Cortisol and testosterone concentrations during the prenatal and postpartum period forecast later caregiving quality in mothers and fathers

Beijers, R.; Breugelmans, S.; Brett, B.; Willemsen, Y.; Bos, P.A.; De Weerth, C.

Hormones and Behavior 142 (2022) 105177

Contents lists available at ScienceDirect

Hormones and Behavior

journal homepage: www.elsevier.com/locate/yhbeh

Cortisol and testosterone concentrations during the prenatal and postpartum period forecast later caregiving quality in mothers and fathers

Roseriet Beijers[a,b,*,1], Sara Breugelmans[a,1], Bonnie Brett[b], Yvonne Willemsen[b], Peter Bos[c], Carolina de Weerth[b]

[a] Department of Social Development, Behavioural Science Institute, Radboud University, the Netherlands
[b] Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, the Netherlands
[c] Institute of Education and Child Studies, Leiden University, the Netherlands

ARTICLE INFO

Keywords:
Caregiving quality
Testosterone
Cortisol
Toddlerhood
Prenatal
Postnatal

ABSTRACT

Given that parental caregiving quality affects child development from birth onwards, it is important to detect parents who are at risk for low-quality caregiving as early as possible, preferably before or soon after birth. This study investigated whether cortisol (CORT) and testosterone (T) measured during the last trimester of pregnancy and six weeks postpartum were associated with observed caregiving quality at child age 3 in mothers ($N = 63$) and fathers ($N = 45$). CORT and T were measured during an interaction with a simulator infant (pregnancy) and their own infant (postpartum). In mothers, no associations were found with CORT and T during pregnancy, but higher postpartum CORT during a mother-infant interaction was related to higher caregiving quality during toddlerhood. In fathers, the association between T during pregnancy and caregiving quality in toddlerhood was more negative for fathers with low CORT. In contrast to mothers, higher postpartum CORT in fathers was associated with lower caregiving quality in toddlerhood. These findings proved robust after applying the Benjamini–Hochberg procedure to control for false discovery rate. Our findings indicate that CORT and T during the perinatal period can forecast caregiving quality in both mothers and fathers. Moreover, our results provided evidence for the dual-hormone hypothesis, but only in fathers. These findings contribute to our growing understanding on how endocrine measures explain individual differences in caregiving quality in mothers and fathers.

High parental caregiving quality contributes positively to a child's cognitive and social-emotional development, whereas low parental caregiving quality increases the risk for a range of poor child outcomes, including behavior problems and psychopathology (Gilbert et al., 2009; Keyes et al., 2012; Pechtel and Pizzagalli, 2011; Vesely et al., 2013). Given that parental caregiving quality can affect child development from birth onwards (Feldman et al., 2004), it is important to detect parents who are at risk for low-quality caregiving as early as possible, preferably soon after or even before birth with the aim of offering preventive interventions. The steroid hormones cortisol (CORT) and testosterone (T) are suggested to account for individual differences in parental caregiving quality (Bos, 2017; Storey et al., 2020). Moreover, in a previous study in the current sample, CORT, T, and their interaction were predictive of parenting quality in the prenatal and postnatal periods, but in fathers only (Bos et al., 2018). Whether these predictive effects extend into toddlerhood remains unknown. The current study aimed to investigate whether CORT, T and their interactions measured during infant caregiving in the prenatal and early postnatal periods predict parental caregiving quality at child age 3.

Previous studies on associations between hormones and parenting focused mostly on mothers and CORT, and fathers and T. CORT is the end product of the hypothalamic-pituitary-adrenal (HPA) axis and is seen as the body's primary stress hormone. In relation to maternal caregiving behavior, the cross-sectional study of Stallings et al. (2001) found that CORT elevations were positively associated with mothers' nurturing and attentive responses to their newborns. However, at later child ages, associations between CORT and maternal caregiving behavior seem to be reversed. For instance, higher CORT baseline (Mills-Koonce et al., 2009) and decreased CORT downregulation (Thompson and Trevathan, 2008) were associated with lower maternal sensitivity

* Corresponding author at: Department of Social Development, Behavioural Science Institute, Radboud University, the Netherlands.
E-mail address: roseriet.beijers@ru.nl (R. Beijers).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.yhbeh.2022.105177
Received 6 October 2021; Received in revised form 21 March 2022; Accepted 12 April 2022
Available online 2 May 2022
0018-506X/

towards 3- to 6-month-old children. Results from Finegood et al. (2016) indicated that this association may extend to the second postpartum year as well: higher maternal salivary CORT was associated with lower maternal sensitivity observed during parent–child interactions across the first 2 postpartum years. Compared to maternal caregiving, much less is known on the role of CORT in paternal caregiving. Higher paternal CORT levels and reactivity to the first time holding their babies forecasted more involvement in play and childcare in fathers 2–4 months later (Kuo et al., 2018). In contrast, in fathers of 4- to 8-month old infants, higher paternal CORT levels correlated with less proprioceptive touch and neutral affect towards the child, and with less gaze synchrony between father and child (Weisman et al., 2013). Similarly, in fathers of 1- to 5-year-old children, a range of fathering behaviors, including cuddling and joint play, were related to lower CORT (Ahnert et al., 2021). In a previous study in the current sample, we found that higher prenatal CORT, measured during an interaction with an unsoothable crying simulator infant, forecasts lower caregiving quality in fathers six weeks postpartum (Bos et al., 2018).

More often studied in relation to paternal caregiving is the hormone T. T is the end product of the hypothalamic-pituitary-gonadal (HPG) axis and is critical for reproductive, sexual and dominance behavior (Bos, 2017). In general, fathers have lower T compared to non-fathers (Gettler et al., 2011; Grebe et al., 2019; Perini et al., 2012). This T down-regulation is suggested to reflect a shift away from mating and towards parenting effort, and suggested to be adaptive since higher T concentrations are associated with aggression and less empathy (Carré et al., 2011; Hermans et al., 2006). Lower T might therefore facilitate higher quality of paternal care. Indeed, fathers with lower T showed more nurturing behavior towards their 1-year-olds (Kuo et al., 2016), and reported greater involvement in direct (e.g., bathing) and indirect (e.g., washing infant's clothes) childcare (Kuo et al., 2018). Also, fathers' engagement in childcare, attachment to the child, and cuddling appeared more pronounced the greater the T diurnal decline (Ahnert et al., 2021). In contrast, fathers' higher T was associated with increased child abuse risk, less observed positive parenting, and more observed negative parenting towards their 18-month-olds (Rodriguez et al., 2020). It is unclear how this relation between T and caregiving quality further develops as children age and men shift back to a mating orientation (Corpuz et al., 2020; Rosenbaum et al., 2018). The study from Corpuz et al. (2020) detected a T rebound in fathers from when infants were 3 months old to when infants were 9–10 months. As predicted, the more accelerated this T rebound, the less time fathers invested in their infants. In contrast, however, more accelerated T rebounds were related to higher levels of paternal care quality during a structured infant challenge activity. These results reveal nuance to the relation between T and paternal care.

The few studies that examined associations between T and maternal caregiving behavior indicated mixed results. While higher T was associated with more sensitivity in one study (Endendijk et al., 2016), and lower maternal engagement in another study (Clowtis et al., 2016), T was not predictive of maternal behavior in other studies (Bos et al., 2018; Gordon et al., 2017; Rodriguez et al., 2020). It thus remains unclear whether and how T is related to maternal caregiving quality.

In general, most studies focused on either CORT or T. However, several studies also reported positive, bidirectional within-person CORT and T relations (e.g. Crewther et al., 2021; Harden et al., 2016; Marceau et al., 2015). This CORT-T “coupling” may be associated with social behavior (Knight et al., 2020). According to the dual hormone hypothesis, T's association with behavior depends on CORT levels. More specifically, the association between T and behavior is hypothesized to be more robust when CORT levels are low, while this association is attenuated when CORT levels are high (Knight et al., 2020; Mehta and Josephs, 2010; Mehta and Prasad, 2015). For instance, Mehta and Josephs (2010) found that T increases dominant behavior, but only for individuals with low CORT. So far, research on the dual-hormone hypothesis has mainly focused on status-seeking behaviors, like aggression and dominance. However, the previously mentioned study by Bos et al. (2018) was the first to extend the dual-hormone hypothesis to parental caregiving behavior. Prenatally, in line with the predictions of the dual-hormone hypothesis, a stronger negative association was found between T and caregiving behavior in fathers with lower CORT compared to fathers with higher CORT. Postnatally, however, a stronger negative association between CORT and caregiving behavior was found in fathers with higher T compared to fathers with lower T. Similarly, a recent study in 79 first-time fathers of 2- to 4-month-old infants revealed no significant associations between paternal sensitivity and basal hormone levels or reactivity, but found that cortisol and testosterone interacted. Fathers with low basal cortisol levels showed more sensitivity with increasing T levels, but fathers with high cortisol levels were less sensitive with increasing T levels (Lotz et al., 2022).

