Received: 28 July 2020 | Revised: 25 September 2020 | Accepted: 10 October 2020

DOI: 10.1002/dev.22055

SPECIAL ISSUE

Developmental Psychobiology WILEY

Exploring the hormonal and neural correlates of paternal protective behavior to their infants

Anna M. Lotz[1,2] | Martine W. F. T. Verhees[1] | Lisa I. Horstman[1,2] | Madelon M. E. Riem[1,4] | Marinus H. van IJzendoorn[3] | Marian J. Bakermans-Kranenburg[1,2] | Renate S. M. Buisman[1]

[1]Clinical Child & Family Studies, Faculty of Behavioral and Movement Sciences, Vrije Universiteit, Amsterdam, the Netherlands

[2]Leiden Institute for Brain and Cognition, Leiden University Medical Center, Leiden, the Netherlands

[3]Department of Psychology, Education and Child Studies, Erasmus University, Rotterdam, The Netherlands

[4]Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

Correspondence
Anna M. Lotz, Clinical Child & Family Studies, Faculty of Behavioral and Movement Sciences, Vrije Universiteit, Amsterdam, the Netherlands.
Email: a.m.lotz@vu.nl

Funding Information
This study was supported by a European Research Council grant (grant number AdG 669249, 2015) awarded to MJB-K and a Spinoza prize (Netherlands Organization for Scientific Research) awarded to MHvIJ.

Abstract

Infant protection is an important but largely neglected aspect of parental care. Available theory and research suggest that endocrine levels and neural responses might be biological correlates of protective behavior. However, no research to date examined associations between these neurobiological and behavioral aspects. This study, preregistered on https://osf.io/2acxd, explored the psychobiology of paternal protection in 77 new fathers by combining neural responses to infant-threatening situations, self-reported protective behavior, behavioral observations in a newly developed experimental set-up (Auditory Startling Task), and measurements of testosterone and vasopressin. fMRI analyses validated the role of several brain networks in the processing of infant-threatening situations and indicated replicable findings with the infant-threat paradigm. We found little overlap between observed and reported protective behavior. Robust associations between endocrine levels, neural responses, and paternal protective behavior were absent.

KEYWORDS
brain imaging, fathers, hormones, parental care, protection

1 | INTRODUCTION

An important aspect of parental care, but largely neglected in parenting research, is infant protection (Bakermans-Kranenburg & Van IJzendoorn, 2017). Parental protection of infants from diseases, accidents, and stranger violence is vital for infant survival during pregnancy and after birth (Hahn-Holbrook et al., 2011). In addition, experiences with protective parents may enhance children's trust in parent's availability to help and protect them in times of danger or distress, promoting secure attachment (Bowbly, 1982). Despite the clear importance of parental protection by both parents, protective behavior, and its neurobiology has been mainly studied in mothers (Hahn-Holbrook et al., 2011). Since fathers have an increasing role in childcare in modern Western societies (Bakermans-Kranenburg et al., 2019), research into fathers' caregiving and paternal protection is timely. For this reason, this study aims to explore the psychobiology of protective behavior in fathers.

Studies in both biparental mammals and humans have examined the psychobiology of paternal protective behavior in a variety of paradigms. In mammals, for example, pup retrieval and attacking intruders as examples of parental protection have been studied (Abraham & Feldman, 2018; Wynne-Edwards & Timonin, 2007). The neural basis of this behavior might be linked to the endocrine system, with an important role for steroids

and neuropeptides (reviewed by Bales & Saltzman, 2016, but see Wynne-Edwards & Timonin, 2007 for a critical note on the suggestion of a causal link between behavior and the endocrine system). In humans, men's behavioral, hormonal, and neural responses have been studied during exposure to infant crying, an indicator for a possible threatening situation (e.g., Alyousefi-Van Dijk et al., 2019; Khoddam et al., 2020; Li et al., 2017; Thijssen et al., 2018; Van Anders et al., 2012), during exposure to video fragments of infant-threatening situations, and via self-report on daily life situations (e.g., Van't Veer et al., 2019). Several literature reviews described a neuro-endocrine basis for paternal behavior (Abraham & Feldman, 2018; Hahn-Holbrook et al., 2011; Rilling & Mascaro, 2017), indicating that specific hormones such as testosterone (T) and vasopressin (AVP), and specific neural activations might be proper correlates or even activators of observed protective behavior. However, only a few studies combined the various aspects of the psychobiology of human paternal protective behavior, and no studies to date integrated all three aspects (hormonal, neural, and behavioral). This study aims to fill this gap by combining neural responses to infant threat, behavioral measures of protective behavior in daily life and in an experimental set-up, and measurements of T and AVP.

