).

Although fussing, crying, and unsoothable crying represent gradations in the intensity of crying that may be differentially displayed in different types of situations, in this study they were combined. This is common practice in research on young infants (Wolke, Bilgin, & Samara, 2017) who are not yet using fussing in an instrumental manner. Furthermore, the original research on the Baby Day Diary (Barr et al., 1988) showed that negative vocalizations tended to cluster together and that omitting fussing resulted in missing an important sign of infant distress. Also, by doing this, we reduced the number of outcome variables, hence reducing chance capitalization, and we obtained a robust measure of fussing and crying that was more in line with the general temperament measures used in the three earlier human studies on milk cortisol and behavior, namely negative affectivity, fearfulness, and fear reactivity (Grey et al., 2013; Glynn et al., 2007; Nolvi et al., 2017). It is important to note here that analyses in which fussing was excluded from the combined crying variable rendered comparable results.

**Control variables.** In line with Grey et al. (2013), Glynn et al. (2007), and Hinde et al. (2014), we controlled for maternal educational level and parity. Though infants with a birth weight lower than 2500 grams were excluded, birth weight was also included as a control variable because it can reflect infant vulnerability and predict both infant crying behavior (Milidou, Søndergaard, Jensen, Olsen, & Henriksen, 2014) and maternal stress (Halpern, Brand, & Malone, 2001). Furthermore, taking cortisol circadian variations into account and in line with previous research (Glynn et al., 2007; Grey et al., 2013; Nolvi et al., 2017), we controlled for collection time of breast milk sampling. Additionally, we controlled for the time interval between waking up and collection time of breast milk sampling.

Pearson correlations showed no significant relation between infant sex and number of Baby Day Diary assessments \( r = -.06, p = .629 \), nor between infant sex and number of breast milk assessments \( r = .07, p = .596 \). There were also no significant relations between breast milk cortisol concentrations and the number of Baby Day Diary assessments \( r = -.09, p = .274 \), and breast milk cortisol concentrations and number of breast milk cortisol assessments \( r = -.05, p = .553 \). Number of Baby Day Diary and breast milk cortisol assessments were thus not included as control variables.

### Missing data

An overview of sample flow and missing data is provided in Figure 1. Of the initial 77 breastfeeding mothers, two mothers did not return any breast milk samples and five mothers did not complete the crying diaries. Subsequently, we excluded samples of mothers reporting to breastfeed for less than 50% of the time, as we speculated that potential effects of breast milk cortisol would only be detectable if infants received substantial amounts of breastfeeding. Additionally, we excluded samples of mothers reporting illness and/or antibiotic use. Samples were also excluded when they were collected before 4:00 a.m. or after 12:00 p.m. since recent research indicated that cortisol in breast milk follows a circadian rhythm similar to cortisol in saliva (Pundir et al., 2017; van der Voorn et al., 2016).
At two weeks of infant age, breast milk samples were missing due to the following reasons: not able to collect (enough) breast milk \( (n=6) \), breastfeeding for less than half of the feedings \( (n=2) \), use of antibiotics \( (n=5) \), and collection before 4:00 a.m. or after 12:00 p.m. \( (n=5) \). Additionally, two mothers did not fill in the crying diary at two weeks of age.

At six weeks of infant age, breast milk samples were missing due to inability to collect breast milk \( (n=1) \), use of antibiotics \( (n=4) \), breast infection \( (n=1) \), stopped breastfeeding \( (n=4) \), use of formula for more than 50% of the time \( (n=1) \), and collection before 4 a.m. or after 12 p.m. \( (n=6) \). Additionally, one mother did not fill in the crying diary at six weeks of age.

At 12 weeks of infant age, breast milk samples were missing due to inability to collect (enough) breast milk \( (n=1) \), use of antibiotics \( (n=4) \), stopped breastfeeding \( (n=7) \), use of formula for more than 50% of the time \( (n=1) \), and collection before 4 a.m. or after 12 p.m. \( (n=5) \).

In total, 68 mothers filled in the crying diary at all time points, one mother missed one time point and one mother only filled in the crying diary once. With respect to milk samples, due to missing samples and exclusion criteria, we were able to include all three breast milk cortisol samples for 41 mothers, two for 12 mothers, and one for 10 mothers. As multilevel analyses are robust for missing data (Tabachnik & Fidell, 2007), families with missing data at one or two of the time points were still included in the overall analysis and no imputation was needed. Thus, the total number of participants in the analyses is 70.

**Statistical analyses**

Two values for milk cortisol were outside the range of \( +/-3 \) SD and were subsequently replaced with the highest cortisol value inside the range. Deleting these outliers instead of replacing them led to comparable results.

First, to investigate the stability of milk cortisol over time, multilevel models with time were used. To investigate whether breast milk cortisol is related to infant crying, three separate analyses were done; one for crying total duration, one for crying frequency, and one for crying bout. Due to the longitudinal nature of our design, multilevel (hierarchical) linear modelling (MLM), also known as mixed model analysis, was used. MLM is robust for missing data and is unaffected by unequal sample sizes (Tabachnik & Fidell, 2007). Therefore, there was no need to control for the fact that not every mother collected a breast milk sample at each time point, and we could run the analyses on the full data set \( (N=70) \).

MLM is conveyed as a set of regression equations. First, the intercepts-only model (a model without variables) is run in order to check whether a multilevel model is required, by means of the intraclass correlation. The intraclass correlations were 0.24 for crying total duration, 0.59 for crying frequency, and 0.72 for crying bout, which shows that multilevel analyses were appropriate. The participant was the level 2 identifier, and the outcome measure and other measures were the level 1 variables. Second, following Tabachnik and Fidell (2007), a build-up strategy was used. To the intercept-only model, variables are added one at a time. After each addition, the -2 log likelihood ratio scale after generalized least square estimation is examined. The -2 log likelihood is a determinant of model fit. If model fit increases, the added variable is kept. If model fit decreases, the added variable is cut from the model.

Linear time and quadratic time were entered first into the model, with linear time considered as a random factor. The time model that best improved the model fit was retained. Then, the control variables parity, infant sex, maternal educational level, birth weight, collection time, and the time interval between waking up and collection time of breast milk sampling were added one by one.

Thereafter, breast milk cortisol was added to the model. Then, interaction terms between breast milk cortisol and time, and between breast milk cortisol and quadratic time were added to examine the associations of breast milk cortisol with crying over time. Finally, the interaction term between infant sex and breast milk cortisol was added to examine sex differences. The final, best fitting models for crying total duration, crying frequency and crying bout are presented in the results. Final models were checked in regression analyses for normality of the residuals; residuals showed nearly normality. See Table 1 for Durbin-Watson indices and Variance inflation factors. Analyses were done using SPSS®, version 22.0.0.1 (IBM Corporation, Armonk NY, USA) for Windows®.

**RESULTS**

**Preliminary analyses**

On average, breast milk cortisol levels increased from 2 weeks through 12 weeks of infant age (see Table 2). To examine stability or changes in breast milk cortisol levels over time, multilevel analyses with time and quadratic time were performed (see Table 3). The time model proved a better fit; time was significant \( (t=2.23, p=.028) \). Bivariate Pearson correlations, not presented here, showed modest correlations \( (r’s \text{ ranging between } -.41 \text{ and } .68) \), between crying duration, frequency and bout, indicating that they measure different constructs, justifying the separate analyses.
Table 1

Regression Statistics

<table>
<thead>
<tr>
<th></th>
<th>Crying Total Duration</th>
<th>Crying Frequency</th>
<th>Crying Bout</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durbin-Watson Indexa</td>
<td>1.793</td>
<td>1.463</td>
<td>1.542</td>
</tr>
<tr>
<td>VIFb</td>
<td>1.070</td>
<td>1.110</td>
<td>1.130</td>
</tr>
<tr>
<td>Birthweight Infant</td>
<td>1.038</td>
<td>1.038</td>
<td>1.043</td>
</tr>
<tr>
<td>Education</td>
<td>1.044</td>
<td>1.044</td>
<td>1.044</td>
</tr>
<tr>
<td>Sex Infant</td>
<td>5.355</td>
<td>5.384</td>
<td>5.448</td>
</tr>
<tr>
<td>Collection Time</td>
<td>5.208</td>
<td>5.215</td>
<td>5.217</td>
</tr>
<tr>
<td>Milk cortisol</td>
<td>1.063</td>
<td>1.064</td>
<td>2.807</td>
</tr>
<tr>
<td>Interaction Infant Sex X Cortisol</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
</tbody>
</table>

aDurbin-Watson statistic is a measure of autocorrelation in residuals, and is always between 0 and 4. A value of 2 means no autocorrelation, values approaching 0 indicate positive and values approaching 4 indicate negative autocorrelation.
bVariance inflation factors (VIF) indicate how much the variance of the estimated regression coefficients are inflated. VIF=1 means not correlated, VIF 1-5 means moderately correlated, VIF 5-10 means highly correlated.