While these results of Bos et al. (2018) and Lotz et al. (2022) suggest that T, CORT and their interaction forecast caregiving behavior in fathers during the perinatal period, it is still unknown whether CORT and T during the perinatal period remain predictive of caregiving quality later in childhood. During the transition from infancy to toddlerhood, parents need to adapt their caregiving behavior in a period of marked child developmental change. While during infancy children mostly depend on their parents for processes including co-regulation, the transition to toddlerhood is characterized by increased child self-regulation and emerging autonomy. During this time of growing independency and autonomy, it remains important for parents to be sensitive and understand the signals of the child, such as requests for autonomy in managing everyday situations, to aid the child in these developmental tasks, and at the same time to maintain closeness (Burrous et al., 2009; Calkins et al., 1998; Sroufe et al., 2005). The current, preregistered study investigated the following questions: (1) Are CORT, T and their interaction measured *in the third trimester of pregnancy* associated with caregiving quality at child age 3?, and (2) Are CORT, T, and their interaction measured *six weeks postpartum* associated with parental caregiving quality at child age 3? To account for potentially concurrent links, we also answered the question: (3) Are CORT, T, and their interaction measured *in toddlerhood* associated with caregiving quality in toddlerhood? For these three research questions, and following the dual-hormone hypothesis, we hypothesized that higher CORT and T would predict lower paternal caregiving quality in toddlerhood, and that the negative association between T and caregiving quality would be stronger for fathers with lower CORT. Given that studies find both positive and negative associations between CORT, T and maternal caregiving behavior (Endendijk et al., 2016; Clowtis et al., 2016; Finegood et al., 2016; Mills-Koonce et al., 2009; Thompson and Trevathan, 2008), while other studies do not find any associations (Bos et al., 2018; Gordon et al., 2017; Rodriguez et al., 2020), we did not set up specific hypotheses on the direction of the associations between T, CORT and maternal caregiving quality in toddlerhood.

1. Method

1.1. Participants

Participants were part of the *BINGO* (Dutch acronym for Biological Influences on Baby's Health and Development) study, a longitudinal study examining early-life predictors of parental caregiving quality and infant health. This study was approved by the ethical committee of the Faculty of Social Sciences of the Radboud University [ECSW2014-1003-189 and ECSW-2018-034]. Participants were recruited via folders distributed in midwife practices and pregnancy courses in the region Arnhem-Nijmegen (the Netherlands). To facilitate a higher number of participants, mothers were also able to participate without their partner.

Exclusion criteria during pregnancy were: insufficient mastery of the Dutch language, drug use, excessive alcohol use, and severe pregnancy complications. In total, 88 expectant mothers and 57 of their partners participated in the study and signed informed consent. Mothers participated without the father because he had no time to participate ($n = 19$), had no interest in participating ($n = 7$), was a donor ($n = 2$), was known to the researchers ($n = 1$), or refused without reason ($n = 2$).

Participants were excluded postnatally if they experienced complications during pregnancy (after initial contact), or if the infant was born prematurely (gestational age ≤ 37 weeks), with a low birth weight (<2500 g), with a low 5-minute Apgar score (<7), or with anomalies. Two families were excluded from further analyses due to prematurity ($n = 1$), and brain damage at birth (n = 1). Seven families stopped participation after birth due to personal reasons. In addition, fifteen families did not participate in the 3-year measurement round. The final sample therefore consisted of 64 mothers and 46 of their partners. Infants (30 boys, 34 girls) were born full-term ($M = 39.81$ weeks, $SD = 1.52$ weeks), with an average birth weight of 3529.66 g ($SD = 421.00$ g). The study design, hypotheses and analyses for the current study were preregistered on AsPredicted (https://aspredicted.org/blind.php?x=s9ym5k).

1.2. 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 the late afternoon or in the early evening ($M = $ 17:28 h, $SD = 1$ h and 53 min). During this visit, the mothers and fathers were independently introduced to an infant simulator (RealCare Baby; Realityworks) and asked to take care of the infant simulator as if it were their real infant. The infant simulator cried unsoothably after the experimenter left. After 15 min, the experimenter re-entered the room and the participant was asked two manipulation check questions: (1) *How difficult did you find it to interact with the infant simulator?*, and (2) *How seriously did you perform the task?* As there is an estimated lapse of around 20 to 30 min between an event and the production and release of cortisol in saliva (Dickerson and Kemeny, 2004; Gunnar and White, 2001), participants provided saliva samples before the interaction task (baseline), and 15 min (response to the interaction) and 35 min (recovery) after the end of the interaction by means of passive drooling. The saliva samples were immediately stored at −20 °C.

When study infants were around 6 weeks old ($M = 6.77$, $SD = 0.82$), parents were visited at home. All visits took place during the late afternoon or in the early evening ($M = $ 17:40 h, $SD = 1$ h and 59 min). Mothers and fathers were asked to independently undress, change and redress their infant as they would normally do. The interaction was videotaped as unobtrusively as possibly by the experimenter for 15 min. When both parents participated, the mother interacted with the infant first. Participants provided three saliva samples: before the start of the interaction with their infant (baseline), and at 15 and 35 min after the end of the interaction. The saliva samples were transported with a portable freezer and stored at −20 °C.

During a home visit at child age 3 years, mother-child and father-child interactions were videotaped during 12-to-15-minute semi-structured play. All visits took place during the morning and afternoon ($M = $ 11:45 h, $SD = 2$ h and 26 min). The mother was asked to read a book, do a jigsaw puzzle and play with building blocks with the child, in this order. The father was asked to have the child put shaped blocks into a form, categorize blocks on shape and color, and play a fishing game, in this order. Both parents were asked to play with their child as they would normally do. When both parents participated, the mother interacted with the child first. Participants provided three saliva samples: before the start of the interaction with their toddler (baseline), and at 15 and 35 min after the end of the interaction. The saliva samples were transported with a portable freezer and stored at −20 °C.

1.3. Measures

1.3.1. Parental caregiving quality in toddlerhood

The videotaped parent-child interactions at child age 3 were rated for parental caregiving quality behavior using 7-point rating scales (Erickson et al., 1985): *Supportive presence* (i.e. the provision of emotional support and positive regard to the child) and *Respect for autonomy* (i.e. acknowledgment of the validity of the child's perspective and non-intrusiveness). These two scales were used since they were found to adequately capture negative patterns of interaction (Smeekens et al., 2007). Trained observers independently rated the interactions. About 30% of the videos were scored twice for reliability. Interrater agreement was good (ICC = 0.93 and 0.90 for *Supportive Presence* and *Respect for Autonomy*, respectively). Since *Supportive presence* and *Respect for autonomy* were highly and positively correlated ($r = 0.65$, $p < .001$), the scores were averaged to obtain one measure of parental caregiving quality.

1.3.2. Cortisol

The saliva samples were analyzed at the University Medical Center of Utrecht University, the Netherlands. Saliva was thawed and assayed. The concentration of CORT in each sample 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. For the prenatal and postpartum samples at infant age 6 weeks, 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% ($n = 10$). For the samples at child age 3 years, the lower limit of detection was 1.0 nmol/l and inter-assay variation was <6% at 2.5–28 nmol/l ($n = 19$). Intra-assay variation was <4% (n = 10).

1.3.3. Testosterone

After determination of CORT, the concentration of salivary T in the prenatal and postpartum samples at infant age 6 weeks were analyzed at Nagasaki University, Japan, lab by enzyme immunoassay (EIA) using a commercially available kit (Salimetrics Europe Ltd., Suffolk, UK). The sample was thawed, centrifuged at 1500 ×g for 15 min, and the aqueous layer was aliquoted for assay. The cumulative intra-assay CV was <5%. 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. Inter-assay variation was 6.7% at 29 ng/ml and 4.0% at 448 ng/ml ($n = 3$). The information about the recovery and specificity of the kit can be found in the EIA kit online manual (https://salimetrics.com/wp-content/uploads/2018/03/testosterone-saliva-elisa-kit.pdf).

The saliva samples at child age 3 years were analyzed at the University Medical Center of Utrecht University, the Netherlands, also by enzyme immunoassay (EIA) using a commercially available kit (1-2402, Salimetrics, USA). The lower limit of detection was 20 pmol/l. Inter-assay variation was <8% at 80 pmol/l and <4.5% at 660 pmol/l ($n = 20$). Intra-assay variation was 5.4% at 85 pmol/L and 4.0% at 425 pmol/l ($n = 10$).

1.3.4. Control variables

To investigate whether T and CORT are associated with caregiving quality in toddlerhood above and beyond early life caregiving quality, prenatal caregiving quality and postpartum caregiving quality at infant age six weeks were included as control variables. Videos of the prenatal interaction with the unsoothable simulator infant were rated for sensitivity and cooperation using 9-point scales (Ainsworth et al., 2015). 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 = 0.92 and 0.88 for sensitivity and cooperation, respectively). Sensitivity and cooperation were highly correlated ($r = 0.88$) and therefore averaged as a measure for prenatal quality of caregiving. Similarly, videos of the interaction with their own 6-week-old infant were rated for sensitivity and cooperation with the same scales. Again, about 30% of the videos were rated twice for reliability, and interrater agreement was good (ICC = 0.82 and

0.75 for sensitivity and cooperation, respectively). Sensitivity and cooperation were highly correlated ($r = 0.81$) and therefore averaged as a measure for postpartum quality of caregiving.

Additionally, parental age and educational level were included as control variables as they have been implicated in the relation between T, CORT, and parental behavior (Bos, 2017; Saltzman and Maestripieri, 2011; Storey and Ziegler, 2016). Also, in the analyses including prenatal caregiving quality towards the simulator infant, the reported difficultness and seriousness in interacting with the simulator infant, filled out on a 7-point scale, were added as control variables. Finally, in toddlerhood, home visits were not restricted to a certain time window (i.e. took place during both the morning and the (early) afternoon). To control for the CORT and T circadian rhythm, time of the home visit was added as control variable in the analyses investigating maternal and paternal toddlerhood CORT and T associations with parental caregiving quality in toddlerhood.