The potential roles of T and AVP in protective parenting have been incorporated in the Steroid/Peptide Theory of Social Bonds (Van Anders et al., 2011). This model is mostly based on nonhuman mammalian literature, although it includes human research as well. First, the model implies that low levels of T might be linked to parental contexts that are perceived as nurturing. This idea is in line with previous research in humans showing that lower salivary T is associated with increased participation of fathers in child care and enhanced quality of caregiving, both prenatally (Bos et al., 2018; Edelstein et al., 2017) and postnatally (Bos et al., 2018; Gettler et al., 2011; Weisman et al., 2014), although combined effect sizes are small (for a meta-analysis see Meijer et al., 2019). Additionally, the Steroid/Peptide Theory of Social Bonds implies that high levels of T are associated with parental contexts that involve a need for protective responses. Indeed, research has shown that cry sounds increase salivary T levels in men when no nurturing action is possible (Van Anders et al., 2012). Additionally, T levels in fathers-to-be are positively associated with neural activation in brain areas involved in social cognition, arousal, and reward learning when listening to infant cry sounds (Khoddam et al., 2020), although it should be noted that another study did not reveal such a relation in fathers of infants between 1 and 2 years old (Mascaro et al., 2014). Moreover, T administration increases men's neural responses to facial threat cues in brain areas associated with threat processing (Goetz et al., 2014). On the basis of these findings, we expect that baseline T levels are positively associated with protective behavior and neural reactivity to infant-threatening stimuli.

Second, the Steroid/Peptide Theory of Social Bonds assigns a specific role to AVP reactivity in protective parenting, with increases of AVP positively related to protective aggression (Van Anders et al., 2011). A number of studies suggest that paternal protection might be associated with AVP. For instance, it has been shown that in father-to-be's administration of AVP increases orientation toward baby avatars (Cohen-Bendahan et al., 2015), increases excessive handgrip force while looking at an image of an unknown infant compared to an image of their own infant (Alyousefi-Van Dijk et al., 2019), and increases activation in several brain areas in response to emotionally versus neutrally labeled infant cry sounds (Thijssen et al., 2018). Moreover, basal AVP levels in fathers have been shown to be negatively related to neural activity in brain areas involved in empathy and social cognition when viewing neutral or positive videos of own infant versus other infants (Atzil et al., 2012). The authors interpreted this as a possible AVP-dependent vigilance toward strangers. In contrast, other studies showed that AVP administration did not increase fathers' neural processing of infant cry and the subjective cry rating (Li et al., 2017), and no correlation between basal AVP levels and explicit and implicit infant caregiving was observed in prospective fathers (Cohen-Bendahan et al., 2015). Thus, mixed findings for both basal and reactive AVP levels in protective behavior have been documented. However, no studies to date specifically looked at AVP in relation to protective paternal behavior and neural processing of infant threat. Based on previous findings and the Steroid/Peptide Theory, it could be predicted that higher basal levels of AVP are associated with more paternal protective behavior and stronger neural responses to infant threat.

As mentioned above, very little research has examined the relation between brain responses and behavior in the context of protective behavior. To our knowledge, only one study to date has focused on the relation between paternal protective behavior and its neural correlates (Van 't Veer et al., 2019). In that study, on a different sample than this study, several brain networks known to be associated with the parental care network, visual processing, and threat detection were shown to be involved in the processing of videos depicting infant-threatening situations. Moreover, father's reported protective behavior in daily life was linked to stronger brain activation in the frontal pole while watching their own (versus an unknown) infant in threatening (versus neutral) situations. Based on these findings, activation of a neural threat component might be positively associated with observed and reported protective behavior.