Table 2

Demographic Statistics for Maternal and Infant Characteristics, Separately for Boys and Girls

<table>
<thead>
<tr>
<th>Age of mother</th>
<th>Total (N=70) M (SD)</th>
<th>Boys (n=36) M (SD)</th>
<th>Girls (n=34) M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational Level</td>
<td>College: n=21 (30%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>University: n=37 (53%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cortisol (nmol/L)</td>
<td>2 weeks: 10.06 (9.72), n=52, 9.59 (9.03), n=27, 10.56 (10.57), n=25</td>
<td>12.26 (9.07), n=53, 12.62 (7.25), n=28, 11.87 (10.89), n=25</td>
<td>13.24 (8.82), n=52, 12.36 (8.34), n=22, 13.95 (9.26), n=29</td>
</tr>
<tr>
<td></td>
<td>6 weeks: 39.89 (1.53), n=70, 40.05 (1.51), n=36, 39.72 (1.55), n=34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gestational age at birth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight (in grams)</td>
<td>3544.23 (414.05), n=66, 3613.26 (462.20), n=34, 3470.88 (348.25), n=32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crying Bout</td>
<td>2 weeks: 17.03 (14.76), n=68, 14.84 (8.74), n=35, 19.35 (19.09), n=33</td>
<td>15.63 (12.73), n=69, 13.90 (6.99), n=36, 17.52 (16.84), n=33</td>
<td>13.60 (11.03), n=70, 12.17 (7.20), n=36, 15.12 (13.96), n=34</td>
</tr>
<tr>
<td></td>
<td>6 weeks: 39.95 (4.26), n=68, 30.35 (4.58), n=35, 8.08 (3.61), n=33</td>
<td>9.96 (4.58), n=69, 11.55 (4.87), n=36, 8.22 (3.35), n=33</td>
<td>7.36 (4.13), n=70, 8.17 (4.33), n=36, 6.50 (3.76), n=34</td>
</tr>
<tr>
<td></td>
<td>12 weeks: 131.85 (64.83), n=68, 135.14 (59.76), n=35, 128.35 (70.58), n=33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crying Frequency</td>
<td>2 weeks*: 9.25 (4.26), n=68, 10.35 (4.58), n=35, 8.08 (3.61), n=33</td>
<td>9.96 (4.58), n=69, 11.55 (4.87), n=36, 8.22 (3.35), n=33</td>
<td>7.36 (4.13), n=70, 8.17 (4.33), n=36, 6.50 (3.76), n=34</td>
</tr>
<tr>
<td></td>
<td>6 weeks*: 134.56 (62.88), n=69, 145.83 (61.45), n=36, 122.27 (63.02), n=33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12 weeks: 89.23 (48.21), n=69, 92.36 (48.98), n=36, 85.82 (47.88), n=33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: For t-tests for the difference between female and male infants: *=p<.10, *=p<.05. Crying levels are in the normal range for healthy infants of these ages (Wolke et al, 2017). Mean bout length of each crying episode in minutes, mean number of episodes per 24 hr, mean duration in minutes per 24 hr

Table 3

Mixed Model Analysis for Breast Milk Cortisol across Time

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>9.65</td>
<td>1.35</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time</td>
<td>0.34</td>
<td>0.15</td>
<td>.028</td>
</tr>
</tbody>
</table>

Deviance 1040.345

N=70

Multilevel analyses

Crying total duration. The best fitting multilevel model for crying total duration showed a significant association between maternal educational level and crying duration (p=.041). Infants from mothers with higher educational levels cried longer, in total, than infants from mothers with lower educational levels. There was no significant association between milk cortisol and infant crying total duration. Finally, there were no significant relations between our other control variables (i.e., parity, infant sex, birth weight, collection time, and time from waking to collection), or any of the interaction terms and infant crying total duration (see Table 5).

Crying frequency. The best fitting multilevel model for crying frequency showed a significant association between quadratic time (p=.006), infant sex (p=.006), maternal educational level (p=.012) and crying frequency. In line with the normative development of infant crying (Barr et al., 2006), crying frequency significantly increased from time 1 to time 2, and decreased from time 2 to time 3. Overall, female infants cried significantly less frequently than male infants. Infants from mothers with higher educational levels cried more frequently than infants from mothers with lower educational levels. There was no significant relation between breast milk cortisol and infant crying frequency. Finally, no significant associations between any of our other control variables (i.e., parity, birth weight, collection time, and time from waking to collection), or any of the interaction terms and infant crying frequency were found (see Table 5).
### Table 4

**Correlations between Breast Milk Cortisol and Infant Fussing and Crying**

<table>
<thead>
<tr>
<th></th>
<th>Whole sample</th>
<th></th>
<th></th>
<th></th>
<th>Male infants</th>
<th></th>
<th></th>
<th></th>
<th>Female infants</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cortisol</td>
<td>2 weeks</td>
<td>6 weeks</td>
<td>12 weeks</td>
<td>Cortisol</td>
<td>2 weeks</td>
<td>6 weeks</td>
<td>12 weeks</td>
<td>Cortisol</td>
<td>2 weeks</td>
<td>6 weeks</td>
</tr>
<tr>
<td>2 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cortisol</td>
<td>1</td>
<td>.36**</td>
<td>.54**</td>
<td></td>
<td>1</td>
<td>.51*</td>
<td>.68**</td>
<td></td>
<td>1</td>
<td>.30</td>
<td>.47*</td>
</tr>
<tr>
<td>6 weeks</td>
<td></td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td>.06</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crying Bout</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 weeks</td>
<td>-.17</td>
<td>.10</td>
<td>-.12</td>
<td></td>
<td>-.25</td>
<td>-.24</td>
<td>-.41</td>
<td></td>
<td>-.17</td>
<td>.35</td>
<td>.06</td>
</tr>
<tr>
<td>6 weeks</td>
<td>-.11</td>
<td>.20</td>
<td>-.20</td>
<td></td>
<td>-.19</td>
<td>-.19</td>
<td>-.45*</td>
<td></td>
<td>-.07</td>
<td>.47*</td>
<td>-.03</td>
</tr>
<tr>
<td>12 weeks</td>
<td>-.03</td>
<td>-.32*</td>
<td>.04</td>
<td></td>
<td>-.13</td>
<td>-.17</td>
<td>-.39</td>
<td></td>
<td>.05</td>
<td>.44**</td>
<td>.32</td>
</tr>
<tr>
<td>Crying Frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 weeks</td>
<td>.18</td>
<td>.14</td>
<td>.15</td>
<td></td>
<td>.21</td>
<td>.20</td>
<td>.25</td>
<td></td>
<td>.21</td>
<td>.09</td>
<td>.14</td>
</tr>
<tr>
<td>6 weeks</td>
<td>.18</td>
<td>.14</td>
<td>.13</td>
<td></td>
<td>.17</td>
<td>.16</td>
<td>.34</td>
<td></td>
<td>.27</td>
<td>.12</td>
<td>.05</td>
</tr>
<tr>
<td>12 weeks</td>
<td>.11</td>
<td>.11</td>
<td>.00</td>
<td></td>
<td>.15</td>
<td>.11</td>
<td>.07</td>
<td></td>
<td>.12</td>
<td>.10</td>
<td>-.02</td>
</tr>
<tr>
<td>Crying Duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 weeks</td>
<td>.02</td>
<td>.35*</td>
<td>-.02</td>
<td></td>
<td>-.07</td>
<td>.21</td>
<td>-.24</td>
<td></td>
<td>.08</td>
<td>.46*</td>
<td>.16</td>
</tr>
<tr>
<td>6 weeks</td>
<td>.10</td>
<td>.12</td>
<td>-.06</td>
<td></td>
<td>-.03</td>
<td>-.01</td>
<td>-.21</td>
<td></td>
<td>.24</td>
<td>.22</td>
<td>.11</td>
</tr>
<tr>
<td>12 weeks</td>
<td>.01</td>
<td>.01</td>
<td>-.12</td>
<td></td>
<td>.01</td>
<td>-.03</td>
<td>-.31</td>
<td></td>
<td>.03</td>
<td>.04</td>
<td>.04</td>
</tr>
</tbody>
</table>

Note: *p < 0.05, **p < 0.01; Correlations shown are based on the full data set and pairwise deletion.

### Table 5

**Estimates for the Best Fitting Model for Fussing and Crying Total Duration, Frequency, and Bout**

<table>
<thead>
<tr>
<th></th>
<th>Total Duration</th>
<th>Frequency</th>
<th>Bout</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
<td>p</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>3.34</td>
<td>5.66</td>
<td>.557</td>
</tr>
<tr>
<td>Time Quadratic</td>
<td>-0.59</td>
<td>0.39</td>
<td>.132</td>
</tr>
<tr>
<td>Education</td>
<td>8.71</td>
<td>4.14</td>
<td>.041</td>
</tr>
<tr>
<td>Birthweight Infant</td>
<td>0.00</td>
<td>0.01</td>
<td>.904</td>
</tr>
<tr>
<td>Collection Time</td>
<td>0.00</td>
<td>0.00</td>
<td>.128</td>
</tr>
<tr>
<td>Interval Between Waking and Collection</td>
<td>0.01</td>
<td>0.09</td>
<td>.892</td>
</tr>
<tr>
<td>Cortisol</td>
<td>0.25</td>
<td>0.55</td>
<td>.652</td>
</tr>
<tr>
<td>Sex Infant¹</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Interaction Sex Infant * Cortisol</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Deviance</td>
<td>1452.327</td>
<td>699.009</td>
<td>924.534</td>
</tr>
</tbody>
</table>

Note: SE=Standard Error, na=not included in the final model; N=70

¹0=male, 1=female
Crying bout. The best fitting multilevel model for crying bout length showed a significant association between time ($p=.016$) and crying bout length. On average, crying bout length decreased over time. There was no significant relation between milk cortisol and infant crying bout length ($p=.189$). Furthermore, there was a significant association between infant sex ($p=.037$) and crying bout length, indicating that female infants tended to have longer crying bouts than male infants. Finally, there were no significant relations between our control variables (i.e., parity, maternal educational level, birth weight, collection time, and time from waking to collection), quadratic time, or any of the interaction terms and crying bout length (see Table 5).