1.3.5. Missing data

During the home visit at child age 3, one mother and father did not complete the interaction with their own child because their child was ill. This family was excluded from further analyses. In addition, it was not possible to rate two videos (one mother, one father) due to technical issues of the videorecording system. Saliva samples of mothers who had used antibiotics during pregnancy ($N = 2$) and after birth (N = 2) were excluded from the hormonal analyses. Because multilevel analyses are robust to missing values, missing values were not imputed.

1.4. Statistical analyses

First, the data was checked for outliers and subsequently these were winsorized (i.e. outliers were replaced with the value of 3SD away from the mean; Tukey, 1962) for the following variables: maternal prenatal T ($N = 1$), maternal postpartum CORT ($N = 2$), maternal postpartum T ($N = 1$), maternal CORT at child age 3 ($N = 1$), and paternal CORT at child age 3 ($N = 4$). The outliers mostly belonged to individual cases, except for paternal CORT at child age 3 where one father had 3 outliers. Subsequent sensitivity analyses with the outliers included and with the outliers excluded rendered similar results.

The same statistical approach as used by Bos et al. (2018) was adopted and is as follows: mothers and fathers were analyzed separately. Because of the longitudinal design (CORT and T were examined three times per measurement wave), multilevel linear modeling (MLM) was used. MLM enables investigations of T and CORT responses (by including time), and their associations with quality of caregiving. Moreover, MLM is robust to missing data (Tabachnick and Fidell, 2007). The parent was the level 2 identifier, and the outcome and predictors were the level 1 variables. To investigate whether caregiving quality in toddlerhood was associated with parental CORT and T during toddlerhood, caregiving quality in toddlerhood was added as a predictor to the multilevel models predicting toddlerhood CORT and T. To investigate whether caregiving quality in toddlerhood is associated with prenatal and postpartum CORT and T, caregiving quality in toddlerhood was added as a predictor to the multilevel models predicting prenatal and postpartum CORT and T. The analyses were run in SPSS (version 25.0).

First, an intercept-only model was run to check whether a multilevel linear model was appropriate. For mothers, the intraclass correlations were 0.49 for prenatal CORT, 0.65 for postpartum CORT, 0.62 for toddlerhood CORT, 0.56 for prenatal T, 0.33 for postpartum T and 0.76 for toddlerhood T. The intraclass correlations for the father mixed model analyses were 0.77 for prenatal CORT, 0.67 for postpartum CORT, 0.78 for toddlerhood CORT, 0.69 for prenatal T, 0.64 for postpartum T and 0.65 for toddlerhood T. Thus, multilevel linear modeling was appropriate.

Afterwards, a build-up strategy was used by adding the variables one by one to the intercept-only model. After each addition, the −2 log likelihood, a determinant for model fit, was examined. Only the variables that improved the model fit significantly were added to the model. First, time was entered to investigate the reactivity of CORT and T on the interaction task. Then, the control variables were entered one at a time, followed by the parental caregiving quality predictors. Thereafter, the interaction between time and caregiving quality in toddlerhood was entered to examine whether the reactivity of CORT and T were predicted by caregiving quality. Finally, interaction terms between caregiving quality in toddlerhood and CORT or T were added (i.e., the interaction between caregiving and CORT was entered into the model predicting T and vice versa). By doing so, it was possible to investigate whether the association between caregiving quality and one hormone was dependent on the concentration of the other hormone (i.e., dual hormone hypothesis). Prior to entering caregiving quality in toddlerhood and its interactions, variables were centered.

To answer question one (i.e., whether CORT, T and their interaction measured during late pregnancy were associated with parental caregiving quality in toddlerhood), two multilevel models were built: (1) prenatal CORT predicted by parental caregiving quality and (2) prenatal T predicted by parental caregiving quality. In both models, control variables were: age, educational level, difficulty and seriousness interacting with the simulator infant, prenatal caregiving quality and caregiving quality six weeks postpartum.

To answer question two (i.e., whether CORT, T and their interaction measured six weeks postpartum were associated with parental caregiving quality in toddlerhood), two multilevel models were built: (1) CORT measured six weeks postpartum predicted by parental caregiving quality (controlled for age, educational level, prenatal caregiving quality and caregiving quality six weeks postpartum), and (2) T measured six weeks postpartum predicted by parental caregiving quality (controlled for age, educational level, prenatal caregiving quality and caregiving quality six weeks postpartum).

To answer question three (i.e., whether CORT, T and their interaction in toddlerhood were associated with parental caregiving quality in toddlerhood), two multilevel models were built: (1) toddlerhood CORT predicted by parental caregiving quality (controlled for age and educational level), and (2) prenatal T predicted by parental caregiving quality (controlled for age and educational level).

As the multilevel models tested several predictors in several models, the Benjamini-Hochberg Procedure was applied to control for the false discovery rate (Vogt, 2005). The Benjamini-Hochberg Procedure was applied as follows: 1) the individual *p*-values were put in ascending order, 2) ranks were assigned to the p-values, 3) each individual p-value's Benjamini-Hochberg critical value was calculated using the formula (i/m)Q, where: i = the individual p-value's rank, m = total number of tests, and Q = the false discovery rate (i.e. 10%), and 4) the original p-value was compared to the Benjamini-Hochberg critical value. The number of tests is 24; 2 predictors of interest (i.e. caregiving quality/interaction with caregiving quality) × 2 multilevel models (i.e. CORT/T) × 3 time periods (i.e. prenatal/postpartum/toddlerhood) × 2 parents (i.e. mothers/fathers). For the results of the Benjamini Hochberg procedure, see Table 4.

2. Results

2.1. Descriptive statistics

Table 1 presents the means and standard deviations for the control variables, caregiving quality and CORT and T responses to the interaction with the toddler, separately for mothers and fathers. The manipulation check showed that parents did not find it easy nor difficult to interact with the simulator infant ($M = 4.19$, $SD = 1.74$). The parents performed the task with the simulator infant rather seriously ($M = 5.76$, $SD = 1.03$). In mothers, prenatal caregiving quality was positively correlated with caregiving quality 6 weeks postpartum ($r = 0.25$, $p = .001$), but not correlated with caregiving quality at child age 3 ($r = 0.05$, $p = .490$). Caregiving quality at 6 weeks postpartum was positively

Table 1
Descriptive statistics, separately for mothers ($N = 63$) and fathers ($N = 45$).

	Mothers: M (SD)	Fathers: M (SD)
Control variables		
Educational level (%)		
High school/job training	17.5%	15.5%
College/university	81.0%	62.2%
Missing	1.6%	22.2%
Age	34.90 (3.73)	36.68 (4.40)
Difficulty	4.19 (1.73)	4.50 (1.71)
Seriousness	5.76 (1.03)	5.50 (1.00)
Quality of caregiving[a]		
Prenatal	3.96 (2.01)	3.67 (1.46)
Six weeks postpartum	5.51 (1.78)	4.55 (1.79)
Toddlerhood	4.20 (1.21)	3.86 (1.20)
Parental CORT at child age 3		
Baseline	7.38 (2.23)	7.44 (3.02)
15 min after interaction	7.05 (2.19)	7.58 (2.99)
35 min after interaction	6.74 (1.92)	6.63 (4.05)
Parental T at child age 3		
Baseline	117.13 (65.41)	334.56 (120.44)***
15 min after interaction	119.32 (61.98)	372.91 (132.53)***
35 min after interaction	120.60 (61.44)	331.61 (149.81)***

[a] Prenatal caregiving quality and caregiving quality six weeks postpartum were measured on a 9-point Likert scale whereas caregiving quality in toddlerhood was measured on a 7-point Likert scale.
*** At child age 3, *t*-tests indicated that paternal T was significantly higher compared to maternal T at all time points.

correlated with caregiving quality at child age 3 ($r = 0.27$, $p < .001$). In fathers, prenatal caregiving quality was positively correlated with caregiving quality at 6 weeks postpartum ($r = 0.29$, $p = .015$) and caregiving quality at child age 3 ($r = 0.41$, $p < .001$). Caregiving quality at 6 weeks postpartum was not correlated with caregiving quality at child age 3 ($r = 0.04$, $p = .695$).

2.2. Multilevel analyses

First, the multilevel models are presented for mothers and then for fathers. In all models, CORT and T were positively associated (all p's < 0.001), except for the models in toddlerhood for fathers. See Table 2 (mothers) and Table 3 (fathers) for the summary of results of the final multilevel models.

2.2.1. Maternal caregiving quality

2.2.1.1. Prenatal CORT and T. Maternal caregiving quality in toddlerhood was not significantly associated with CORT or T.

2.2.1.2. Postpartum CORT and T. For postpartum CORT, there was a positive effect of caregiving quality in toddlerhood ($p = .013$). Higher maternal caregiving quality during toddlerhood was associated with higher CORT during a mother-infant interaction 6 weeks postpartum. Furthermore, there was a significant interaction effect between time and caregiving quality in toddlerhood ($p = .011$); CORT decreased somewhat more in response to the mother-infant interaction 6 weeks postpartum for those mothers with high caregiving quality in toddlerhood. The slopes were not significantly different from zero ($p = .777$, $p = .114$, for low and high caregiving quality respectively). Maternal caregiving quality in toddlerhood was not significantly associated with T.

2.2.1.3. Toddlerhood CORT and T. CORT was not significantly associated with maternal caregiving quality in toddlerhood. For T, there was a significant interaction effect between time and caregiving quality in toddlerhood ($p = .032$); T decreased slightly more in response to the

Table 2
Maternal multilevel models.