This study explores the psychobiological correlates of paternal protection. To this end, paternal protective responses were measured using behavioral observations during the exposure to a loud/alarming sound in a lab setting, self-reported protective behavior, and the neural processing of videos depicting infant-threatening situations. Moreover, basal salivary T and AVP levels were determined, and relations between neural, hormonal, and behavioral measures were examined. We hypothesized that these three measures would be positively related, for example, higher neural responses to infant threat in brain areas involved in the parental care network, visual processing areas and threat detection would be associated with higher basal hormone levels and more observed and self-reported protective behavior (see Figure 1a). Moreover, based on the

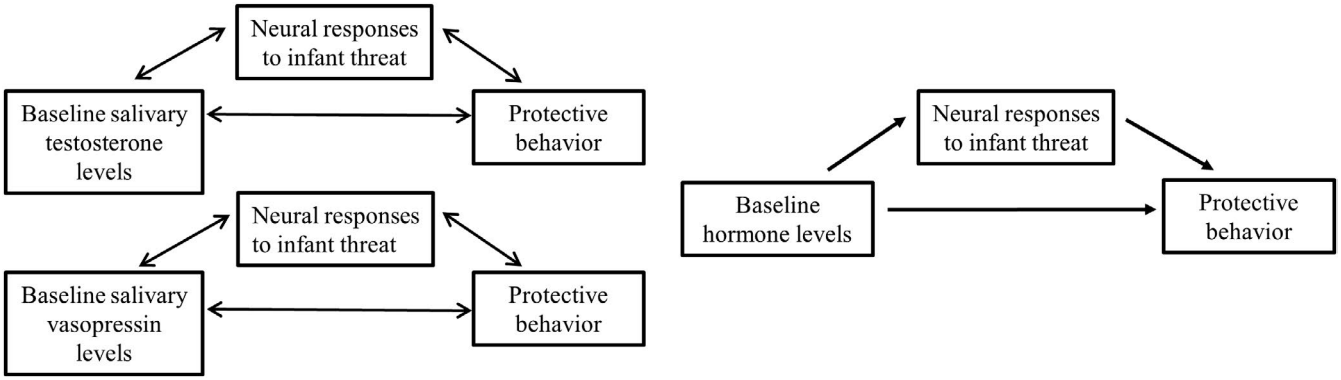

FIGURE 1 Visual overview of working hypotheses. Note. (a) The planned associations between the three dependent variables. Protective behavior represented by either observed or reported protective behavior. (b) Neural responses as a mediator in the relation between baseline hormone levels and protective behavior

previously reported hormone administration effects on neural processing (e.g., Goetz et al., 2014; Thijssen et al., 2018), we speculated that the associations between T and AVP levels and protective behavior would be mediated by neural responses to infant threat in brain areas involved in the parental care network, visual processing, and threat detection (see Figure 1b).

2 | METHODS

2.1 | Participants

Seventy-seven first-time fathers participated in this study. Participants were recruited via municipal records, infant welfare centers, midwife practices, and social media (see Figure S1 for an overview of the recruitment). To maximize sample size for the current analyses, we decided to deviate from a priori stated inclusion criteria in nine cases (MRI contraindication $n = 5$, Cardiovascular disease $n = 1$, Use of medication potentially interfering with the endocrine system $n = 1$, Birth < 37 weeks $n = 1$, Not biological father $n = 1$). Participants who were not eligible to undergo an fMRI scan, for example due to diabetes or metallic foreign objects in body, were invited for a research visit at the participant's home. All but one participant were the biological father of the child. All participants cohabited with the biological mother of the child. Moreover, participants were in good mental and physical health (i.e., had no psychiatric, neuroendocrine, or neurological diagnosis, and no upper torso injury that could affect the use of a baby carrier), except for four who reported a psychiatric, neuroendocrine, or cardio-vascular diagnosis, or had taken medication on the assessment day that could interfere with the endocrine system. Participants reported no significant intake of alcohol or drugs at the time of inclusion. Participants spoke Dutch fluently and reported not to use a baby carrier over 5 hr per week at time of inclusion, which was relevant for various research questions. All children, except for one (born after 36 weeks and 6 days), were born full-term (i.e., born after 37 week gestation) and all were in good health. See Table 1 for sample characteristics. The study was approved by the Ethics Committees of the Leiden University Medical Centre and of the Department of Education and Child studies at Leiden University. The study was carried out in accordance with the declaration of Helsinki and all participants gave written informed consent.

