Prenatal and postnatal cortisol and testosterone are related to parental caregiving quality in fathers, but not in mothers

Peter A. Bos[a],[*], Christine Hechler[b], Roseriet Beijers[b], Kazuyuki Shinohara[c], Gianluca Esposito[d],[e], Carolina de Weerth[b]

[a] *Experimental Psychology, Utrecht University, The Netherlands*
[b] *Radboud University, The Netherlands*
[c] *Graduate School of Medicine, Nagasaki University, Japan*
[d] *University of Trento, Italy*
[e] *Nanyang Technological University, Singapore*

ARTICLE INFO

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 post-natally 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.

1. Introduction

The quality of parental caregiving is a critical factor in a child's cognitive and social-emotional development, with insensitive car-egiving practices increasing the risk for developing various types of psychopathology (Gilbert et al., 2009; Keyes et al., 2012; Morris et al., 2013; Pechtel and Pizzagalli, 2011). It is therefore of great importance to understand the underlying factors that bring forth variation in par-ents' 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 fa-thers. At the same time, human fathers contribute significantly to par-ental care across cultures and also have a strong impact on the out-comes of their child's development and wellbeing (Cabrera et al., 2018; Gray and 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 in-vestigated mothers in the postnatal period, and focused on CORT (e.g. Finegood et al., 2016; Gonzalez et al., 2012; Mills-Koonce et al., 2009). Early postnatal levels of CORT in human mothers have shown positive relations with affectionate infant-directed behavior (Fleming et al., 1987), responsiveness to and attractiveness of baby odor (Fleming et al., 1997), and sympathy towards infant crying (Stallings et al., 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 and 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 et al., 2011). Furthermore, lower paternal T levels were related to more self-reported parental investment (Mascaro et al., 2013), more responsiveness to infant cues (Storey et al., 2000), and stronger sympathy in response to infant crying (Fleming et al., 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 et al., 2017) and with less interactive behavior when fathers interacted with their 3-to-8-month-old infants (Weisman et al., 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 et al., 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 et al., 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 et al., 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 and Hagstrom, 1996). In nulliparous adults, physiological endocrine responses to crying have also been related to intended harsh caregiving (Out et al., 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 asses prenatal responses to infant crying and parental caregiving behavior (e.g. Rutherford et al., 2015, 2017; van Anders et al., 2012; Voorthuis et al., 2013).

Also, most of the 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 (Mehta and 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 and 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 et al., 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, 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 and Hagstrom, 1996), we anticipate negative relations between endocrine responses and quality of caregiving.

2. Methods

2.1. 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 ethical committee of the Faculty of Social Sciences of the Radboud University [ECSW2014-1003-189]. Families were recruited during pregnancy by distributing flyers in midwife practices and pregnancy courses in the region Arnhem-Nijmegen (the Netherlands). Fathers were encouraged to participate, but mothers could also participate in the study without their partner.

Initial prenatal exclusion criteria were: drug use, excessive alcohol use, insufficient mastery of the Dutch language, and an unhealthy pregnancy so far. In total, 88 expectant mothers and 57 of their partners qualified for participation 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 unknown reasons (n = 2). The majority of participants were born in the Netherlands (83 mothers, 51 fathers) and were employed (83 mothers, 55 fathers). Postnatal exclusion criteria were: complications during pregnancy (after initial contact), prematurity (gestational age ≤ 37 weeks), birth weight < 2500 g, 5-minute Apgar score < 7, and child anomalies. Two infants were born in week 36 of the pregnancy. As these infants were completely healthy, the families were included as well. Two families were excluded from the analyses due to premature birth of the child (gestational week 35, n = 1), and brain damage detected at birth (n = 1). The final sample thus consisted of 86 mothers and 56 of their partners. Seven families stopped participation after birth due to personal reasons. Infants (41 boys, 38 girls) were born at term ($M_{\text{gestational age}}$ = 39.78, SD = 1.53), with an average birth weight of 3531.07 g (SD = 428.43).