DISCUSSION

The present study aimed to longitudinally investigate the association between breast milk cortisol and infant crying during the first twelve postnatal weeks. Higher concentrations of milk cortisol were expected to be related to higher levels of infant fussing and crying. Furthermore, we explored sex differences in the link between milk cortisol and infant fussing and crying. Results showed that milk cortisol concentrations increased from 2 weeks through 12 weeks of infant age. Milk cortisol concentrations were unrelated to the total duration, frequency, and bout length of infant fussing and crying for both boys and girls.

The results are not in line with the two previous studies that showed higher breast milk cortisol concentrations to be related to more infant temperamental negativity (Grey et al., 2013; Nolvi et al., 2017). Based on these studies, we expected milk cortisol to be related to infant fussing and crying. There are several possible explanations for the current study’s lack of relations between milk cortisol and infant fussing and crying in the first three months postpartum. First, differences in the designs and measures used may explain the differences in results. While in this study fussing and crying over three 72-hr periods were assessed within the first 3 months, the Grey et al. (2013) study used maternal reports of temperamental negative affectivity (including sadness and fear) at 3 months and the Nolvi et al. (2017) study observed laboratory-induced fear reactivity at 8 months. This variation in designs and outcome measures may be behind the differences in results between the three studies. Also, while Nolvi et al. (2017) found a relation between milk cortisol and girls’ fear reactivity in a laboratory setting, but no relation between milk cortisol and maternal reported fearfulness, Grey et al. (2013) did find a relation between milk cortisol and maternal reported infant negative affectivity. This positive relation between milk cortisol and negative affectivity seemed to be strongest for the fear subscale of negative affectivity (Grey et al., 2013). Thus, it is possible that milk cortisol is more specifically related to child fearfulness instead of to more general measures of fussing and crying, as were assessed in the current study. Clearly, more research is needed in this field before conclusions on relations between milk cortisol and infant behavior can be drawn.

A second explanation for the lack of relation between breast milk cortisol and infant fussing and crying might be related to our restriction of milk collection to the morning. Just as in saliva, cortisol follows a circadian rhythm in milk, with high concentrations in the morning and a decline throughout the day (Pundir et al., 2017; van der Voorn et al., 2016). Our restriction of milk collection to the morning, when milk cortisol is naturally high, might explain why we found no association between breast milk cortisol and infant behavior. In laboratory stressors, greater cortisol reactivity is obtained when tests are conducted in the afternoon (Dickerson & Kemeny, 2004). Thus, if our assumption that higher milk cortisol reflects environmental stressors of the mother is true, differences between mothers might be most notable in milk samples collected in the afternoon or evening. Additionally, future research into the relations between milk cortisol and infant behavior should take several milk samples throughout the day, including afternoon and evening samples, to calculate the total cortisol output an infant is exposed to, as also suggested by Finken et al. (2017).

The longitudinal design of the present study allowed us to describe the time course and stability of cortisol concentrations in breast milk throughout the first three months. On average, milk cortisol levels increased from 2 weeks through 12 weeks of infant age. One explanation for this interesting finding is that many mothers in the Netherlands, given the length of maternity leave, transition back to work around 10-12 weeks postpartum after giving birth. This transition poses many stressors for mothers, including readjusting to work, combining working and family life, and getting used to being separated from the infant (Costigan, Cox, & Cauce, 2003; Nichols & Roux, 2004; Wiese & Heidemeier, 2012). Although transitioning back to work also brings positive emotions, research indicates that mothers perceive it primarily as stressful (Nichols & Roux, 2004). Transitioning back to work, and/or the anticipation thereof, might thus have led to increasing milk cortisol concentrations over the first three months of lactation.

Furthermore, breast milk cortisol concentrations at 2 weeks were positively and significantly related to breast milk cortisol concentrations at 6 and 12 weeks, suggesting some intra-individual stability over time. Yet breast milk cortisol levels at 6 weeks were not related to breast milk cortisol concentrations at 12 weeks. As the link between milk cortisol and infant fussing and crying behavior might be bi-directional, it is possible that infant fussing and crying leads to more maternal stress (Barr et al., 2014; Brand et al., 2014; Fujiwara et al., 2011), and subsequently, higher milk cortisol concentrations. Due to a relatively small sample size, we were unable to employ a cross-lagged panel
design to shed more light on the bi-directional relations between crying behavior and milk cortisol concentrations. However, at around 6 weeks of age, infants are at the peak of the normal developmental crying curve (Barr et al., 2006), and this increased crying might have influenced maternal milk cortisol concentrations at 6 weeks, explaining the lack of a relation between milk cortisol concentrations at 6 weeks and at 12 weeks. Nevertheless, previous research did not find an association between maternal perceived stress, anxiety, and depression and breast milk cortisol concentrations (Grey et al., 2013; Nolvi et al., 2017). Hence, more research is needed to investigate milk cortisol stability and its determinants, including maternal psychological complaints.

**Future directions**

There are several suggestions for future investigations. First, in line with the lactational programming hypothesis, we hypothesized that the mother, via her milk, transfers environmental signals to the infant who in turn adapts his or her behavior accordingly. However, this process has been challenged by the idea that the mother not simply signals information about the environment, but also fine-tunes these signals. Wells (2014) suggested that maternal physiology is capable of smoothing over short-term signals which buffer against ecological stresses. In line with this idea, future research could further examine when, and under which conditions, ecological stresses are reflected in maternal milk cortisol concentrations, and also when, and under which conditions, maternal milk cortisol concentrations predict child outcomes.

Second, research on breast milk cortisol is as yet scarce. Fundamental research on the biology of lactational programming is needed to better understand potential mechanisms underlying hypothesized effects of milk cortisol. For example, research focusing on possible changes in intestinal permeability and receptor expression in the infants’ intestines could help disentangle underlying mechanisms. Animal research has shown that especially early in life, the infant’s intestinal tract has a high number of glucocorticoid receptors, which have been assumed to create an intersection between the gut-brain axis and the hypothalamic-pituitary-adrenal (HPA) axis (Allen-Blevis et al., 2015). Intestinal glucocorticoid receptors decline after the mother is no longer breastfeeding, suggesting a special role for milk-borne cortisol on offspring behavior (Hinde et al., 2014), as milk cortisol would bind to these receptors producing effects that are likely to influence infant behavior (Allen-Blevis et al., 2015). In future (animal) research, it would be interesting to examine whether breast milk cortisol has an effect on infant biobehavioral regulation, and whether intestinal permeability and receptor expression determine a certain window of sensitivity to breast milk cortisol.

Third, other infant outcomes and their relation with breast milk cortisol should be investigated. Next to temperament and infant negativity, we recommend examining the relation between breast milk cortisol and other developmental outcomes influenced by the limbic system, including long-term memory and learning. A variety of animal and human research has already shown that high exposure to cortisol, albeit in the blood stream, can impair explicit learning and plasticity in the limbic system (reviewed in Sapolsky, 2003). Furthermore, animal studies showed that higher cortisol concentrations in breast milk were related to improved learning in rat offspring (Catalani et al., 2002).

Fourth, another direction for future research is related to possible influences of breast milk cortisol on infant physical health. Hinde et al. (2014) suggested that milk cortisol binding to infants’ intestinal tract cortisol receptors influences the maturation of the gastrointestinal tract. A healthy development of the gastrointestinal tract in turn, is related to the infant’s developing immune system and health (Bäckhed, 2011; Dimmitt et al., 2010; Sudo et al., 2004). For example, a recent study found that 3-month-old infants exposed to higher breast milk cortisol concentrations showed lower body mass index percentile gains over the first two postnatal years and lower body mass index percentile than infants exposed to lower levels of cortisol (Hahn-Holbrook, Le, Chung, Davis, & Glynn, 2016). Future research should therefore also focus on the relation between breast milk cortisol and infant physical health, including illnesses and gut microbiota.

Fifth, on a more general note, future studies on relations between breast milk cortisol and infant outcomes, including biobehavioral regulation, cognitive outcomes, and physical health, would also benefit from taking milk volumes into account. Milk volumes are expected to vary across the first three months of life, as a function of infant mass, gastrointestinal capacity, and behavioral development (Hinde, 2013; Hinde et al., 2014). Thus, even if milk cortisol concentrations were more or less stable, since volume is increasing, the total transfer of cortisol to the infant may be actually increasing.

Lastly, as our current longitudinal study only spans a short period of time and previous research is mostly cross-sectional, future research should also investigate whether early breast milk cortisol relates to offspring behavior at later ages. Consistent with the idea of lactational programming, the effects of early breast milk cortisol exposure might extend to, or even become clearer, in childhood or adulthood.