	Prenatal			Postpartum			Toddlerhood		
	Estimate	SE	*p*	Estimate	SE	*p*	Estimate	SE	*p*
Cortisol									
Intercept	12.82	5.29	0.019*	−9.71	8.11	0.234	7.05	2.69	0.011*
Time	−0.01	0.00	0.034*	−0.01	0.00	0.001**	−0.00	0.00	0.007**
Age	0.05	0.11	0.677	0.01	0.07	0.852	0.02	0.07	0.779
Educational level	−0.35	0.29	0.243	0.05	0.16	0.747	−0.34	0.19	0.077
Caregiving quality (prenatal)	0.05	0.23	0.816	0.19	0.14	0.178	–	–	–
Caregiving quality (postpartum)	0.03	0.26	0.924	0.13	0.15	0.402	–	–	–
Caregiving quality (3y)	0.20	0.36	0.579	3.91	1.56	0.013*	−0.10	0.22	0.458
T prenatal	0.10	0.01	<0.001**	–	–	–	–	–	–
T postpartum	–	–	–	0.09	0.01	<0.001	–	–	–
T toddlerhood	–	–	–	–	–	–	0.01	0.00	<0.001**
Caregiving quality (3y) * Time	–	–	–	−0.00	0.00	0.011*	–	–	–
Caregiving quality (3y) * T prenatal	0.01	0.01	0.326	–	–	–	–	–	–
Caregiving quality (3y) * T postpartum	–	–	–	−0.02	0.01	0.074	–	–	–
Caregiving quality (3y) * T toddlerhood	–	–	–	–	–	–	0.00	0.00	0.470
Deviance	646.011			417.433			644.307		
Testosterone									
Intercept	−79.47	27.58	0.005**	−6.17	12.28	0.617	186.39	73.01	0.013*
Time	0.08	0.02	<0.001**	0.02	0.01	0.109	0.00	0.00	0.092
Age	–	–	–	–	–	–	−5.10	1.83	0.007**
Educational level	3.97	2.07	0.062	1.04	0.90	0.255	7.83	5.22	0.138
Caregiving quality (prenatal)	–	–	–	−1.11	0.77	0.157	–	–	–
Caregiving quality (postpartum)	−0.78	1.81	0.668	−0.78	0.85	0.366	–	–	–
Caregiving quality (3y)	−3.34	2.64	0.213	0.02	1.13	0.988	1.14	6.30	0.857
CORT prenatal	3.96	0.42	<0.001**	–	–	–	–	–	–
CORT postpartum	–	–	–	3.14	0.48	<0.001**	–	–	–
CORT toddlerhood	–	–	–	–	–	–	7.21	1.83	<0.001**
Caregiving quality (3y) * Time	–	–	–	–	–	–	−0.00	0.00	0.032*
Caregiving quality (3y) * CORT prenatal	−0.14	0.34	0.654	–	–	–	–	–	–
Caregiving quality (3y) * CORT postpartum	–	–	–	−0.40	0.34	0.242	–	–	–
Caregiving quality (3y) * CORT toddlerhood	–	–	–	–	–	–	0.11	1.72	0.947
Deviance	1158.325			889.763			1713.173		

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

Table 3
Paternal multilevel models.

	Prenatal			Postpartum			Toddlerhood		
	Estimate	SE	p	Estimate	SE	p	Estimate	SE	p
Cortisol									
Intercept	6.41	6.89	0.358	7.81	3.24	0.021*	12.05	5.93	0.047*
Time	−0.00	0.01	0.840	−0.00	0.00	0.118	−0.00	0.00	0.372
Age	–	–	–	–	–	–	0.18	0.12	0.107
Educational level	–	–	–	–	–	–	−0.99	0.33	0.005**
Caregiving quality (prenatal)	−0.02	0.59	0.980	0.54	0.14	0.001**	–	–	–
Caregiving quality (postpartum)	−0.63	0.40	0.134	−0.44	0.15	0.008**	–	–	–
Caregiving quality (3y)	1.21	1.12	0.287	−5.50	2.02	0.011*	−1.59	0.79	0.047*
T prenatal	0.06	0.02	0.002**	–	–	–	–	–	–
T postpartum	–	–	–	0.04	0.01	<0.001**	–	–	–
T toddlerhood	–	–	–	–	–	–	0.00	0.00	0.160
Caregiving quality (3y) * Time	–	–	–	0.00	0.00	0.007**	–	–	–
Caregiving quality (3y) * T prenatal	−0.00	0.01	0.722	–	–	–	–	–	–
Caregiving quality (3y) * T postpartum	–	–	–	0.00	0.01	0.838	–	–	–
Caregiving quality (3y) * T toddlerhood	–	–	–	–	–	–	0.01	0.00	0.001**
Deviance	287.925			195.680			476.693		
Testosterone									
Intercept	89.79	43.73	0.047*	152.03	77.29	0.057	1117.58	203.04	<0.001**
Time	−0.05	0.04	0.238	−0.11	0.05	0.041*	0.00	0.01	0.365
Age	–	–	–	−0.46	1.11	0.685	−9.03	4.32	0.044*
Educational level	–	–	–	–	–	–	−21.84	13.24	0.108
Caregiving quality (prenatal)	–	–	–	−2.46	3.28	0.460	–	–	–
Caregiving quality (postpartum)	2.70	2.81	0.348	6.18	3.17	0.060	–	–	–
Caregiving quality (3y)	7.75	5.75	0.185	−1.04	6.80	0.879	−88.85	25.91	0.001**
CORT prenatal	4.28	0.95	<0.001**	–	–	–	–	–	–
CORT postpartum	–	–	–	8.60	1.56	<0.001**	–	–	–
CORT toddlerhood	–	–	–	–	–	–	3.19	3.56	0.372
Caregiving quality (3y) * CORT prenatal	−1.88	0.66	0.006**	–	–	–	–	–	–
Caregiving quality (3y) * CORT postpartum	–	–	–	−0.59	0.99	0.553	–	–	–
Caregiving quality (3y) * CORT toddlerhood	–	–	–	–	–	–	9.45	2.76	0.001**
Deviance	509.263			521.351			1214.562		

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

mother-toddler interaction for mothers with high caregiving quality, whereas T increased somewhat more in response to the mother-toddler interaction for mothers with low caregiving quality. The slopes were not significantly different from zero (p = .310, p = .651, for low and high caregiving quality respectively). Moreover, there was a significant negative effect of maternal age (p = .007). Mothers who were older

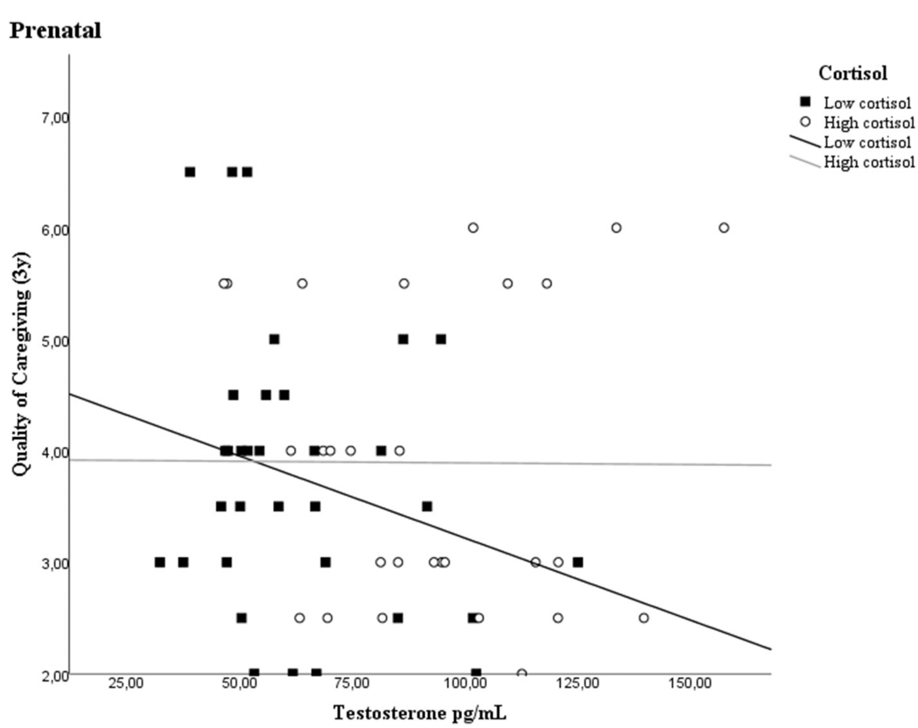

Fig. 1. Interaction between paternal quality of caregiving in toddlerhood and prenatal CORT on T.

showed lower T during the interaction with their toddler.

2.2.2. Paternal caregiving quality

2.2.2.1. Prenatal CORT and T. For prenatal CORT, there were no associations with caregiving quality in toddlerhood. For prenatal T, a significant negative interaction effect was found between prenatal CORT and caregiving quality in toddlerhood ($p = .006$). Using a median split, low and high CORT groups were created to visualize this interaction. Fig. 1 shows that the association between prenatal T and caregiving quality in toddlerhood was slightly more negative for the low CORT group than the high CORT group. The slopes were not significantly different from zero ($p = .153$ and $p = .976$, for low and high CORT respectively).

2.2.2.2. Postpartum T and CORT. For postpartum CORT, there was a significant negative effect of paternal quality of caregiving in toddlerhood ($p = .011$); lower caregiving quality during an interaction with their toddler was associated with higher CORT measured six weeks postpartum. In addition, there was a significant interaction effect between time and caregiving quality in toddlerhood ($p = .007$); the CORT decrease was more pronounced during the interaction with their 6-week-old infant for the fathers with high quality caregiving compared to fathers with low quality caregiving. The slopes were not significantly different from zero ($p = .117$, $p = .301$, for low and high caregiving quality respectively). For postpartum T, there was no significant effect of caregiving quality in toddlerhood.