2.2. Procedure

Participants visited the lab during the third trimester of pregnancy (M = 33.93 weeks, SD = 2.24 weeks). All lab visits took place during late afternoon (after 15:00) or in the early evening (M = 17:28, SD = 01:53). When both the mother and father participated, they were

tested separately. A dice was rolled to decide whether the mother or the father was tested first. Participants first filled in some questionnaires and then performed a working memory task and a handgrip dynamometer task, both unrelated to the current study. Between the working memory and handgrip dynamometer task, participants provided a saliva sample (approximately 2 ml) by means of passive drooling (T1). This sample served as baseline measurement. Subsequently, participants interacted with an unsoothable crying simulator infant for 15 min while being video recorded. The simulator infant (RealCare Baby; Realityworks, Eau Claire, WI, USA) was used to elicit prenatal caregiving behavior towards a crying infant. The simulator infant resembles a real infant aged 0–3 months in weight, size, and appearance as well as in expressed needs. Similar to a real infant, the simulator starts fussing to express a need which eventually turns into crying if the need is not met.

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 -unbeknownst 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 min 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 min (T2) and 35 min (T3) after the end of the interaction with the simulator infant, participants provided saliva samples again. Saliva samples were immediately stored at -20℃.

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 et al., 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 min long (in cases when the parent finished before the 15 min he/she was asked to continue interacting with the infant until 15 min were completed). Changing an infant at this age constitutes a mild physical stressor that may elicit crying and fussing (Jansen et al., 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 min before the start of the interaction (T1; between the working memory and handgrip dynamometer task), at 15 min (T2), and 35 min (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℃. During both visits, the parents were blind to the exact nature of the tasks and the goals of the present study.

2.3. Measures

2.3.1. Prenatal quality of caregiving

The videos were rated for parental sensitivity and cooperation using 9-point rating scales (Ainsworth et al., 2015), 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.

2.3.2. Postnatal quality of caregiving

The videos were rated for sensitivity and cooperation using the same 9-point scales (Ainsworth et al., 2015) 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.

2.3.3. 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 in house competitive radio-immunoassay employing a polyclonal anticortisol-antibody (K7348). [1,2-^{3}H(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% ($n = 10$).

2.3.4. 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 $1500 \times g$ for 15 min, 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.

2.3.5. 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 and Maestripieri, 2011; Storey and 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 difficultness and seriousness in interacting with the simulator infant. Control variables that did not improve the model significantly, were removed from the analyses.

2.4. 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 (Tabachnick and 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 Tabachnick 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.

3. Results

3.1. Preliminary analyses

3.1.1. 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 (Tabachnick and Fidell, 2007), missing values were not imputed.

3.1.2. 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; $M = 5.60$, $SD = 1.13$).

3.1.3. 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 ($t = 3.10$, $p = .002$).

3.1.4. Prenatal and postnatal T and CORT responses to the interactions

Fig. 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

Table 1
Descriptive Statistics, Separately for Mothers and Fathers.

	Mothers M (SD)	Fathers M (SD)
Control Variables		
Educational Level		
University	n = 41 (48%)	n = 27 (48%)
College	n = 30 (35%)	n = 16 (29%)
High school	n = 1 (1%)	n = 3 (5%)
Job training	n = 13 (16%)	n = 10 (18%)
Parity		
First-child	73 (85%)	44 (79%)
Second-child	13 (15%)	12 (21%)
Age	31.62 (3.79)	32.93 (4.16)
Difficulty[a]	4.35 (1.66)	4.60 (1.86)
Seriousness[b]	5.70 (1.11)	5.44 (1.15)
Quality of caregiving		
Prenatal	3.84 (1.87)	3.39 (1.56)
Postnatal[*]	5.45 (1.74)	4.41 (1.76)

* Difference between mothers and fathers is p = .026, as indicated by paired sample *t*-test.

[a] How difficult expectant parents found it to interact with the simulator infant as if it were real.