**Strengths and limitations**

A strong point of this study is its longitudinal nature. Mothers collected breast milk samples and filled in infant behavior diaries for three consecutive days when the infant was 2, 6, and 12 weeks of age. The current study also has limitations to note. First, even though the diary used to measure infant crying and fussing behavior has been shown to be a valid maternal self-report measure that is comparable to audio recordings (Barr, et al. 1988), it might still be possible that mothers over- or underreported infant crying and fussing behavior. Furthermore, the cry diary uses 5-minute intervals, so the diary-
recorded crying bout might also reflect intermittent clusters of crying behaviors (Barr, Paterson, MacMartin, Lehtonen, & Young, 2005). Future research should therefore also include audio recordings or observations. Second, although the study controlled for time of day effects by only including morning samples and controlling for collection time, only one sample was collected at each assessment moment. As discussed above, multiple daily samples may provide additional and more reliable information on breast milk cortisol and related lactational programming processes. Finally, it is important to note that, due to our exclusion criteria or because mothers stopped breastfeeding, only 41 of the 70 mothers could be included at all three time points. As our multilevel analyses are robust for missing data, and unequal sample sizes pose no problems (Tabachnik & Fidell, 2007), this pattern did not affect our results. However, future studies should find ways to decrease missingness in order to improve overall sample size.

CONCLUSION

While the health benefits of breast milk are long known, the role that biological constituents present in breast milk may have in programming child development are just beginning to be uncovered. The present longitudinal study is the first to investigate the association between human milk cortisol and infant crying over the first three months of life. On average, milk cortisol levels increased from 2 weeks through 12 weeks of infant age. No evidence was found for relations between milk cortisol and total duration, frequency, and bout length of fussing and crying. As the research on breast milk cortisol and lactational programming is in its initial phases, future studies, preferably with larger populations, are highly needed to replicate and extend these findings.

REFERENCES


Are Cortisol Concentrations in Human Breast Milk Associated with Infant Crying?

Chapter 6


Chapter 7

Summary, Conclusion, and General Discussion
7.1 Summary of the Thesis

The broader goal of the current thesis was to contribute scientific knowledge to the field of early development. The thesis had two more specific aims. The first aim was to shed more light on possible prenatal predictors of the quality of parental caregiving in the first months after the birth of the infant (Chapters 2, 3, and 4). The second aim was to examine possible mechanisms underlying the link between maternal stress and child outcomes (Chapters 5 and 6). With the exception of Chapter 2 (which used a student sample), the studies in the current thesis were carried out within the longitudinal BINGO (Dutch acronym for Biological Influences on Baby's Health and Development) project, in which mothers and fathers were followed from the third trimester of pregnancy till 12 weeks after the child was born.

Chapters 2, 3, and 4 concern the first aim of the current thesis: to investigate possible prenatal predictors of the quality of postpartum parental caregiving. The quality of parental caregiving has been shown to affect children's development from birth onwards. Therefore, it is important to detect parents at risk for low-quality caregiving as early as possible – preferably already before birth.

Chapter 2. The study in Chapter 2, conducted with a student sample, was used as a preliminary study. The task completed by the students was used with expectant parents in the next study (Chapter 3). In a sample of 120 university students, cognitive and emotional reactions to infant crying were investigated. The aim was to determine whether young childless adults show cognitive disturbances and negative emotions when exposed to infant crying, compared to other disturbing noises, and whether these cognitive disturbances and negative emotions were associated. We also examined whether these cognitive and emotional reactions to infant crying were related to personality characteristics (i.e., empathy, ego-resiliency, hostility, and impulsivity). Participants performed a working memory task while being subjected to different disturbing noises, including infant crying.

Results showed that participants had the least correct trials on the task and showed the most negative emotions when hearing infant crying, as compared to the other noises. Personality characteristics were not associated with the participants' cognitive performance and emotional reactions to infant crying. Negative emotions were associated with more incorrect trials on the working memory task in all conditions except in the infant crying condition. Thus, cognitive disturbances and negative emotions may be general adult responses to infant crying, which suggests a broadly present human emotional and cognitive response to infant crying.

Chapter 3. In this study, observations of expectant mothers’ and fathers’ behavior when exposed to infant crying were examined as predictors of the quality of caregiving towards their own infant six weeks postpartum. Eighty-eight expectant mothers and 57 of their male partners were tested during the third trimester of pregnancy. Both parents were filmed individually while caring for a crying simulator infant during a 15-minute period. The quality of their caregiving was observed afterwards and indicated by ratings on their sensitivity and cooperation towards the simulator infant. Furthermore, cognitive interference on a working memory task (see Chapter 2) and the ability to regulate physical force when exposed to infant crying were assessed. Six weeks after the birth of their infant, parents were filmed and rated for sensitivity and cooperation during a 15-minute interaction with their own infant at home.

Results showed that, in both mothers and fathers, prenatal quality of caregiving behavior towards the simulator infant significantly and positively predicted postnatal quality of caregiving towards the own infant. Cognitive interference and the ability to regulate physical force when exposed to infant crying were not associated with postnatal quality of caregiving. Thus, prenatal observations of caregiving with a simulator infant can predict the postnatal quality of caregiving. However, more research is needed to translate these findings into practice, for example for screening parents-at-risk, and for teaching and training purposes.

Chapter 4. As in Chapter 3, here too prenatal prediction of the quality of parental caregiving towards the own infant six weeks postpartum was the focus of interest. In the study in this chapter (Chapter 4), baseline and reactivity measures of the hormones cortisol and testosterone were used as prenatal predictors. In the BINGO study sample of 88 expectant mothers and 57 of their male partners, cortisol and testosterone in saliva were measured before, during, and after an interaction with a simulator infant during late pregnancy. Expectant parents were, as described in Chapter 3, also rated for prenatal quality of caregiving behavior. When the infant was 6 weeks old, parents were visited at home and filmed individually during a 15-minute interaction with their infant. Cortisol and testosterone in saliva were again measured. The quality of their caregiving behavior was measured by rating the parents on sensitivity and cooperation.

Baseline and reactivity measures of cortisol and testosterone were used as prenatal predictors of postnatal quality of caregiving, controlled for prenatal quality of caregiving, and postnatal cortisol and testosterone levels. Results showed that there were no relations between maternal caregiving behavior and cortisol and testosterone concentrations. However, in fathers, higher overall prenatal cortisol levels were associated with lower postnatal quality of caregiving. Additionally, the interaction between cortisol and testosterone significantly predicted fathers' quality of caregiving, both in the prenatal and postnatal period. The findings are interpreted in the context of hyper-reactive physiological responses observed in parents at risk for insensitive caregiving. Further research is necessary to explain these results; for example, to understand why cortisol and caregiving behavior is related across the perinatal period in fathers, but not in mothers.
Chapters 5 and 6 pertain to the second aim of this thesis, namely to examine possible underlying biological mechanisms by which maternal stress is related to infant development. Recent research indicates that maternal psychosocial stress during pregnancy and maternal early postpartum stress potentially affects infant birth outcomes and child emotional and behavioral development, but the mechanisms underlying this relation are only partly understood. The present thesis focused on two biological mechanisms that received little attention in the past.

Chapter 5. One mechanism explaining the relation between maternal prenatal stress and child outcomes might be that maternal prenatal psychosocial stress affects child development and health via the mother’s, and in turn the child’s, intestinal microbiota. The study reported in Chapter 5 investigated the first step of this potential mechanism, namely the relation between psychosocial stress and fecal microbiota in pregnant mothers. Seventy expectant mothers collected a stool sample during the third trimester of pregnancy to investigate intestinal microbiota. To measure psychosocial stress, these expectant mothers were also asked to fill in questionnaires on general and pregnancy-specific stress and anxiety. It was hypothesized that mothers with high psychosocial stress would have phylum-level microbial compositions characterized by more Proteobacteria and less Actinobacteria and genus-level microbial compositions characterized by more Lactobacilli, compared to mothers with low reported psychosocial stress.

Contrary to the hypothesis, results showed no association between psychosocial stress and phylum-level microbial composition. A significant association between maternal general anxiety and microbial composition on genus level was found. Mothers with high general anxiety had, amongst others, higher relative abundance of Peptostreptococcaceae Incertae Sedis and unclassified genera in the Peptococcaceae family, than mothers with low general anxiety. Furthermore, mothers with low general anxiety had, amongst others, higher relative abundance of Eubacterium and Oscillospira than mothers with high general anxiety. No statistically significant associations between other measured psychosocial stress variables and the relative abundance of genus-level bacteria were detected.

This study investigated the first step of a mechanism potentially underlying links between prenatal psychosocial stress and infant outcomes, namely the relation between psychosocial stress and fecal microbiota in pregnant mothers. Contrary to the hypothesis, it was not found that mothers with high psychosocial stress had phylum-level microbial compositions characterized by more Proteobacteria and less Actinobacteria, and less Lactobacilli on genus level, compared to mothers with low reported psychosocial stress. However, a significant association between late pregnancy general anxiety and the women’s fecal microbial composition at genus level was found. This study therefore offers insights into associations between maternal mental health and gut microbial composition during pregnancy and provides a starting point for future investigations in which maternal diet as well as infant microbiota and development should also be assessed.

Chapter 6. This study focused on another possible biological mechanism that could explain the link between maternal stress and infant developmental outcomes; this mechanism entails that maternal stress would affect the infant through the composition of maternal breast milk. The present study examined part of this mechanism by longitudinally investigating whether maternal milk cortisol is associated with infant fussing and crying behavior. Seventy mothers provided milk samples when the infant was 2, 6, and 12 weeks of age. At the same time points, mothers filled in an infant cry diary for three subsequent days. Three measures of crying, that included crying, unsovable crying, and fussing, were derived from the cry diary: total duration (mean duration in minutes per 24 hours), frequency (mean number of episodes per 24 hours), and bout length (mean bout length of each episode in minutes). Results showed that breast milk cortisol concentrations, on average, increased from 2 to 12 weeks of infant age. However, breast milk cortisol was not related to any of the three measures of infant crying. While the health benefits of breast milk are long known, the role that biological constituents present in breast milk may have in programming child development are just beginning to be uncovered. Future studies, preferably with larger populations, are highly needed to replicate and extend these findings.