2.2.2.3. Toddlerhood CORT and T. For CORT, a significant negative effect of caregiving quality in toddlerhood was found ($p = .047$). However, this effect did not survive the Benjamini-Hochberg Procedure (see Table 4). Additionally, there was a significant effect of the interaction between T and caregiving quality of fathers in toddlerhood ($p = .001$). Using a median split, low and high T groups were created. Fig. 2 shows that the association between CORT and caregiving quality in toddlerhood. The association between CORT and caregiving quality was positive for the high T group, but slightly negative in the low T group. When examining the slopes, the slope for the high T group was significant ($p < .05$), while the slope for the low T group was not significant ($p = .463$). Furthermore, a lower educational level of the father was associated with higher CORT during the interaction with their toddler ($p = .005$).

For T, a significant negative effect of caregiving quality in toddlerhood was found ($p = .001$). Lower caregiving quality was associated with higher T measured during an interaction with their toddler. In addition, there was a significant effect of the interaction between CORT and caregiving quality in toddlerhood ($p = .001$). Using a median split, low and high CORT groups were created. Fig. 3 shows that the association between T and caregiving quality in toddlerhood was more negative for fathers with low CORT. When examining the slopes, we found that the slope for the low CORT group was significant ($p = .002$), while the slope for the high CORT group was not significant ($p = .732$). Furthermore, fathers who were younger had higher T during the interaction with their toddler ($p = .044$).

3. Discussion

In our previous study in the current sample (Bos et al., 2018), we found that interactions between CORT and T were cross-sectionally related to caregiving quality during the prenatal and postpartum period, but in fathers only. Also in fathers, higher CORT during pregnancy forecasted lower caregiving quality during the postpartum period. In mothers, no evidence was found for associations between the endocrine measures and caregiving quality. In the current preregistered study, we extended these findings in the same cohort by investigating whether maternal and paternal CORT and T during the prenatal and postpartum periods were predictive of their caregiving quality towards their 3-year-olds. The following questions were answered: (1) Are CORT, T and their interaction measured in the third trimester of pregnancy associated with caregiving quality in toddlerhood?, and (2) Are CORT, T, and their interaction measured six weeks postpartum associated with parental caregiving quality in toddlerhood? The results indicated that, for mothers, no associations were found between CORT, T and interactions during pregnancy and caregiving quality in toddlerhood. However, mothers who showed higher CORT in response to the interaction with their 6-week-old infant showed higher caregiving quality when their child was 3 years old. For fathers, the association between T during pregnancy and caregiving quality in toddlerhood was more negative for fathers with low CORT. In contrast to mothers, fathers who showed higher CORT in response to the interaction with their 6-week-old infant showed lower caregiving quality when their child was 3 years old.

To account for potentially concurrent links, we also answered the question whether CORT, T, and their interaction measured in toddlerhood were associated with caregiving quality in toddlerhood. For mothers, T, but not CORT, was associated with caregiving quality: T decreased in response to the mother-toddler interaction for mothers with high caregiving quality, whereas T increased in response to the mother-toddler interaction for mothers with low caregiving quality. For fathers, CORT, T, and their interactions were associated with caregiving quality. High T was associated with lower paternal caregiving quality during a father-toddler interaction. Also, the interactions between CORT and T provided evidence for the dual-hormone hypothesis, and indicated that the association between T and caregiving quality in toddlerhood was more negative for fathers with low CORT. Below we will

Table 4
Benjamini Hochberg corrections.

Model	Variable	P value	Rank	Benjamini-Hochberg critical value - (i/m)Q
Paternal toddlerhood CORT*	Caregiving quality × T toddlerhood	0.001	1	0.004167
Paternal toddlerhood T*	Caregiving quality	0.001	2	0.008333
Paternal toddlerhood T*	Caregiving quality × CORT toddlerhood	0.001	3	0.012500
Paternal prenatal T*	Caregiving quality × CORT prenatal	0.006	4	0.016667
Paternal postpartum CORT*	Caregiving quality × time	0.007	5	0.020833
Paternal postpartum CORT*	Caregiving quality	0.011	6	0.025000
Maternal postpartum CORT*	Caregiving quality × time	0.011	7	0.029167
Maternal postpartum CORT*	Caregiving quality	0.013	8	0.033333
Maternal toddlerhood T*	Caregiving quality × time	0.032	9	0.037500
Paternal toddlerhood CORT	Caregiving quality	0.047	10	0.041667
Maternal postpartum CORT	Caregiving quality × T postpartum	0.074	11	0.045833

Note: i = the individual p-value's rank, m = total number of tests ($N = 24$ tests), Q = the false discovery rate (=10%).
*The original *p*-values below the Benjamini-Hochberg critical value (i/m)Q are highlighted in grey and considered significant.

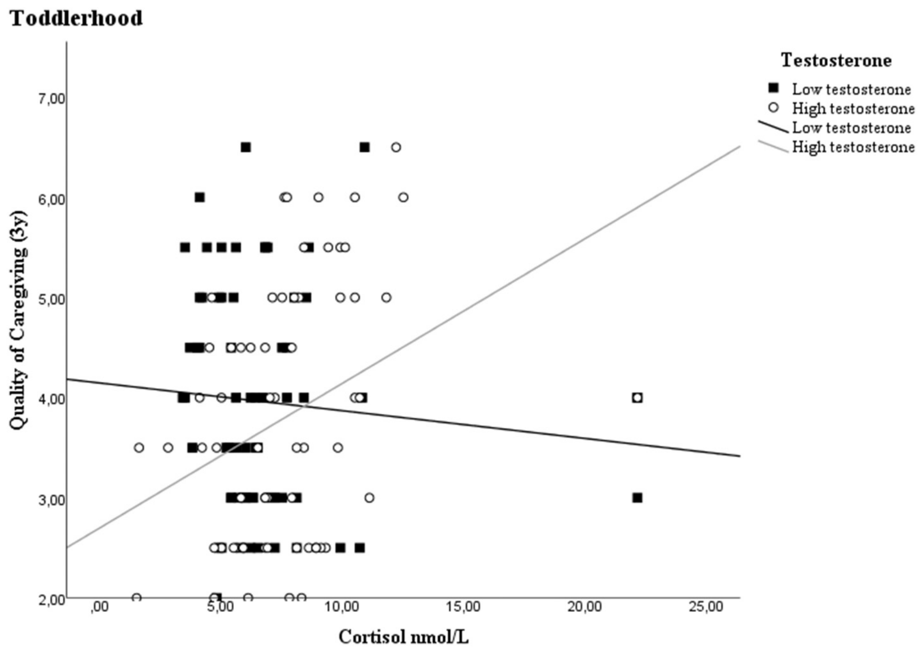

Fig. 2. Interaction between paternal quality of caregiving in toddlerhood and toddlerhood T on CORT.

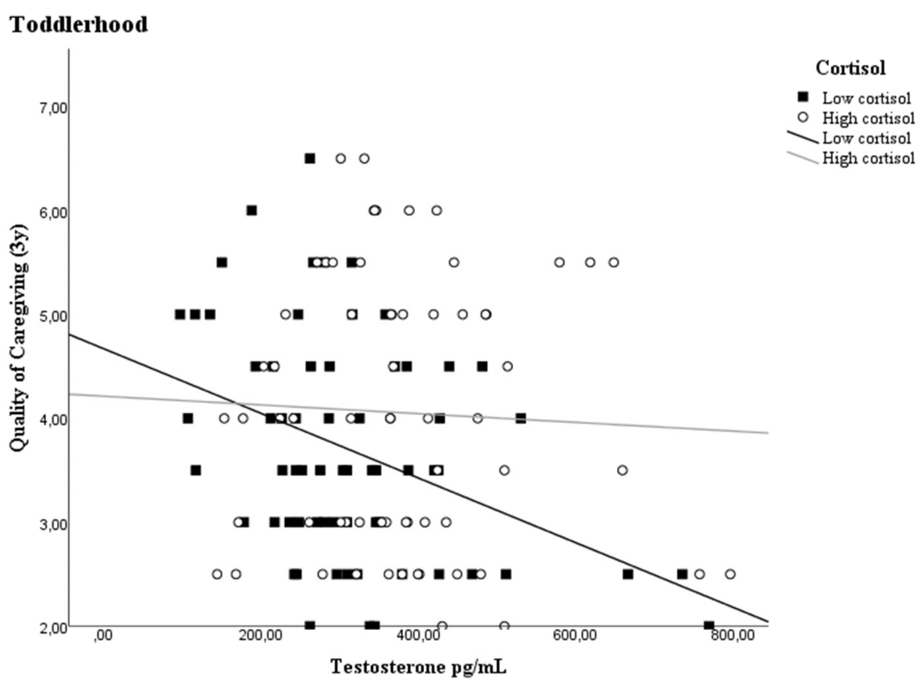

Fig. 3. Interaction between paternal quality of caregiving in toddlerhood and toddlerhood CORT on T.

discuss these findings, first for the mothers and then for the fathers. An overview of all results, including the previous findings in the same cohort (Bos et al., 2018), can be found in Fig. 4.