[b] How seriously expectant parents performed the interaction with the simulator infant.
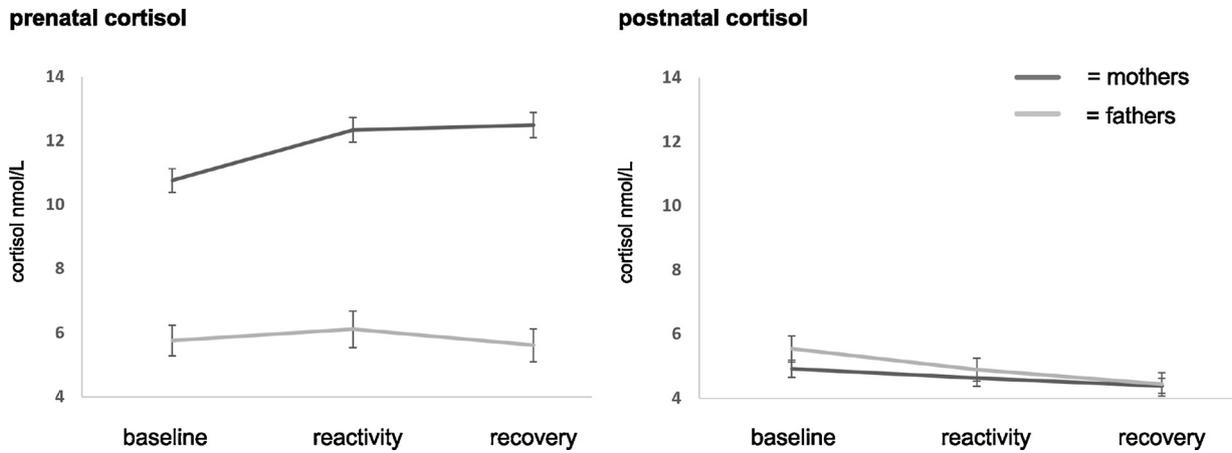
prenatal cortisol

postnatal cortisol

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

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.

3.2. Main multilevel analyses

3.2.1. Is prenatal quality of caregiving behavior associated with prenatal T and CORT baseline and reactivity?

3.2.1.1. *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.

3.2.1.2. *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 Fig. 3). The figure shows that for the low CORT group, T is more negatively associated with quality of caregiving than for the

prenatal testosterone
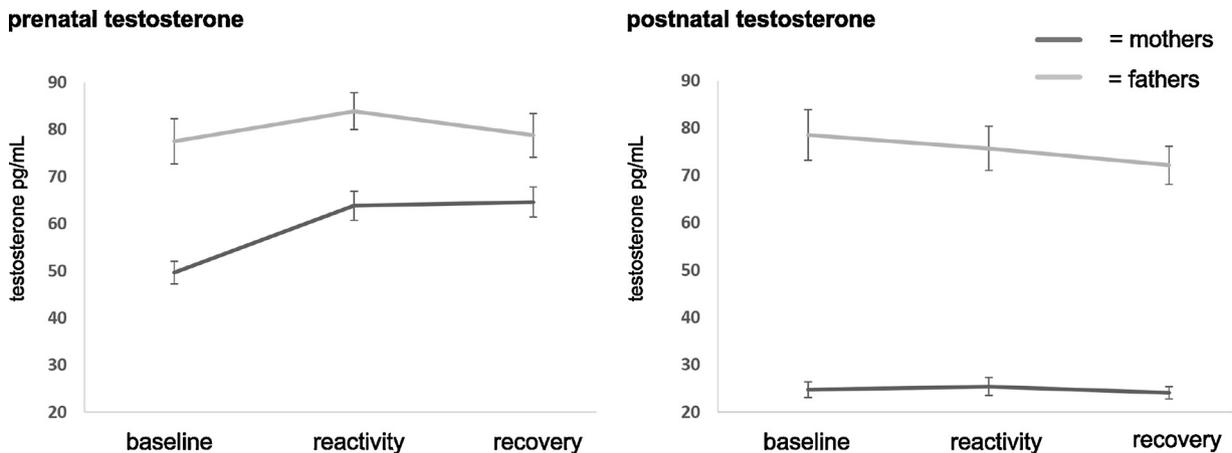

postnatal testosterone

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

Table 2
Multilevel Time- only Models.

		Mothers			Fathers		
		Estimate	SE	p	Estimate	SE	p
Prenatal	Cortisol						
	Intercept	17.94	2.64	< .001	14.95	3.57	< .001
	Time	−0.01	0.00	.027	−0.01	0.00	.012
	Time Quadratic	–	–	–	–	–	–
	Deviance		1258.398			694.762	
	Testosterone						
	Intercept	4.58	22.10	.836	656.05	250.53	.010
	Time	0.05	0.02	.014	−1.01	0.45	.026
	Time Quadratic	–	–	–	0.00	0.00	.029
	Deviance		1851.911			1007.304	
Postnatal	Cortisol						
	Intercept	35.54	8.97	< .001	18.32	2.45	< .001
	Time	−0.05	0.02	.003	−0.01	0.00	< .001
	Time Quadratic	−0.00	0.00	.009	–	–	–
	Deviance		846.190			626.141	
	Testosterone						
	Intercept	26.90	10.63	.012	− 353.85	152.89	.023
	Time	−0.00	0.01	.854	0.86	0.28	.003
	Time Quadratic	–	–	–	−0.00	0.00	.002
	Deviance		1352.739			1091.728	

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.