7.2 Conclusions

Taken together, the results of the five studies presented in this thesis can be summarized in the following main conclusions:

- Exposure to infant crying when performing a working memory task is related to decreased performance and increased negative emotions in young childless adults.
- Postnatal quality of parental caregiving behavior can, for both mothers and fathers, be predicted from prenatal observations of caring for a crying simulator infant, and, for fathers, also from cortisol concentrations during the interaction with this simulator infant.
- No evidence was found for associations between maternal psychosocial stress and intestinal phylum-level bacteria composition during late pregnancy. Associations between maternal general anxiety and intestinal bacteria composition at the end of pregnancy were found at genus-level.
- No evidence was found for an association between breast milk cortisol and infant fussing and crying in the first three months of life.
7.3. General Discussion

7.3.1 Early Development: Links Between Caregiving, Maternal Stress, and Child Outcomes

The broader goal of the current thesis was to contribute scientific knowledge to the field of early development, by focusing on two specific areas that were approached through two research aims. The first research aim was to shed more light on possible prenatal predictors of the quality of parental caregiving in the first months after birth. The second research aim was to examine possible mechanisms underlying the link between maternal stress and child outcomes. Both aims were addressed within the same participant group from the same longitudinal study. Specific research questions regarding the first and second aim of this thesis were examined in separate studies, reported in separate chapters of this thesis. For a schematic overview of the associations between the main study variables examined in the different chapters, see Figure 1.2 in the introductory chapter of this thesis. In short, this thesis looked at early caregiving and stress separately. However, there is evidence in the literature that caregiving and maternal stress are associated and may interact with each other in their impact on child development. In the following paragraphs, possible links between caregiving and maternal stress, and how they might be addressed in future studies, will be discussed.

Previous research has acknowledged that caregiving and stress can act together in influencing child development. For example, the Early Life Stress (ELS) model by Loman and Gunnar (2010) recognizes that early life stress can negatively influence child development, but that sensitive parental caregiving can potentially buffer these negative effects. Research has found evidence for the proposed buffering effect of high quality caregiving. For example, children exposed to intimate partner violence between their parents showed heightened stress responses at 24 months of age, but not when the mother was sensitive (Hibel, Granger, Blair, & Cox, 2011), and children of anxious mothers showed more negativity, but not when the mother was sensitive (Kertz, Smith, Chapman, & Woodruff-Borden, 2008).

While the current thesis could have linked quality of caregiving and maternal stress by adding analyses, caregiving and stress were not investigated together because the sample size did not permit including extra variables in the current studies, due to power issues. Low power reduces the likelihood that a statistically significant result in the current data set is actually a true effect in the main population. However, the current thesis provides relevant findings that serve as a starting point for future research. Several suggestions for these further investigations can be made, see Figure 7.3.1.
The first suggestion of linking quality of caregiving and stress is based on the investigation of breast milk cortisol and its relation to infant crying, see arrow A in Figure 7.3.1. Since no effect of breast milk cortisol on infant crying was found (see Chapter 6), it would be interesting to see whether breast milk cortisol would be related to infant crying in subgroups based on maternal quality of caregiving. For example, it might be that high levels of breast milk cortisol are related to high levels of infant crying, but only when the mother shows low quality of caregiving. Furthermore, as the fathers participated in the BINGO study as well, future research could investigate whether the father has a buffering effect in the relation between maternal stress and infant outcomes, and in the relation between low quality maternal caregiving and infant outcomes. As has been shown previously in other studies, care provided by people other than the mother can buffer many negative effects of the mother on the child (NICHHD Early Child Care Research Network, 1997). For example, in cases where mothers show low sensitivity because of depression, highly sensitive fathers can buffer the negative effects of maternal depression on family cohesion (Vakrat, Apter-Levy, & Feldman, 2017).

A second suggestion to link stress and quality of caregiving is represented by arrow B in Figure 7.3.1 and involves the idea that stress itself can be a determinant of the quality of caregiving. Different sources of stress, such as low partner or social support, low parental ego-resiliency, difficult infant temperament, and caring for triplets as compared to twins or singletons have been shown to negatively affect the quality of parenting (see, e.g., Belsky, 1984; Booth, Macdonald, & Youssef, 2018; Feldman, Eidelman, & Rotenberg, 2004; Rodgers, 1998; van Bakel & Riksen-Walraven, 2002). Additionally, mothers high in trait-anxiety - which can also be considered a form or source of psychosocial stress - were less sensitive and responsive towards their infants (Nicol-Harper, Harvey, & Stein, 2007). There is less research on the link between stress and quality of caregiving in fathers, but one study suggests that high reported stress within the first 6 months of infant life is related to less positive paternal behavior toward the infant (Darke & Goldberg, 1994).

There is also some evidence that even prenatal stress can affect postnatal quality of caregiving. Animal studies showed that caregiving behavior is affected by stress during pregnancy (Rutherford et al., 2014; Bosch, Müssch, Bredewold, Slattery, & Neumann, 2007; Smith, Seckl, Evans, Costall, & Smythe, 2004). Human studies mostly show indirect effects. For example, pregnancy requires psychosocial and emotional adjustments, which can be stressful (Da Costa, Larouche, Drista, & Bender, 1999). These adjustments can be less stressful when the mother has effective psychosocial support (Elsenbruch et al., 2007; Westdahl et al., 2007). This prenatal support is also, indirectly, positively related to postnatal quality of caregiving (Thomas, Letourneau, Bryce, Campbell, & Giesbrecht, 2017). However, there is currently no research in humans showing associations between prenatal stress and postnatal quality of caregiving.

In this thesis, it was shown that the maternal and paternal prenatal quality of caregiving when interacting with a simulator infant (see Chapter 3), and for fathers also the cortisol concentrations during this interaction (see Chapter 4), predicted postnatal quality of caregiving when interacting with the own infant. Future, larger studies might want to add psychosocial stress, measured prenatally, to the predictions of postnatal quality of caregiving to see whether 1) prenatal psychosocial stress is related to postnatal parental caregiving, and 2) whether psychosocial stress predicts parental caregiving above and beyond parental prenatal caregiving behavior and paternal prenatal cortisol concentrations. An investigation of the relationship between prenatal stress and postnatal quality of care would add to the preliminary evidence that prenatal stress predicts postnatal quality of caregiving.

7.3.2. Sample Limitations

The total sample of the current thesis was, compared to other research, relatively small, and in some cases the relatively small number of participants prevented additional analyses. For example, in Chapter 6, the bi-directional relationship between breast milk cortisol and infant crying could not be tested, since the necessary cross-lagged panel design analyses would need more participants. Also, fathers were less interested in participation than mothers and/or were not willing or able to take the time for participation. Future research should find out which incentives are necessary for fathers to participate in research on pregnancy and infant development. Nonetheless, 57 fathers agreed to participate, and 88 mothers started the study. Given the time constraints and the longitudinal nature of the BINGO study with two labour-intensive visits and additional measurement times, the number of participants who completed the BINGO study can be considered satisfactory.

Another limitation is that the sample characteristics limit generalizability to the whole population, as our sample was healthy and well-educated. Future research should put effort in drawing a sample that is more representative of the whole population. For example, the number of participants from different socioeconomic backgrounds and from high- and low-risk groups could be matched.

7.3.3 Future Directions

Next to the ideas proposed above, there are several other recommendations for future research. First, the current thesis investigated whether breast milk cortisol is related to infant crying and found no such relationship (see Chapter 6). Nonetheless, as research on breast milk cortisol and its influences on infant development is still scarce, there are several important suggestions and implications for future research. For example, future investigations could examine when, and under which conditions, ecological stresses
are reflected in maternal milk cortisol concentrations, and also when, and under which conditions, maternal milk cortisol concentrations predict child outcomes. Furthermore, future research could examine whether breast milk cortisol has an effect on infant biobehavioral regulation, and, by using animal models, whether intestinal permeability and receptor expression determine a certain window of sensitivity to breast milk cortisol. Animal research has shown that receptor expression is relatively high during early infancy (Allen-Blevis, Sela, & Hinde, 2015) and declines after weaning (Hinde et al., 2014), suggesting that milk cortisol has a higher influence on infant behavior that develops early in life or has its critical or sensitive period in early infancy. Related to this, future research should extend the period of investigation, to examine whether effects of early breast milk cortisol exposure extend to, or become clearer, in childhood or even adulthood. Finally, maternal prenatal and postnatal stress and its influence on breast milk cortisol could be investigated, to examine whether reported stress and breast milk cortisol concentrations are associated. In saliva, reported stress and cortisol concentrations are not always correlated (Dickerson & Kemeny, 2004), so the same might be true for breast milk cortisol concentrations.