For mothers, no associations were observed between prenatal CORT, T or their interactions and maternal quality of caregiving at child age 3. In our previous study, the endocrine measures during late pregnancy were also not related to observed quality of caregiving in mothers when interacting with their 6-week-old infant (Bos et al., 2018). There are multiple explanations for why CORT and T during pregnancy do not predict later caregiving quality in mothers. First, mothers undergo many physiological adjustments during pregnancy to facilitate fetal growth and development, including changes in steroid hormones (Noyola-Martínez et al., 2019). For instance, CORT increases such that by the end of pregnancy, CORT levels are two to three times higher than normal (de Weerth and Buitelaar, 2005; Dörr et al., 1989). As a consequence, the HPA-axis may respond not-at-all or less to stressors. These higher CORT levels may have also dampened the stress response to the interaction with the unsoothable crying simulator infant in our study and obscured associations between prenatal endocrine measures and later caregiving quality. Another explanation is that the association of prenatal CORT or T and maternal caregiving quality depends on other endocrine hormones not taken into account in this study. Gordon et al. (2017), for instance, found that T was only associated with maternal caregiving behavior in interaction with oxytocin. Possibly more or other maternal

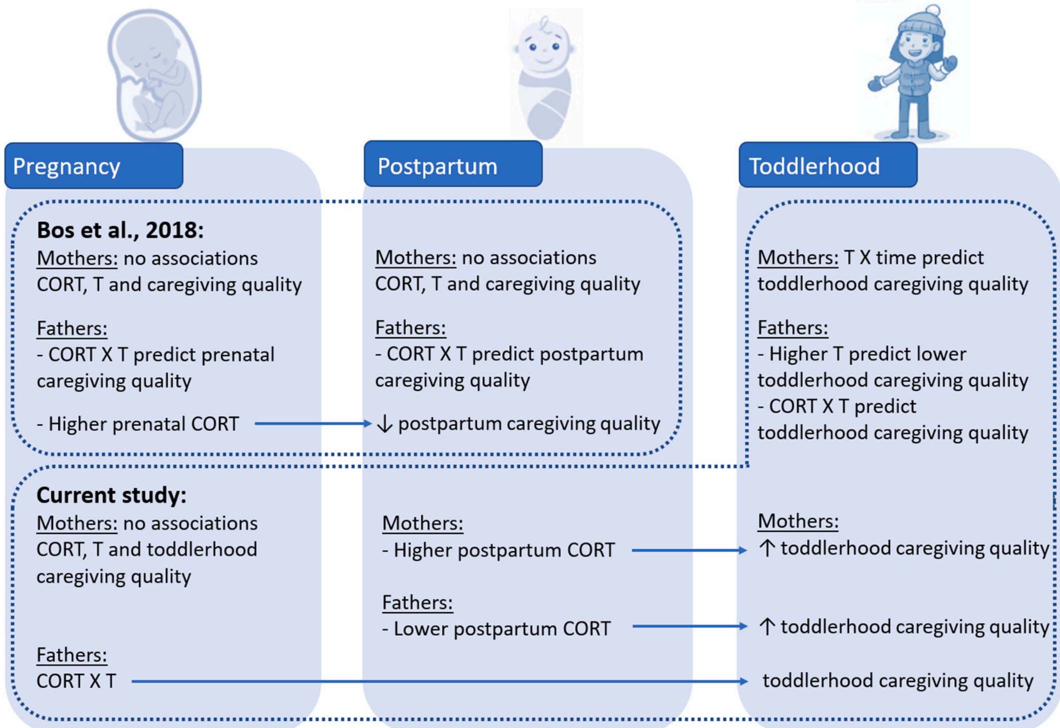

Fig. 4. Graphical abstract of the findings in the BINGO cohort on parental CORT, T and CORT × T associations with caregiving quality from pregnancy to age 3 years.

endocrine measures during pregnancy are needed to fully elucidate how different hormones during pregnancy forecast differences in postpartum caregiving quality.

Mothers who showed higher CORT in response to the interaction with their 6-week-old infant showed higher caregiving quality when their child was 3 years of age. These results are in line with those of Stallings et al. (2001), who found higher CORT to facilitate maternal caregiving behavior in the first postnatal days. The question remains why high maternal CORT during the postpartum period is associated with higher caregiving quality in toddlerhood. There is surprisingly little known on how endocrine factors mark human development in the timeframe of parturition and the early postpartum period (Bos, 2017). Giving birth is associated with a hormonal cascade in mothers resulting in high levels of OXT, AVP, catecholamines and also CORT (Bos, 2017). Furthermore, it is known that CORT increases in response to infant crying and predicts alertness to such crying (Fleming et al., 2002). Also, CORT administration upregulates hippocampal responses to infant cry sounds (Bos et al., 2014), while activation of the hippocampus has been related to maternal bonding (Musser et al., 2012). Higher CORT in mothers during the postpartum period might thus facilitate arousal-dependent initiation of caregiving and stimulate bonding processes, through activation of the hippocampus and also potentially the amygdala. Both these structures are densely distributed with CORT receptors, and induced neural plasticity during the sensitive phase of pregnancy and early postpartum might also explain why the results found are long term (Bos, 2017). Moreover, such effects on the neural structures involved in motivation might explain why mothers with higher CORT hold more positive attitudes towards their own children (Fleming et al., 1997), spilling over into more deliberate and reflective behaviors. The interactions between endocrine and motivational responses becomes evident from research which shows that women who feel pressure to do well as mothers show higher stress levels (Henderson et al., 2016). Mothers who are more motivated to be consistently responsive and available caregivers may feel more alert and ready for action, and may be somewhat stressed, when interacting with their own child, but clearly more research is needed.

Furthermore, T was also related to maternal quality of caregiving, but only cross-sectionally during toddlerhood. A significant T by time interaction indicated that decreasing T during a mother-toddler interaction was associated with higher maternal caregiving quality, while increasing T was associated with lower maternal caregiving quality. As such, these results are in line with the higher T associations with lower maternal engagement (Clowtis et al., 2016), and add to the small body of research on T associations with caregiving quality in mothers.

Note that we did not find evidence for the dual-hormone hypothesis in mothers. Studies indeed suggest that the effect described in the dual-hormone hypothesis might be more pronounced in men compared to women (e.g., Welker et al., 2014; Dekkers et al., 2019), potentially due to differences in the hormonal systems that vary across women and men. For instance, T concentrations are considerably lower in women than in men (van Anders et al., 2015), as also can be seen in our study (see Table 1). Additionally, in women, T functions via conversion to estradiol, and as such, CORT may therefore interact more strongly with estradiol than with T. Indeed, Tackett et al. (2015) provided evidence that the dual-hormone hypothesis could be extended to estradiol and CORT interactions. Thus, our findings support the idea that the direction of the effects related to the dual-hormone hypothesis, at least with respect to CORT and T, differ across men and women.

For fathers, our results did provide evidence for the dual-hormone hypothesis in the direction expected (i.e., the association between T and behavior is more robust when CORT levels are low). First, a significant interaction was found between prenatal T and caregiving quality in toddlerhood on CORT. More specifically, the negative association between prenatal T and caregiving quality was stronger for fathers with low CORT. Similarly, the association between toddlerhood T and caregiving quality in toddlerhood was more negative for fathers with low CORT. These results provide support of the notion that effects of T are dependent on levels of CORT and underscore the need for more studies that incorporate both steroid hormones and their interaction when studying social behaviors including parenting.

Lastly, we found that lower CORT –measured six weeks postpartum when fathers were interacting with their 6-week-old infant– was associated with higher caregiving quality in toddlerhood. Generally, studies have shown a negative association between parenting stress and caregiving quality (Pelchat et al., 2003). The results of our study are in line with these findings and suggest that lower CORT in fathers, which may indicate lower stress during father-child interactions, are associated with higher caregiving quality, even after the first postpartum year. The question remains why this association is parent-specific, as mothers who showed higher CORT in response to the interaction with their 6-week-old infant showed higher caregiving quality when their child was 3 years of age. Although this is one of the few studies addressing sex-specific changes, and future comparative work should further explore these apparent differences between maternal and paternal neuroendocrine development, we can speculate about the causes of the observed sex-specific effect. One probable reason is that endocrine responses during pregnancy, as described above, differently prime neural circuits involved in caregiving. Studies indicate that mothers, but not fathers, show cortical thinning in regions involved in empathic responding directly after birth (Hoekzema et al., 2017). In fathers, structural neural changes are observed 3 months postpartum (Kim et al., 2014). This might align with the different role CORT has in maternal and paternal behavior, where fathers lack the positive effect of CORT-induced neural plasticity and increased caregiving motivation directly after birth (Hoekzema et al., 2017), but do show the same negative relation between caregiving and stress physiology later in life, similarly to mothers (Finegood et al., 2016). Our findings suggest that further investigation including third factors, such as parental neural changes during the transition to parenthood, and motivation and emotions experienced during parent-infant interactions, may help to unravel the differential directions found between postpartum hormones and toddlerhood caregiving quality in mothers and fathers.

This study has several strengths. First, by including both CORT and T, we were able to investigate their independent and dependent associations with caregiving quality. Also, most studies to date used self-report measures as index of caregiving quality, while the current study used multiple observations of parent-child interactions. Nevertheless, a number of limitations should also be borne in mind. The first limitation concerns the sample sizes of the group of mothers and fathers. Especially the size of the group of fathers was relatively small because fewer fathers than mothers were willing to participate. Though we used repeated endocrine measurements to increase power, it remains important that these findings are replicated. Related to the previous point, because almost 30% of the mothers participated in our study without a/their partner ($N = 18$ of the total sample of 64 mothers), mothers were always asked first to interact with the child, and the partner thereafter. This way, all mother-child interactions were comparable to one another. The downside of this decision is that the behavior and hormones during subsequent father-child interactions might have been affected by the infant's experiences during the mother-child interaction. Future research should either avoid this contagion possibility by randomly selecting the parent who goes first or by spreading parent-interactions across visits. However, it would also be of interest to study contagion in greater depth in larger samples and investigate whether and how behavior and hormones from one parent-child interaction affect the behavior and hormones of the other parent-child interaction. Another limitation of our study is that the sample mostly consisted of highly educated families, which limits the generalizability of our results. This is especially relevant given that secondary analyses of our study indicated that lower-educated fathers had higher CORT during the interaction with their toddler (see Table 3). Moreover, this study lacked information on maternal hormonal contraceptives use at 6 weeks postpartum and during toddlerhood. Including contraceptive use would have been relevant as there is evidence that T concentrations and responses may be suppressed in women on hormonal birth control (Voorthuis et al., 2019).