3.2.2. Is postnatal quality of caregiving associated with postnatal T and CORT baseline and reactivity?
3.2.2.1. 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

Table 3
Maternal Multilevel Models.

	Prenatal				Postnatal				Pre- to Postnatal			
	Estimate	SE	p	VIF	Estimate	SE	p	VIF	Estimate	SE	p	VIF
Cortisol												
Intercept	15.60	2.82	< .001	–	31.91	9.71	**.003**	–	15.21	3.12	< .001	–
Time	−0.01	0.00	**.007**	–	−0.05	0.02	**.009**	–	−0.01	0.00	**.012**	–
Time Quadratic	–	–	–	–	−0.00	0.00	**.021**	–	–	–	–	–
Parity	−1.18	0.76	.124	1.04	–	–	–	–	−0.94	0.81	.250	1.03
Educational Level	−0.36	0.19	.063	1.05	−0.06	0.13	.657	1.04	−0.36	0.21	.085	1.05
Testosterone Prenatal	0.08	0.01	< .001	1.02	–	–	–	–	0.08	0.01	< .001	1.01
Testosterone Postnatal	–	–	–	–	0.07	0.01	< .001	1.01	–	–	–	–
Quality of caregiving Prenatal	0.16	0.18	.377	1.16	–	–	–	–	0.29	0.19	.130	1.09
Quality of caregiving Postnatal	–	–	–	–	0.19	0.12	.120	1.03	−0.01	0.21	.973	1.10
Interaction Parity*Sensitivity	0.70	0.45	.125	1.19								
Deviance		921.293[a]				635.199[c]				850.118[e]		
Testosterone												
Intercept	−68.46	22.68	**.003**	–	−10.16	11.92	.397	–	−69.66	24.82	**.005**	–
Time	0.07	0.02	< .001	–	0.02	0.01	.077	–	0.07	0.02	< .001	–
Parity	–	–	–	–	3.20	3.11	.308	1.03	–	–	–	–
Educational Level	2.37	1.57	.130	1.03	0.80	0.78	.312	1.03	2.42	1.69	.151	1.05
Cortisol Prenatal	4.14	0.39	< .001	1.03	–	–	–	–	4.06	0.41	< .001	1.04
Cortisol Postnatal	–	–	–	–	2.63	0.44	< .001	1.04	–	–	–	–
Quality of caregiving Prenatal	−1.15	1.37	.402	1.05	–	–	–	–	−1.26	1.52	.410	1.11
Quality of caregiving Postnatal	–	–	–	–	−0.36	0.74	.630	1.08	−0.14	1.63	.933	1.08
Deviance		1698.162[b]				1252.429[d]				1552.627[f]		

[a] Durbin-Watson = 1.31.
[b] Durbin-Watson = 1.10.
[c] Durbin-Watson = 1.01.
[d] Durbin-Watson = 1.44.
[e] Durbin-Watson = 1.31.
[f] Durbin-Watson = 1.15.

Table 4
Paternal Multilevel Models.