Second, additional data could be collected to gain a deeper understanding of the mechanisms behind the links that are often reported between maternal stress and infant development. Chapter 6 investigated a link between breast milk cortisol and infant crying and found no relationship between the two. The chapter focused on infant crying as this is one of the most prominent and evolutionarily adaptive infant behaviors in the first three months of life, so theoretically, crying was assumed to be a relevant outcome measure. Nonetheless, there are other infant outcomes that could be investigated in relation to milk cortisol. For example, future research might investigate possible influences of breast milk cortisol on infant health. Hinde et al. (2014) suggested that the infants’ intestinal tract has cortisol receptors that bind ingested cortisol, thereby influencing the maturation of the gastrointestinal tract. A healthy development of the gastrointestinal tract, in turn, is related to the infants’ developing immune system and health (Bäckhed, 2011; Dimmitt et al., 2010; Sudo et al., 2004). Thus, ingested milk-borne cortisol could assist a healthy development of the infants’ intestines, which in turn could lead to an immune system that can more readily fight off pathogens and facilitate protection against potentially injurious microorganisms (Dimmitt et al., 2010; Sudo et al., 2004). This, in turn, can benefit the infants’ overall health. Future research might therefore include reports of infant illnesses, to examine whether breast milk cortisol concentrations are related to infant health. Additionally, it has been suggested that the limbic region in the brain is especially receptive for cortisol ingested via breastmilk (Glynn et al., 2007). The limbic region is involved in the regulation of anxiety, fear, and behavioral inhibition (Glynn et al., 2007). Future research could therefore include reports of such behaviour, to investigate whether breast milk cortisol is associated with these types of behavior.

Another area in which additional data could be collected to gain a deeper understanding of the mechanisms behind the links that are often reported between maternal stress and infant development, is within the field of microbiota. Chapter 5 investigated a possible link between maternal prenatal psychosocial stress and maternal microbiota and found that mothers with higher reported anxiety had a different microbiome composition than mothers with lower reported anxiety. However, the bacteria found in most mothers and which seemed to underlie the differences, *Peptostreptococcaceae* and *Peptococcaceae*, are strongly affected by diet (e.g., Bridgewater et al., 2017; Fan, Liu, Song, Chen, & Ma, 2017; Zhang et al., 2013). Thus, in order to understand the found differences in microbiome composition, it might be necessary to include dietary information in future studies. These future studies could potentially assess whether anxiety directly affects the microbiome through activation of the stress system, or whether anxiety leads to a change in food intake, which in turn leads to a change in microbial composition. Future research investigating the link between maternal psychosocial stress and maternal microbiota should therefore try to also include dietary information.

Finally, Chapter 5 assumed that prenatal stress would affect maternal microbiota, which in turn would affect the infant microbiota. Infant microbiota would in turn influence infant development and health. This assumption was based on previous research showing that 1) infant intestines are virtually sterile at birth and the microorganisms important for the colonization of the intestine originate mainly from the mother in case of vaginal deliveries (de Weerth, Fuentes, & de Vos, 2013; Grönlund, Grzeskowiak, Isolauri, & Salminen, 2011; Jiménez et al., 2008; Jiménez et al., 2005), 2) maternal prenatal stress is related with differences in infant microbiota (Golubeva et al., 2015; Jašarević, Howard, & Bale, 2015; Zijlmans, Korpela, Riksen-Walraven, de Vos, & de Weerth, 2015), 3) the intestinal bacteria have a central position in human health and disease, and are suggested to also play a role in the regulation of emotions, behavior, and higher cognitive functions (Adlerberth & Wold, 2009; Cryan & Dinan, 2012; de Weerth, 2017; Diaz Heijtz, 2016; Luczynski et al., 2016; Saulnier et al., 2013; Sekirov et al., 2010) and 4) lower diversity and stability in the microbiome of the intestines have been shown to be related to infant colic (de Weerth et al., 2013). The current thesis investigated the first assumption, namely that prenatal stress is related to maternal microbiota. However, the whole gut microbial pathway, from prenatal maternal stress to infant health and development, has not been investigated. This is important as there are probably other mechanisms behind the link between prenatal stress and infant outcomes that are not explained by the association with microbiota. For example, prenatal psychosocial stress could increase cortisol concentrations in the mother, which crosses the placenta and increases cortisol.
concentrations in the fetus (Duthie & Reynolds, 2013). In turn, these heightened cortisol levels can affect the developing HPA axis of the fetus, resulting in increased basal cortisol concentrations and cortisol reactivity in the infant after birth (Tollenaar, Beijers, Jansen, Riksen-Walraven, & de Weerth, 2011). Cortisol, in turn, can change the permeability of the gut and affect the immune cells in the gut, affect gut motility and secretion, and produce increases in bile acid, all of which can potentially influence the infant intestinal microbiota (Zijlmans et al., 2015; Cryan & Dinan, 2012). Examining exactly how and to what degree prenatal psychosocial stress influences maternal microbiota and in turn infant microbiota and infant development might be an insightful endeavour for future research.

7.3.4 Practical Implications

The research in the current thesis consisted of fundamental research, primarily relevant for scientific insights. Nonetheless, in the long term it might have practical implications. For example, in Chapters 3 and 4, the simulator infant was used to assess prenatal quality of caregiving behavior and hormonal reactions, which were then found to predict postnatal quality of care. This is a promising finding that could open promising perspectives for future research and clinical applications in the long term. It has been suggested, for example, that an infant simulator could be used to test psychological and physiological correlates of parenting behavior, disentangle effects of maternal, paternal, and infant characteristics on parental quality and test the efficiency of specific interventions to boost parenting quality (Bakermans-Kranenburg, Alink, Biro, Voorthuis, & van Uzendoorn, 2015). That would indeed be a very valuable application, but more scientific work needs to be done first before this may be realized. For example, the studies in Chapter 3 and 4 of the present thesis should be replicated in additional samples with more diversity in demographic background and parental characteristics. Also, the feasibility of using an infant simulator with different groups of persons at risk should be tested, as well as the validity of the procedure when used with these groups of persons.

Additionally, Chapter 5 suggests that prenatal psychosocial stress is related to maternal microbiota. The microbiota is transferred from the mother to the child during vaginal delivery, and as stated above, may have effects on the health and development of the infant. Subsequently, future research could test whether 1) stress-reducing interventions have an effect on maternal microbiota, and 2) whether probiotics, i.e., “live microorganisms, which when consumed in adequate amounts, confer a health benefit on the host” (Pineiro & Stanton, 2007, pp. 8505-8515) - or prebiotics (food ingredients that induce activity or growth of beneficial microorganisms, Hutkins, et al., 2016) - which both have an effect on microbiota, reduce maternal psychosocial stress. Both interventions could potentially influence the infant's microbiota, and thus indirectly promote infant health and development.

REFERENCES


Appendices

Nederlandse samenvatting
Deutsche Zusammenfassung
Curriculum Vitae
Publications
Acknowledgements
**NEDERLANDSE SAMENVATTING**

De algemene doelstelling van dit proefschrift was een wetenschappelijke bijdrage te leveren aan het onderzoeksgebied van de vroege ontwikkeling. Het proefschrift had twee meer specifieke doelen. Het eerste doel was het verkrijgen van meer inzicht in de rol van mogelijke prenatale voorspellers van de kwaliteit van zorggedrag in de eerste maanden na de geboorte van de baby (Hoofdstuk 2, 3 en 4). Het tweede doel was het onderzoeken van mogelijke onderliggende mechanismen in de relatie tussen stress bij de moeder en ontwikkelingsuitkomsten bij het kind (Hoofdstuk 5 en 6). Op Hoofdstuk 2 na (waarin gebruik werd gemaakt van een steekproef van studenten) zijn de studies in dit proefschrift uitgevoerd in het kader van de longitudinale BINGO (Biologische Invloeden op baby’s Gezondheid en Ontwikkeling) studie, waarin moeders en vaders zijn gevolgd vanaf het derde trimester van de zwangerschap tot 12 weken na de geboorte van het kind.

Hoofdstuk 2, 3 en 4 betreffen het eerste doel: het onderzoeken van mogelijke voorspellers tijdens de zwangerschap van de kwaliteit van ouderlijk zorggedrag na de geboorte van het kind. Eerdere studies hebben aangetoond dat de kwaliteit van zorggedrag al vanaf de geboorte effect heeft op de ontwikkeling van kinderen. Het is daarom van belang om zo vroeg mogelijk, het liefst al vóór de geboorte, te kunnen vaststellen welke ouders mogelijk risico lopen om hun kind zorg van lage kwaliteit te bieden.

Hoofdstuk 2. De studie in Hoofdstuk 2 diende als preliminaire studie. De werkgeheugentaak die de studenten in deze studie uitvoerden werd later gebruikt voor de aanstaande ouders in de volgende studie (Hoofdstuk 3). In een steekproef van 120 universitaire studenten zijn cognitieve en emotionele reacties op babygehuil onderzocht. Het doel was om te achterhalen of jong volwassenen zonder kinderen negatieve emoties en kognitieve verstoringen laten zien als ze worden blootgesteld aan babygehuil dan wanneer ze andere geluiden horen. Deelnemers hebben een werkgeheugentaak uitgevoerd waarbij ze blootgesteld werden aan verschillende geluiden, waaronder babygehuil.

De resultaten lieten zien dat de deelnemers minder opgaven goed maakten naarmate zij meer negatieve emoties hadden. Dus de cognitieve verstoringen en de negatieve emoties zijn bij volwassenen mogelijk algemene reacties op babygehuil, wat suggereert dat er een algemene menselijke emotionele en cognitieve reactie op babygehuil bestaat.