Finally, while during the prenatal and postpartum assessment rounds all mothers and fathers were assessed in the late afternoon and early evening to control for the circadian variation in CORT and T concentrations, the visits during toddlerhood took place in the morning and in the afternoon. This was because the main focus of this visit was to investigate children's executive functioning and we aimed to test children to the best of their abilities (for more information on the procedure and content of the visit, see Willemsen et al., 2021). Though we controlled for assessment time when investigating maternal and paternal toddlerhood CORT and T associations with caregiving quality, future research should aim to restrict assessment times to the second part of the day when cortisol concentrations are slowly decreasing (Adam et al., 2017).

Future research should take the role of breastfeeding status into account and time assessment times of maternal saliva samples in relation to mothers' nursing times. While in the current study there was little variation in breastfeeding status (i.e. >75% of mothers breastfed their 6-week-old infants while only one mother breastfed her 3-year-old child), breastfeeding might facilitate mother-infant cortisol attunement (Jonas et al., 2018) and reduces endocrine stress-responses (Heinrichs et al., 2002). As such, associations between CORT, T and caregiving quality may be more pronounced in more diverse samples in terms of breastfeeding status. Another important avenue for future research is to investigate why and how parents come to vary in their CORT and T concentrations and responses. Our study already indicated that higher-educated fathers had lower CORT during an interaction with their toddler (see Table 3). Other studies, for example, point in the direction of romantic relationship quality and adult attachment orientation as factors that are associated with individual differences in parental T concentrations during pregnancy and parent-infant interactions (Chin et al., 2020; Edelstein et al., 2019). Such investigations are important in unravelling the hormonal bases of human parental care.

To conclude, the current study built on the results of Bos et al. (2018) by examining in the same cohort whether parental perinatal measures of CORT and T forecast toddlerhood caregiving quality in both mothers and fathers. Importantly, this study indicates that CORT, T, and their interaction measured during pregnancy and the early postpartum period are associated with parental caregiving quality when the child is 3 years of age. Also, our results provide evidence for the dual-hormone hypothesis, but only in fathers. These results indicate that multiple endocrine measures need to be included in future investigations to better understand how hormones can explain individual differences in parental caregiving quality. Eventually, this research may help to detect parents at risk for low caregiving quality soon after or even before birth, which in turn opens a window for timely interventions.

Declaration of competing interest

None.

Acknowledgements

We thank all families who kindly participated in the BINGO study, and the PhD-students, research assistants and students for their assistance with data collection.

Source of funding

This research was supported by a Netherlands Organisation for Scientific Research VENI grant (016.195.197-to Beijers), a VICI grant (016. Vici.185.038-to de Weerth), a Jacobs Foundation Advanced Research Fellowship (to de Weerth), and a European Union's Horizon 2020 research and innovation program grant (728018-to de Weerth).

References

Adam, E.K., Quinn, M.E., Tavernier, R., McQuillan, M.T., Dahlke, K.A., Gilbert, K.E., 2017. Diurnal cortisol slopes and mental and physical health outcomes: a systematic review and meta-analysis. Psychoneuroendocrinology 83, 25–41. https://doi.org/10.1016/j.psyneuen.2017.05.018.

Ahnert, L., Deichmann, F., Bauer, M., Supper, B., Piskernik, B., 2021. Fathering behavior, attachment, and engagement in childcare predict testosterone and cortisol. Dev. Psychobiol. 63 (6), e22149. https://doi.org/10.1002/dev.22149.

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.

Bos, P.A., 2017. The endocrinology of human caregiving and its intergenerational transmission. Dev. Psychopathol. 29 (3), 971–999. https://doi.org/10.1017/S0954579416000973.

Bos, P.A., Montoya, E.R., Terburg, D., van Honk, J., 2014. Cortisol administration increases hippocampal activation to infant crying in males depending on childhood neglect. Hum. Brain Mapp. 35, 5116–5126.

Bos, P.A., Hechler, C., Beijers, R., Shinohara, K., Esposito, G., de Weerth, C., 2018. Prenatal and postnatal cortisol and testosterone are related to parental caregiving quality in fathers, but not in mothers. Psychoneuroendocrinology 97, 94–103. https://doi.org/10.1016/j.psyneuen.2018.07.013.

Burrous, C.E., Crockenberg, S.C., Leerkes, E.A., 2009. Developmental history of care and control, depression and anger: correlates of maternal sensivity in toddlerhood. Infant Ment. Health J. 30, 103–123. https://doi.org/10.1002/imhj.20206.

Calkins, S.D., Smith, C.L., Gill, K.L., Johnson, M.C., 1998. Maternal interactive style across contexts: relations to emotional, behavioral and physiological regulation during toddlerhood. Soc. Dev. 7, 350–369.

Carré, J.M., McCormick, C.M., Hariri, A.R., 2011. The social neuroendocrinology of human aggression. Psychoneuroendocrinology 36 (7), 935–944. https://doi.org/10.1016/j.psyneuen.2011.02.001.

Chin, K., Chopik, W.J., Wardecker, B.M., LaBelle, O.P., Moors, A.C., Edelstein, R.S., 2020. Longitudinal associations between prenatal testosterone and postpartum outcomes in a sample of first-time expectant lesbian couples. Horm. Behav. 125, 104810 https://doi.org/10.1016/j.yhbeh.2020.104810.

Clowtis, L.M., Kang, D.H., Padhye, N.S., Rozmus, C., Barratt, M.S., 2016. Biobehavioral factors in child health outcomes: the roles of maternal stress, maternal-child engagement, salivary cortisol, and salivary testosterone. Nurs. Res. 65 (5), 340–351.

Corpuz, R., D'Alessandro, S., Collom, G.K.S., 2020. The postnatal testosterone rebound in first-time fathers and the quality and quantity of paternal care. Dev. Psychobiol. 63, 1415–1427. https://doi.org/10.1002/dev.22064.

Crewther, B.T., Hecht, M., Cook, C.J., 2021. Diurnal within-person coupling between testosterone and cortisol in healthy men: evidence of positive and bidirectional time-lagged associations using a continuous-time model. Adapt. Hum. Behav. Psychiol. https://doi.org/10.1007/s40750-021-00162-8.

de Weerth, C., Buitelaar, J.K., 2005. Cortisol awakening response in pregnant women. Psychoneuroendocrinology 30 (9), 902–907. https://doi.org/10.1016/j.psyneuen.2005.05.003.

Dekkers, T.J., van Rentergem, J.A.A., Meijer, B., Popma, A., Wagemaker, E., Huizenga, H.M., 2019. A meta-analytical evaluation of the dual-hormone hypothesis: does cortisol moderate the relationship between testosterone and status, dominance, risk taking, aggression, and psychopathy? Neurosci. Biobehav. Rev. 96, 250–271. https://doi.org/10.1016/j.neubiorev.2018.12.004.

Dickerson, S.S., Kemeny, M.E., 2004. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychol. Bull. 130, 355–391. https://doi.org/10.1037/0033-2909.130.3.355.01482.

Dörr, H.G., Heller, A., Versmold, H.T., Sippell, W.G., Herrmann, M., Bidlingmaier, F., Knorr, D., 1989. Longitudinal study of progestins, mineralocorticoids, and glucocorticoids throughout human pregnancy. J. Clin. Endocrinol. Metab. 68 (5), 863–868. https://doi.org/10.1210/jcem-68-5-863.

Edelstein, R.S., Chin, K., Saini, E.K., Kuo, P.X., Schultheiss, O.C., Volling, B.L., 2019. Adult attachment and testosterone reactivity: Fathers' avoidance predicts changes in testosterone during the strange situation procedure. Horm. Behav. 112, 10–19. https://doi.org/10.1016/j.yhbeh.2019.01.009.

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. https://doi.org/10.1016/j.yhbeh.2016.01.016.

Erickson, M.F., Sroufe, L.A., Egeland, B., 1985. The relationship between quality of attachment and behavior problems in preschool in a high-risk sample. Monogr. Soc. Res. Child Dev. 50 (1–2), 147–166.

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 (6), 1774–1791. https://doi.org/10.1111/j.1467-8624.2004.00816.x.

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 (7), 1073–1087. https://doi.org/10.1037/dev0000123.

Fleming, A.S., Ruble, D., Krieger, H., Wong, P.Y., 1997. Hormonal and experiential correlates of maternal responsiveness during pregnancy and the puerperium in human mothers. Horm. Behav. 31, 145–158.

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 (39), 16194–16199. https://doi.org/10.1073/pnas.1105403108.

Gilbert, R., Widom, C.S., Browne, K., Fergusson, D., Webb, E., Janson, S., 2009. Burden and consequences of child maltreatment in high-income countries. Lancet (London, England) 373 (9657), 68–81. https://doi.org/10.1016/S0140-6736(08)61706-7.

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. https://doi.org/10.1016/j.yhbeh.2017.05.016.