	Prenatal				Postnatal				Pre- to Postnatal			
	Estimate	SE	p	VIF	Estimate	SE	p	VIF	Estimate	SE	p	VIF
Cortisol												
Intercept	5.55	4.41	.212	–	5.25	4.09	.204	–	4.29	4.67	.361	–
Time	−0.00	0.00	.215	–	−0.01	0.00	**.041**	–	−0.00	0.00	.712	–
Age	–	–	–	–	0.10	0.08	.233	1.01				
Educational level	–	–	–	–	–	–	–	–	–	–	–	–
Testosterone prenatal	0.52	0.01	**< .001**	1.01	–	–	–	–	0.05	0.01	**< .001**	1.01
Testosterone postnatal	–	–	–	–	0.05	0.01	**< .001**	1.04	–	–	–	–
Quality of caregiving prenatal	0.21	0.31	.508	1.01	–	–	–	–	0.39	0.32	.232	1.06
Quality of caregiving postnatal	–	–	–	–	−0.25	0.18	.175	1.04	−0.58	0.29	**.050**	1.05
Interaction quality of caregiving postnatal*testosterone postnatal	–	–	–	–	−0.01	0.00	**< .001**	1.03	–	–	**–**	–
Deviance		516.010[a]				445.162[c]				476.920[e]		
Testosterone												
Intercept	650.00	227.75	**.005**	–	−67.74	257.48	.793	–	613.63	257.64	**.019**	–
Time	−0.95	0.40	**.021**	–	0.23	0.46	.619	–	−0.90	0.46	.053	–
Time quadratic	0.00	0.00	**.023**	–	−0.00	0.00	.554	–	0.00	0.00	.057	–
Cortisol prenatal	3.10	0.63	**< .001**	1.02	–	–	–	–	2.95	0.68	**< .001**	1.02
Cortisol postnatal	–	–	–	–	6.01	0.88	**< .001**	1.11	–	–	–	–
Seriousness	−8.32	3.48	**.022**	1.20	–	–	–	–	−8.89	3.50	**.016**	1.20
Quality of caregiving prenatal	−1.78	2.71	.515	1.30	–	–	–	–	−1.00	2.69	.711	1.25
Quality of caregiving postnatal	–	–	–	–	2.66	2.12	.217	1.11	1.77	2.12	.410	1.07
Interaction quality of caregiving prenatal*cortisol prenatal	−1.09	0.51	**.035**	1.11	–	–	–	–	–	–	–	–
Deviance		949.706[b]				1013.851[d]				880.943[f]		

[a] Durbin-Watson = 0.89.
[b] Durbin-Watson = 1.42.
[c] Durbin-Watson = 1.11.
[d] Durbin-Watson = 1.18.
[e] Durbin-Watson = 0.93.
[f] Durbin-Watson = 1.33.
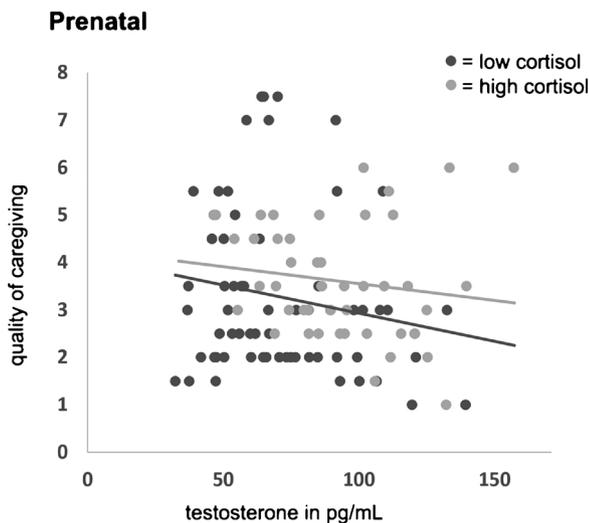
Fig. 3. Interaction between cortisol and testosterone on quality of prenatal caregiving for fathers, with regression lines for the high and low cortisol group.

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.

3.2.2.2. 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 445.162 (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 Fig. 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.

3.2.3. Is postnatal quality of caregiving related to prenatal T and CORT?
3.2.3.1. Mothers. To investigate whether postnatal quality of caregiving is related to prenatal T and CORT, controlling for
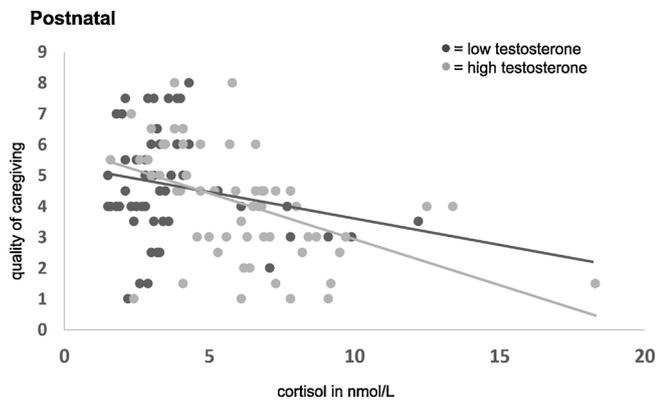
Postnatal

Fig. 4. Interaction between testosterone and cortisol on quality of postnatal caregiving for fathers with regression lines for the high and low testosterone group.

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.

3.2.3.2. 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.