**Hoofdstuk 3.** In deze studie is het gedrag van aanstaande moeders en vaders, terwijl ze werden blootgesteld aan gehuil van baby’s, tijdens de zwangerschap geobserveerd. Deze observaties werden gebruikt als voorspellers van zorggedrag voor de eigen baby zes weken naar de geboorte. Achtentachtig aanstaande moeders en 57 mannelijke partners werden getest in het derde trimester van de zwangerschap. Beide ouders werden 15 minuten gefilmd tijdens een individuele interactie met een huilende simulator-baby. Het gedrag van de ouders werd later aan de hand van de films geobserveerd, waarbij de kwaliteit van de zorggedrag werd beoordeeld op sensitiviteit en coöperatie. Daarnaast werden kognitieve verstoringen en cognitieve variaties met elkaar samenhangen getest. In de BINGO studie werd bij 88 aanstaande moeders en 57 mannelijke partners in het derde trimester van de zwangerschap cortisol en testosteron in speeksel gemeten voor, tijdens, en na hun interactie met een simulator-baby. Het gedrag van de aanstaande ouders was, zoals beschreven in Hoofdstuk 3, ook beoordeeld op kwaliteit van zorggedrag. Toen de eigen baby 6 weken oud was zijn de ouders thuis beoordeeld en gefilmd tijdens een individuele interactie met hun eigen baby. De kwaliteit van zorggedrag werd bepaald door het gedrag van de ouders te beoordelen op sensitiviteit en coöperatie. Cortisol en testosteron in speeksel werden ook gemeten tijdens deze interactie.
De basale levels en de reactiviteit van cortisol en testosteron werden gebruikt als prenatale voorspellers van postnataal zorggedrag, waarbij werd gecontroleerd voor prenataal zorggedrag en voor postnatale cortisol en testosteron concentraties. De resultaten lieten zien dat het postnatale zorggedrag van de moeders niet samenhangt met hun prenatale cortisol en testosteron concentraties. Bij de vaders echter hing een lagere kwaliteit postnataal zorggedrag samen met hogere prenatale cortisol concentraties. Daarnaast was het zorggedrag van de vader in zowel de prenatale als de postnatale periode gerelateerd aan interacities tussen cortisol en testosteron. Deze bevindingen zijn geïnterpreteerd in de context van hypereffectieve fysiologische responses die geobserveerd zijn bij ouders met risico op lagere kwaliteit van zorggedrag. Verder onderzoek is nodig om deze resultaten te verduidelijken; bijvoorbeeld om te kunnen begrijpen waarom cortisol en zorggedrag wel bij vaders, maar niet bij moeders met elkaar samenhangen over de perinatale periode.

Hoofdstuk 5 en 6 betreffen het tweede doel van dit proefschrift, namelijk het onderzoeken van mogelijke onderliggende mechanismen in de relatie tussen stress bij de moeder en ontwikkelingsuitkomsten bij het kind. Recent onderzoek laat zien dat zowel psychosociale stress tijdens de zwangerschap als postnatale stress bij de moeder vlak na de geboorte de gezondheid en sociaal-emotionele ontwikkeling van het kind kunnen beïnvloeden. De mechanismen die ten grondslag liggen aan deze relatie zijn echter nog niet helemaal bekend. In dit proefschrift zijn twee biologische mechanismen onderzocht die in het verleden weinig aandacht hebben gehad.

**Hoofdstuk 5.** Het eerste mechanisme houdt in dat prenatale psychosociale stress bij de moeder de gezondheid en ontwikkeling van het kind beïnvloedt via de darmbacteriën van de moeder. De darmbacteriën van de moeder zijn bij vaginale bevallingen de belangrijkste bron voor de kolonisatie van de darmen en de ontwikkeling van de darmbacteriën van het kind, die op hun beurt weer belangrijk zijn voor het zich ontwikkelende darmstelsel, het immuunsysteem, en mogelijk ook de hersenontwikkeling van het kind. In de studie in Hoofdstuk 5 is de eerste stap in dit potentiële mechanisme onderzocht, namelijk de relatie tussen psychosociale stress en darmbacteriën bij zwangere vrouwen. Zeventig aanstaande moeders hebben tijdens het derde trimester van de zwangerschap een monster van hun ontlasting genomen, waarin vervolgens de darmbacteriën werden onderzocht. De aanstaande moeders werden vervolgens ook gevraagd om vragenlijsten over algemene en zwangerschaps-specifieke stress en angst in te vullen, om psychosociale stress te kunnen meten. Op fylum niveau werd verwacht dat de darmbacteriën van moeders met hoge psychosociale stress vooral gekenmerkt zouden worden door meer Proteobacteria en minder Actinobacteria, en op genus niveau door meer Lactobacilli, vergeleken met moeders met lage psychosociale stress.


In deze studie is de eerste stap onderzocht in een mechanisme dat mogelijk ten grondslag ligt aan de relatie tussen prenatale psychosociale stress bij moeders en ontwikkelingsuitkomsten bij hun kind, namelijk de relatie tussen psychosociale stress en darmbacteriën bij aanstaande moeders. Tegen de verwachting in werden de darmbacteriën van moeders met hoge psychosociale stress niet gekenmerkt door meer *Proteobacteria* en minder *Actinobacteria* op fylum niveau, en minder Lactobacilli op genus niveau, vergeleken met moeders met lage psychosociale stress. Er werd echter wel een significante samenhang gevonden tussen algemene angst en darmbacteriën op genus niveau. Deze studie biedt inzichten in het verband tussen mentale gezondheid en darmbacteriën van moeders tijdens de zwangerschap en is een startpunt voor verder onderzoek, waarin ook het eetgedrag van de moeder en de darmbacteriën en de ontwikkeling van de baby dienen te worden onderzocht.

**Hoofdstuk 6.** In deze studie is een ander biologisch mechanisme onderzocht dat de relatie tussen stress bij de moeder en ontwikkelingsuitkomsten bij het kind zou kunnen verklaren; dit mechanisme houdt in dat stress bij de moeder invloed heeft op het kind via de samenstelling van de moedermelk. In deze studie is een deel van dit mechanisme onderzocht, namelijk de relatie tussen cortisol in moedermelk en de relatieve hoeveelheid bacteriën op genus niveau. Deze studie biedt inzichten in het verband tussen mentale gezondheid en darmbacteriën van moeders tijdens de zwangerschap en is een startpunt voor verder onderzoek, waarin ook het eetgedrag van de moeder en de darmbacteriën en de ontwikkeling van de baby dienen te worden onderzocht.

**Hoofdstuk 6.** In deze studie is een ander biologisch mechanisme onderzocht dat de relatie tussen stress bij de moeder en ontwikkelingsuitkomsten bij het kind zou kunnen verklaren; dit mechanisme houdt in dat stress bij de moeder invloed heeft op het kind via de samenstelling van de moedermelk. In deze studie is een deel van dit mechanisme onderzocht, namelijk de relatie tussen cortisol in moedermelk en de relatieve hoeveelheid bacteriën op genus niveau. Deze studie biedt inzichten in het verband tussen mentale gezondheid en darmbacteriën van moeders tijdens de zwangerschap en is een startpunt voor verder onderzoek, waarin ook het eetgedrag van de moeder en de darmbacteriën en de ontwikkeling van de baby dienen te worden onderzocht.
onderzoek, het best met een groter steekproef, is van belang om de resultaten van deze studie te repliceren en meer kennis te verkrijgen.

Conclusies
Bij elkaar genomen kunnen de resultaten van de vijf studies in dit proefschrift samengevat worden in de volgende hoofdconclusies:

- Wanneer jongvolwassenen zonder kinderen tijdens een werkgeheugentaak worden blootgesteld aan babygehuil, verminderen hun prestaties en ervaren zij meer negatieve emoties.
- De kwaliteit van zorggedrag voor de eigen baby kan bij zowel moeders als vaders worden voorspeld uit observaties van de kwaliteit van hun zorggedrag voor een simulator-baby tijdens de zwangerschap, en voor vaders ook door cortisol concentraties tijdens de interactie met deze simulator-baby.
- Er is geen bewijs gevonden voor een verband tussen psychosociale stress en darmbacteriën op fylum niveau bij zwangere vrouwen. Wel is er bij de aanstaande moeders een verband gevonden tussen algemene angst en darmbacteriën op genus niveau.
- Er is geen bewijs gevonden voor een samenhang tussen cortisol in moedermelk en het huilen van baby’s in de eerste drie maanden na de geboorte.

DEUTSCHE ZUSAMMENFASSUNG


Die Ergebnisse zeigten, dass die Teilnehmer mehr Fehler machten und mehr negative Emotionen zeigten, wenn Säuglingsweinen gehört wurde, als wenn andere störende Geräusche gehört wurden. Persönlichkeitsmerkmale standen in keinem Zusammenhang mit der kognitiven Leistung und den emotionalen Reaktionen der Teilnehmer auf das Weinen von Säuglingen. In allen Durchläufen waren die negativen
Emotionen mit mehr Fehlern bei dem Gedächtnistest verbunden, außer wenn Säuglingsweinen gehört wurde. Daher könnten kognitive Störungen und negative Emotionen allgemeine Reaktionen von Erwachsenen auf das Weinen von Säuglingen sein, was auf eine allgemein vorhandene kognitive und emotionale Reaktion des Menschen auf das Weinen von Säuglingen schließen lässt.