Grebe, N.M., Sarafin, R.E., Strenth, C.R., Zilioli, S., 2019. Pair-bonding, fatherhood, and the role of testosterone: a meta-analytic review. Neurosci. Biobehav. Rev. 98, 221–233. https://doi.org/10.1016/j.neubiorev.2019.01.010.

Gunnar, M.R., White, B.P., 2001. Salivary cortisol measures in infant and child assessment. In: Singer, L.T., Zeskind, P.S. (Eds.), Biobehavioral Assessment of the Infant. Guilford Press, New York, pp. 167–189.

Harden, K.P., Wrzus, C., Luong, G., Grotzinger, A., Bajbouj, M., Rauers, A., Wagner, G.G., Riediger, M., 2016. Diurnal coupling between testosterone and cortisol from adolescence to older adulthood. Psychoneuroendocrinology 73, 79–90.

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

Henderson, A., Harmon, S., Newman, H., 2016. The price mothers pay, even when they are not buying it: mental health consequences of idealized motherhood. Sex Roles 512–526. https://doi.org/10.1007/s11199-015-0534-5.

Hermans, E.J., Putman, P., Van Honk, J.J.P., 2006. Testosterone administration reduces empathetic behavior: a facial mimicry study. Psychoneuroendocrinology 31 (7), 859–866. https://doi.org/10.1016/j.psyneuen.2006.04.002.

Hoekzema, E., Barba-Müller, E., Pozzobon, C., Picado, M., Lucco, F., García-García, D., Soliva, J.C., Vilarroya, O., 2017. Pregnancy leads to long-lasting changes in human brain structure. Nature Neurosci. 20, 287–296. https://doi.org/10.1038/nn.4458.

Jonas, W., Bisceglia, R., Meaney, M.J., Dudin, A., Fleming, A.S., Steiner, M., 2018. The role of breastfeeding in the association between maternal and infant cortisol attunement in the first postpartum year. Acta Paediatr. 107, 1205–1217. https://doi.org/10.1111/apa.14256.

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 (2), 107–115. https://doi.org/10.1192/bjp.bp.111.093062.

Kim, P., Rigo, P., Mayes, L.C., Feldman, R., Leckman, J.F., Swain, J.E., 2014. Neural plasticity in fathers of human infants. Soc. Neurosci. 9, 522–535. https://doi.org/10.1080/17470919.2014.933713.

Knight, E.L., Sarkar, A., Prasad, S., Mehta, P.H., 2020. Beyond the challenge hypothesis: the emergence of the dual-hormone hypothesis and recommendations for future research. Horm. Behav. 123 https://doi.org/10.1016/j.yhbeh.2019.104657.

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 (3), 303–314. https://doi.org/10.1002/dev.21370.

Kuo, P.X., Braungart-Rieker, J.M., Burke Lefever, J.E., Sarma, M.S., O'Neill, M., Gettler, L.T., 2018. Fathers' cortisol and testosterone in the days around infants' births predict later paternal involvement. Horm. Behav. 106, 28–34. https://doi.org/10.1016/j.yhbeh.2018.08.011.

Lotz, A.M., Buisman, R.S.M., Alyousefi-van Dijk, K., Witte, A.M., Bakermans-Kranenburg, M.J., Verhees, M.W.F.T., 2022. Exploring the role of endocrine factors in sensitive parenting in men. Horm. Behav. 140 https://doi.org/10.1016/j.yhbeh.2022.105118.

Marceau, K., Ruttle, P.L., Shirtcliff, E.A., Hastings, P.D., Klimes-Dougan, B., Zahn-Waxler, C., 2015. Within-person coupling of changes in cortisol, testosterone, and DHEA across the day in adolescents. Dev. Psychobiol. 57, 654–669.

Mehta, P.H., Josephs, R.A., 2010. Testosterone and cortisol jointly regulate dominance: evidence for a dual-hormone hypothesis. Horm. Behav. 58 (5), 898–906. https://doi.org/10.1016/j.yhbeh.2010.08.020.

Mehta, P.H., Prasad, S., 2015. The dual-hormone hypothesis: a brief review and future research agenda. Curr. Opin. Behav. Sci. 3, 163–168. https://doi.org/10.1016/j.cobeha.2015.04.008.

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. Developmental Psychobioly 51 (8), 650–661. https://doi.org/10.1002/dev.20400.

Musser, E.D., Kaiser-Laurent, H., Ablow, J.C., 2012. The neural correlates of maternal sensitivity: an fMRI study. Dev. Cognit. Neurosci. 2, 428–436.

Noyola-Martínez, N., Halhali, A., Barrera, D., 2019. Steroid hormones and pregnancy. Gynecol. Endocrinol. 35 (5), 376–384. https://doi.org/10.1080/09513590.2018.1564742.

Pechtel, P., Pizzagalli, D.A., 2011. Effects of early life stress on cognitive and affective function: an integrated review of human literature. Psychopharmacology 214 (1), 55–70. https://doi.org/10.1007/s00213-010-2009-2.

Pelchat, D., Bisson, J., Bois, C., Saucier, J.F., 2003. The effects of early relational antecedents and other factors on the parental sensitivity of mothers and fathers. Infant Child Dev. 12 (1), 27–51. https://doi.org/10.1002/icd.335.

Perini, T., Ditzen, B., Hengartner, M., Ehlert, U., 2012b. Sensation seeking in fathers: the impact on testosterone and paternal investment. Horm. Behav. 61, 191–195. https://doi.org/10.1016/j.yhbeh.2011.12.004.

Rodriguez, C.M., Granger, D.A., Leerkes, E.M., 2020. Testosterone associations with parents' child abuse risk and at-risk parenting: a multimethod longitudinal examination. Child Maltreat. 26, 50–62. https://doi.org/10.1177/1077559520930819.

Rosenbaum, S., Gettler, L.T., McDade, T.W., Bechayda, S.S., Kuzawa, C.W., 2018. Does a man's testosterone "rebound" as dependent children grow up, or when pairbonds end? A test in Cebu, Philippines. Am. J. Hum. Biol. 30 (6), e23180 https://doi.org/10.1002/ajhb.23180.

Saltzman, W., Maestripieri, D., 2011. The neuroendocrinology of primate maternal behavior. Prog. Neuro-Psychopharmacol. Biol. Psychiatry 35, 1192–1204. https://doi.org/10.1016/j.pnpbp.2010.09.017.

Smeekens, S., Riksen-Walraven, J.M., van Bakel, H.J., 2007. Multiple determinants of externalizing behavior in 5-year-olds: a longitudinal model. J. Abnorm. Child Psychol. 35 (3), 347–361. https://doi.org/10.1007/s10802-006-9095-y.

Sroufe, L.A., Egeland, B., Carlson, E.A., Collins, W.A., 2005. The Development of the Person: The Minnesota Study of Risk and Adaptation From Birth to Adulthood. Guilford Press, New York.

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 Sci. Pract. 1 (1–2), 71–100. https://doi.org/10.1207/S15327922PAR011&2_5.

Storey, A.E., Ziegler, T.E., 2016. Primate paternal care: interactions between biology and social experience. Horm. Behav. 77, 260–271. https://doi.org/10.1016/j.yhbeh.2015.07.024.

Storey, A.E., Alloway, H., Walsh, C.J., 2020. Dads: Progress in understanding the neuroendocrine basis of human fathering behavior. Horm. Behav. 119 https://doi.org/10.1016/j.yhbeh.2019.104660.

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

Tackett, J.L., Reardon, K.W., Herzhoff, K., Page-Gould, E., Harden, K.P., Josephs, R.A., 2015. Estradiol and cortisol interactions in youth externalizing psychopathology. Psychoneuroendocrinology 55, 146–153. https://doi.org/10.1016/j.psyneuen.2015.02.014.

Thompson, L.A., Trevathan, W.R., 2008. Cortisol reactivity, maternal sensitivity, and learning in 3-month-old infants. Infant Behav. Dev. 31 (1), 92–106. https://doi.org/10.1016/j.infbeh.2007.07.007.

Tukey, J.W., 1962. The future of data analysis. Ann. Math. Stat. 33 (1), 1–67. https://doi.org/10.1214/aoms/1177704711. JSTOR 2237638.

van Anders, S.M., Steiger, J., Goldey, K.L., 2015. Effects of gendered behavior on testosterone in women and men. Proc. Natl. Acad. Sci. U. S. A. 112 (45), 13805–13810. https://doi.org/10.1073/pnas.1509591112.

Vesely, C.K., Brown, E.L., Mahatmya, D., 2013. It takes two: sensitive caregiving across contexts and children's social, emotional, and academic outcomes. Early Educ. Dev. 24 (7), 960–978. https://doi.org/10.1080/10409289.2013.825185.

Vogt, W.P., 2005. Dictionary of Statistics & Methodology: A Nontechnical Guide for the Social Sciences. SAGE.

Voorthuis, A., Bakermans-Kranenburg, M.J., van IJzendoorn, M.H., 2019. Testosterone reactivity to infant crying and caregiving in women: the role of oral contraceptives and basal cortisol. Infant Behav. Dev. 56 https://doi.org/10.1016/j.infbeh.2017.08.002.

Weisman, O., Zagoory-Sharon, O., Feldman, R., 2013. Oxytocin administration alters HPA reactivity in the context of parent-infant interaction. Eur. Neuropsychopharmacol. 23, 1724–1731. https://doi.org/10.1016/j.euroneuro.2013.06.006.

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. https://doi.org/10.1016/j.physbeh.2014.02.057.

Willemsen, Y., Beijers, R., Vasquez, A.A., de Weerth, C., 2021. Do breastfeeding history and diet quality predict inhibitory control at preschool age? Nutrients 13. https://doi.org/10.3390/nu13082752.