4. 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 and 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 sensitivity 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 quality of postnatal 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 and Panneton, 2017), and abusive parents generally show hyperreactive stress-responses to infant distress (McCanne and 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 sensitivity in fathers was only observed in those less stressed by the infant simulator. This finding seems in line with the dual-hormone hypothesis (Mehta and 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 et al., 2014). Furthermore, this finding for the postnatal session is in line with the idea that hyperreactive physiological responses are negatively related to parental sensitivity (McCanne and 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), fathers' 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 et al., 2003), facilitate positive social interaction and caregiving quality in fathers interacting with their own children (Naber et al., 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 level, as well as decreased quality of parental caregiving (Bailey et al., 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 et al., 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, 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 et al., 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 post-partum 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 et al., 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 et al., 2014, 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 et al., 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, a 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.

Conflict of interest statement

Acknowledgements and role of the funding source

References

Ainsworth, M.D.S., Blehar, M.C., Waters, E., Wall, S.N., 2015. Patterns of Attachment: A Psychological Study of the Strange Situation. Taylor & Francis.

Bailey, J., Hill, K., Oesterle, S., Hawkins, J., 2009. Parenting practices and problem behavior across three generations: monitoring, harsh discipline, and drug use in the intergenerational transmission of externalizing behavior. Dev. Psychol. 45, 1214–1226.

Barr, R.G., Trent, R.B., Cross, J., 2006. Age-related incidence curve of hospitalized shaken baby syndrome cases: convergent evidence for crying as a trigger to shaking. Child Abuse Negl. 30, 7–16.

Bos, P.A., 2017. The endocrinology of human caregiving and its intergenerational transmission. Dev. Psychopathol. 29, 971–999.

Buckels, E.E., Beall, A.T., Hofer, M.K., Lin, E.Y., Zhou, Z., Schaller, M., 2015. Individual differences in activation of the parental care motivational system: assessment, prediction, and implications. J. Pers. Soc. Psychol. 108, 497.

Cabrera, N.J., Volling, B.L., Barr, R., 2018. Fathers are parents, too! Widening the lens on

parenting for children's development. Child Dev. Perspect. https://doi.org/10.1111/cdep.12275.

Deater-Deckard, K., Panneton, R., 2017. Parental Stress and Early Child Development. Springer International Publishing.

Eisenegger, C., Haushofer, J., Fehr, E., 2011. The role of testosterone in social interaction. Trends Cogn. Sci. (Regul. Ed.) 15, 263–271.

Endendijk, J.J., Hallers-Haalboom, E.T., Groeneveld, M.G., van Berkel, S.R., van der Pol, L.D., Bakermans-Kranenburg, M.J., Mesman, J., 2016. Diurnal testosterone variability is differentially associated with parenting quality in mothers and fathers. Horm. Behav. 80, 68–75.

Feldman, R., 2015. Sensitive periods in human social development: new insights from research on oxytocin, synchrony, and high-risk parenting. Dev. Psychopathol. 27, 369–395.

Feldman, R., Eidelman, A.I., Rotenberg, N., 2004. Parenting stress, infant emotion regulation, maternal sensitivity, and the cognitive development of triplets: a model for parent and child influences in a unique ecology. Child Dev. 75, 1774–1791.

Finegood, E.D., Blair, C., Granger, D.A., Hibel, L.C., Mills-Koonce, R., 2016. Psychobiological influences on maternal sensitivity in the context of adversity. Dev. Psychol. 52, 1073.

Fleming, A.S., Steiner, M., Anderson, V., 1987. Hormonal and attitudinal correlates of maternal behaviour during the early postpartum period in first-time mothers. J. Reprod. Infant Psychol. 5, 193–205.

Fleming, A.S., Steiner, M., Corter, C., 1997. Cortisol, hedonics, and maternal responsiveness in human mothers. Horm. Behav. 32, 85–98.

Fleming, A.S., Corter, C., Stallings, J., Steiner, M., 2002. Testosterone and prolactin are associated with emotional responses to infant cries in new fathers. Horm. Behav. 42, 399–413.

Gettler, L.T., McDade, T.W., Feranil, A.B., Kuzawa, C.W., 2011. Longitudinal evidence that fatherhood decreases testosterone in human males. Proc. Natl. Acad. Sci. U. S. A. 108, 16194–16199.

Gilbert, R., Widom, C., Browne, K., Fergusson, D., Webb, E., Janson, S., 2009. Burden and consequences of child maltreatment in high-income countries. Lancet 373, 68–81.