Die Ergebnisse zeigten, dass sowohl bei Müttern als auch bei Vätern die pränatale Qualität der Fürsorge gegenüber dem Simulator-Säugling signifikant und positiv die postnatale Qualität der Fürsorge gegenüber dem eigenen Säugling vorhersagte. Kognitive Störungen und die Fähigkeit körperliche Kräfte zu regulieren, beim Hören von Säuglingsweinen, waren nicht mit der postnatalen Qualität der Fürsorge verbunden. Daher können pränatale Beobachtungen der Fürsorge für einen Simulator-Säugling die postnatale Qualität der Fürsorge vorhersagen. Für die Umsetzung der Ergebnisse in die Praxis sind jedoch weitere Forschungsarbeiten erforderlich, beispielsweise für das Screening gefährdeter Eltern bei denen ein Risiko für minderwertige Fürsorge besteht, sowie für Lehr- und Ausbildungszwecke.


Siebzig werdende Mütter haben während des dritten Trimesters der Schwangerschaft eine Stuhlprobe abgegeben, worin die Darmbakterien untersucht wurden. Um psychosozialen Stress zu messen, wurden die werdenden Mütter gebeten, Fragelisten zu allgemeinen und schwangerschaftsspezifischen Stress- und Angstzuständen auszufüllen. Es wurde die Hypothese aufgestellt, dass Mütter mit hohem psychosozialen Stress, verglichen mit Müttern mit niedrigem psychosozialen Stress, mikrobielle Zusammensetzungen aufweisen, die, auf Phylum-Ebene, durch mehr Proteobakterien und weniger Actinobakterien, und, auf Genus-Ebene, durch mehr Laktobazillen, gekennzeichnet sind.


Fazit
Zusammengehalten können die Ergebnisse der fünf in dieser Doktorarbeit vorgestellten Studien in den folgenden Hauptschlussfolgerungen zusammengefasst werden:

- Das Hören von Weinen eines Säulings während der Ausführung eines Gedächtnistests hängt zusammen mit einer verminderten kognitiven Leistung und erhöhten negativen Emotionen bei jungen kinderlosen Erwachsenen.
- Die postnatale Qualität der elterlichen Fürsorge kann, bei sowohl Müttern als Vätern, aus pränatalen Beobachtungen der Fürsorge für einen weinenden Simulator-Säugling vorhergesagt werden, und bei Vätern auch aus den Cortisol Konzentrationen während der Interaktion mit diesem weinenden Simulator-Säugling.
- Es wurde kein Hinweis auf einen Zusammenhang zwischen Cortisol aus der Muttermilch und dem Weinen des Kindes während der ersten drei Monate nach der Geburt gefunden.
CURRICULUM VITAE

Christine Hechler was born on May 14th, 1986 in Mainz, Germany. After completing the Abitur, Christine took a time out before moving to the Netherlands and starting the Bachelor programme Psychology at Radboud University Nijmegen in 2008. She received her Bachelor’s degree in 2011 and was admitted to the Research Master Behavioral Science at Radboud University that same year. In 2013, Christine completed her Master’s degree with a thesis on individual differences in reactions to infant crying. Immediately after this, she started her PhD project at the department of Developmental Psychology of Radboud University, under the supervision of prof. dr. Marianne Riksen-Walraven, prof. dr. Carolina de Weerth, and Dr. Roseriet Beijers. This project resulted in the current thesis. After completing the PhD project, Christine worked as a teacher at the department of Developmental Psychology of Radboud University.

PUBLICATIONS


ACKNOWLEDGEMENTS

I would like to thank everyone who, in one way or another, helped in the process of accomplishing this thesis.

First, thanks to the many families that participated in the BINGO project. Thank you all so much for coming to our lab and letting me and my students visit your home. Filling out questionnaires, doing our tasks, collecting diverse samples - all of this while expecting a child and having a newborn around - is hard work, and I am very grateful for your willingness to participate in the BINGO project!

Of course, none of this would have happened if I had not been giving the opportunity to even do a PhD project. Marianne, thank you so much for accepting this project, even though you were about to retire, and retired soon after the start of this project. Your insights have made the project and this thesis what it is today, and your feedback and suggestions were always spot on. Thank you for joining us for discussions throughout the years, your kindness and calm manners always helped me to stay focused and keep going.

In contrast, thank you Carolina for (not literally) kicking my butt! The occasional NSFW humor aside, you always pushed me to stay focused, work hard and keep going. Since I managed to survive the “Carolina curse” by not getting pregnant during the time we worked together, I leave you the BINGO child instead. You already have plans for its future, and I wish you all the luck (and money...) for its further development.

Roseriet, as my daily supervisor you were always the first to go to and the first to help me. Thank you so much for all the input and work you have done! I admire your quick thinking, your critical attitude and your statistical knowledge. You not only helped a lot with all the chapters from this thesis, you also always reminded me to have fun and take breaks (even though this was not really possible at times). Thank you for keeping me focused on the important things.

Keeping BINGO up and running also required help from students, and I would like to thank them all. First, thanks to all the clinical master students who collected samples and helped me during the visits as the second tester. Maudy, Kathrin, Thibaut, Michelle, Leonie, Jill, Garlef, thank you for helping me!

Second, I would like to thank a few research master students. Kelly and Marije, thank you for stepping in as a tester, even though you were not involved in the BINGO project. Thank you, Nancy, for choosing BINGO for your internship. Without you, the organisation of the project would have been so much harder! Thank you also for scoring all these videos, for exchanging ideas, and for finding the first participants. Emma, you were the best thing that could have happened to BINGO. You are smart, hard-working, quick-thinking, and determined, and you deserve to be in the academic world more than I do. Thank you for all the hard work you put into BINGO; lab visits on the weekends, home visits till late at night, taking over BINGO while I was on holiday, supervising new testers, discussing ideas, finding solutions to problems, etc. But most of all I would like to thank you for your company and friendship. The late nights and 70-hour work weeks were much more bearable due to your company, your stories and your jokes (even though Germans don’t have humor). Thank you for tolerating my (occasional?) road rage and for letting me enjoy your pregnancy alongside you. I admire your strength for keeping up with all that has happened, none of which you deserved, and I wish you all the best for the future!

Finally, I would like to thank all the people who joined BINGO voluntarily. Simone, thank you for taking over while I was on holiday. Germaine, thank you so much for testing and typing in all of those cry diaries! Eva, Sina, and Marleen, thank you for being my second testers. Claire, thank you for helping me to find new participants, for being so positive and kind, and for making me laugh countless times.

Not only colleagues but also my paranymphs, I would like to thank Kelly and Yvonne. Kelly, we shared a similar faith, and I enjoyed talking to you about all the stress, the fun times, and also the weird things that can happen during testing. Yvonne, you have taken over the BINGO project, and you have already started testing. I hope you enjoy the work with the families as much as I did.

I would like to thank you both for being my paranymphs and helping me with finishing this thesis and organizing the defence.

I would also like to thank all the people on the 8th floor. Meta, you kept making me laugh during lunch so many times, and I needed that. Thank you for always asking me if I am okay, and for making me feel that not only my work here is valued, but that I am a person too. I would also like to thank you, Michelle, my roommate during the first three years. Thank you for having “an open ear” for all of my problems, both professionally and personally, for discussing problems and finding solutions, and for putting everything in perspective when I was lost. I enjoyed our runs during the Sevenheuvelenacht, and the pizza at your place afterwards. Michelle was replaced by Sven, and I would like to thank him for almost always keep his pants on (Sven, you know what I mean; all others: it’s a joke. Hard to believe for a German, I know).
A very special thanks to **Ronny**, the BSI lab manager and engine of the BSI. Ronny, I am thankful for all the times that I could come to you, whether because I needed something for my project or whether I just wanted to enjoy a good talk and relax. You are the best lab manager the BSI could hope for, and I appreciate all the work you do! Thank you for trusting me with the keys and for helping me managing the lab visits. Also, thank you for making me laugh and reminding me that the world will not end if something does not go as planned.

Thanks to the **TSG staff** for all the (mostly short-notice) help you provided. Special thanks to **Gerard**, who helped with so many things, you are a genius! Thank you also for providing good company. Thank you, **Mark**, for building the handgrip that I used. **Chi**, you came to my rescue before so many lab visits. And thank you, **Philip**, for fixing my Matlab code.

For all the statistics, I relied on two men: **Bill**, thank you for all the help with the analyses, and **Pierre**, thank you for providing us with the codes for the cry diary data. I am grateful that I could always ask for your help.

To all the men and women at the front desk of the Spinoza building: thank you! Thank you for literally opening the door for me; the front door when I came back from visits after hours, the door to your little storage room when I was too late to head to my office, and the BSI lab door when I tested after hours. You always greeted me with a big smile and kind words, and I really needed that!

Last but not least, I would like to thank my husband **Peter**. No academic title in the world can compare to the love and happiness I receive from you! Thank you for keeping up with all the negative side-effects of having such a smart-ass career woman on your side, and I apologize for being such a dragon during the testing phase. Thank you for being so understanding when I cancelled yet another plan, when I neglected our social life, and when I had to miss your promotion due to one of the many visits. Thank you for being at my side and getting me through all of this! As promised, I will not thank you for cooking and cleaning. Ich liebe Dich Babe forever. For ever? For ever ever? For ever ever!