Gonzalez, A., Jenkins, J.M., Steiner, M., Fleming, A.S., 2012. Maternal early life experiences and parenting: the mediating role of cortisol and executive function. J. Am. Acad. Child Adolesc. Psychiatry 51, 673–682.

Gordon, I., Pratt, M., Bergunde, K., Zagoory-Sharon, O., Feldman, R., 2017. Testosterone, oxytocin, and the development of human parental care. Horm. Behav. 93, 184–192.

Gray, P.B., Anderson, K.G., 2010. Fatherhood: Evolution and Human Paternal Behavior. Harvard University Press.

Heinrichs, M., Neumann, I., Ehlert, U., 2002. Lactation and stress: protective effects of breast-feeding in humans. Stress 5, 195–203.

Heinrichs, M., Baumgartner, T., Kirschbaum, C., Ehlert, U., 2003. Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. Biol. Psychiatry 54, 1389–1398.

Helmerhorst, K.O., Riksen-Walraven, J.M., Vermeer, H.J., Fukkink, R.G., Tavecchio, L.W., 2014. Measuring the interactive skills of caregivers in child care centers: development and validation of the caregiver interaction profile scales. Early Educ. Dev. 25, 770–790.

Jansen, J., Beijers, R., Riksen-Walraven, M., de Weerth, C., 2010. Cortisol reactivity in young infants. Psychoneuroendocrinology 35, 329–338.

Keyes, K.M., Eaton, N.R., Krueger, R.F., McLaughlin, K.A., Wall, M.M., Grant, B.F., Hasin, D.S., 2012. Childhood maltreatment and the structure of common psychiatric disorders. Br. J. Psychiatry 200, 107–115.

Kuo, P.X., Saini, E.K., Thomason, E., Schultheiss, O.C., Gonzalez, R., Volling, B.L., 2016. Individual variation in fathers' testosterone reactivity to infant distress predicts parenting behaviors with their 1-year-old infants. Dev. Psychobiol. 58, 303–314.

Mascaro, J.S., Hackett, P.D., Rilling, J.K., 2013. Testicular volume is inversely correlated with nurturing-related brain activity in human fathers. Proc. Natl. Acad. Sci. 110, 15746–15751.

McCanne, T.R., Hagstrom, A.H., 1996. Physiological hyperreactivity to stressors in physical child abusers and individuals at risk for being physically abusive. Aggress. Violent Behav. 1, 345–358.

Mehta, P.H., Prasad, S., 2015. The dual-hormone hypothesis: a brief review and future research agenda. Curr. Opin. Behav. Sci. 3, 163–168.

Mills-Koonce, W.R., Propper, C., Gariepy, J.L., Barnett, M., Moore, G.A., Calkins, S., Cox, M.J., 2009. Psychophysiological correlates of parenting behavior in mothers of young children. Dev. Psychobiol. 51, 650–661.

Morris, A.S., Cui, L., Steinberg, L., 2013. Parenting research and themes: what we have learned and where to go next. In: Larzelere, R.E., Morris, A.S., Harrist, A.W. (Eds.), Authoritative Parenting: Synthesizing Nurturance and Discipline for Optimal Child Development. American Psychological Association, Washington, DC, US, pp. 35–58.

Naber, F., van Ijzendoorn, M.H., Deschamps, P., van Engeland, H., Bakermans-Kranenburg, M.J., 2010. Intranasal oxytocin increases fathers' observed responsiveness during play with their children: a double-blind within-subject experiment. Psychoneuroendocrinology 35, 1583–1586.

Out, D., Bakermans-Kranenburg, M.J., Pelt, Jv., IJzendoorn, M.H.V., 2012. Salivary α-amylase and intended harsh caregiving in response to infant crying. Child Maltreat. 17, 295–305.

Pechtel, P., Pizzagalli, D., 2011. Effects of early life stress on cognitive and affective function: an integrated review of human literature. Psychopharmacology 214, 55–70.

Rilling, J.K., 2013. The neural and hormonal bases of human parental care. Neuropsychologia 51, 731–747.

Rutherford, H.J., Booth, C.R., Luyten, P., Bridgett, D.J., Mayes, L.C., 2015. Investigating the association between parental reflective functioning and distress tolerance in motherhood. Infant Behav. Dev. 40, 54–63.

Rutherford, H.J., Guo, X.M., Graber, K.M., Hayes, N.J., Pelphrey, K.A., Mayes, L.C., 2017. Intranasal oxytocin and the neural correlates of infant face processing in non-parent women. Biol. Psychol. 129, 45–48.

Saltzman, W., Maestripieri, D., 2011. The neuroendocrinology of primate maternal behavior. Prog. Neuro-Psychopharmacol. Biol. Psychiatry 35, 1192–1204.

Stallings, J., Fleming, A.S., Corter, C., Worthman, C., Steiner, M., 2001. The effects of infant cries and odors on sympathy, cortisol, and autonomic responses in new mothers and nonpostpartum women. Parenting 1, 71–100.

Storey, A.E., Walsh, C.J., Quinton, R.L., Wynne-Edwards, K.E., 2000. Hormonal correlates of paternal responsiveness in new and expectant fathers. Evol. Hum. Behav. 21, 79–95.

Storey, A.E., Noseworthy, D.E., Delahunty, K.M., Halfyard, S.J., McKay, D.W., 2011. The effects of social context on the hormonal and behavioral responsiveness of human fathers. Horm. Behav. 60, 353–361.

Storey, A.E., Ziegler, T.E., 2016. Primate paternal care: interactions between biology and social experience. Horm. Behav. 77, 260–271.

Sturge-Apple, M.L., Davies, P.T., Cicchetti, D., Cummings, E.M., 2009. The role of mothers' and fathers' adrenocortical reactivity in spillover between interparental conflict and parenting practices. J. Fam. Psychol. 23, 215.

Tabachnick, B.G., Fidell, L.S., 2007. Using Multivariate Statistics.

Tharner, A., Luijk, M.P., Raat, H., IJzendoorn, M.H., Bakermans-Kranenburg, M.J., Moll, H.A., Jaddoe, V.W., Hofman, A., Verhulst, F.C., Tiemeier, H., 2012. Breastfeeding and its relation to maternal sensitivity and infant attachment. J. Dev. Behav. Pediatr. 33, 396–404.

Thompson, L.A., Trevathan, W.R., 2008. Cortisol reactivity, maternal sensitivity, and learning in 3-month-old infants. Infant Behav. Dev. 31, 92–106.

Thompson, R., Gupta, S., Miller, K., Mills, S., Orr, S., 2004. The effects of vasopressin on human facial responses related to social communication. Psychoneuroendocrinology 29, 35–48.

van Anders, S.M., Tolman, R.M., Volling, B.L., 2012. Baby cries and nurturance affect testosterone in men. Horm. Behav. 61, 31–36.

van Anders, S.M., Tolman, R.M., Jainagaraj, G., 2014. Examining how infant interactions influence men's hormones, affect, and aggression using the Michigan infant nurturance simulation paradigm. Fathering 12, 143.

van Honk, J., Schutter, D.J., Bos, P.A., Kruijt, A.-W., Lentjes, E.G., Baron-Cohen, S., 2011. Testosterone administration impairs cognitive empathy in women depending on second-to-fourth digit ratio. Proc. Natl. Acad. Sci. U. S. A. 108, 3448–3452.

Voorthuis, A., Out, D., van der Veen, R., Bhandari, R., van Ijzendoorn, M.H., Bakermans-Kranenburg, M.J., 2013. One doll fits all: validation of the Leiden Infant Simulator Sensitivity Assessment (LISSA). Attach. Hum. Dev. 15, 603–617.

Voorthuis, A., Bakermans-Kranenburg, M.J., van Ijzendoorn, M.H., 2017. Testosterone reactivity to infant crying and caregiving in women: the role of oral contraceptives and basal cortisol. Infant Behav. Dev. https://doi.org/10.1016/j.infbeh.2017.08.002.

Weisman, O., Zagoory-Sharon, O., Feldman, R., 2014. Oxytocin administration, salivary testosterone, and father–infant social behavior. Prog. Neuro-Psychopharmacol. Biol. Psychiatry 49, 47–52.

Welker, K.M., Lozoya, E., Campbell, J.A., Neumann, C.S., Carré, J.M., 2014. Testosterone, cortisol, and psychopathic traits in men and women. Physiol. Behav. 129, 230–236.

Zeifman, D.M., 2001. An ethological analysis of human infant crying: answering Tinbergen's four questions. Dev. Psychobiol. 39, 265–285.