2.2 | Procedure

The research visit was located either in the Leiden University Medical Centre or at the participant's home, depending on possible MRI contraindications (e.g., diabetes, metallic foreign objects in participant's body) and preference of the participant. The assessment consisted of the following measures: Behavioral measurements, including a 10-min Free play (Witte et al., 2019), a handgrip force paradigm (Alyousefi-Van Dijk et al., 2019), and the Five Minute Speech Sample (Lotz et al., 2020); Saliva and hair sampling for hormonal measurements; and questionnaires about, for example, health, medication, and current mental state. When the assessment took place in the LUMC, the Auditory Startling Task (AST, Witte et al., 2019) was performed and neural measurements were conducted with (f)MRI and DTI. Following the assessments, participants and partners completed some online questionnaires at home (including the Parental Protection Questionnaire, PPQ; Van 't Veer et al., 2019).

2.3 | Measures

2.3.1 | Protective behavior

Auditory startling task

To observe paternal protective behavior, we applied the AST (Witte et al., 2019). During the AST, a short loud sound fragment was played "unexpectedly" by a hidden audio installation while the participant was videotaped playing with his infant. The sound consisted of white noise (80 dB) and was programmed for 10 s with short breaks. At the end of the sound fragment, the researcher entered the room and apologized for the sound, referring to technical problems (the purpose of the sound fragment was explained to the participants at the end of the study). Paternal protective behavior was coded from

TABLE 1 Sample characteristics

	M(*SD*)/*N*(%)	Range
Participant age (years, *N* = 77)	33.14 (5.39)	25.06–56.50
Education (years past primary education, *N* = 76)	8.25 (1.86)	3.00–10.00
Country of birth (*N* = 76)		
The Netherlands	70 (92%)	
Other	6 (8%)	
Race (*N* = 77)		
Caucasian	71 (92%)	
Other	6 (8%)	
Infant age (weeks, *N* = 77)	11.40 (3.10)	7.57–21.43
Infant sex (*N* = 77)		
Male	41 (53%)	
Female	36 (47%)	
Testosterone (pg/ml, *N* = 75)	43.17 (35.15)	2.14–216.96
Vasopressin (pg/ml, *N* = 75)	1.64 (0.46)	0.87–2.67
AST (*N* = 59)	5.56 (1.70)	3.00–10.00
PPQ (*N* = 72)	3.46 (0.39)	2.29–4.00
fMRI cluster 1 (*N* = 64)	0.67 (0.96)	−2.15-2.53
fMRI cluster 2 (*N* = 64)	0.76 (0.97)	−1.98-2.90
fMRI cluster 3 (*N* = 64)	1.02 (0.95)	−1.04-2.94
fMRI cluster 4 (*N* = 64)	0.89 (0.78)	−0.79-2.62

Note: Sample characteristics are calculated based on non-transformed complete cases data. Participants' race was coded by researchers based on videotapes of fathers and was defined as Caucasoid type or other. AST: protective behavior observed during the Auditory Startling task. PPQ: self-reported protective behavior measured with the Paternal Protection Questionnaire. fMRI cluster 1–4 are individual mean *z*-values based on the contrast threat > neutral. Cluster 1: left cuneal cortex, left lateral occipital cortex and left occipital pole. Cluster 2: right cuneal cortex, right lateral occipital cortex and right occipital pole. Cluster 3: left supramarginal gyrus, planum temporale, lateral occipital cortex, central opercular cortex, middle temporal cortex, parietal operculum cortex. Cluster 4: insular cortex, frontal orbital cortex, temporal pole, superior frontal gyrus, frontal operculum cortex, central operculum cortex, thalamus, putamen, precentral gyrus, right planum temporale, right supramarginal gyrus, right middle temporal gyrus, right inferior temporal cortex, right superior temporal gyrus, anterior cingulate gyrus, amygdala.

the video, as well as baby states 30 s before and during the sound fragment. The coding system was developed for this study. The coding scheme for protective behavior consists of a 10-point scale, with higher scores reflecting more protective behavior, see Table S1. The coding scheme for baby states before and during noise was based on a 5-point scale (Mah et al., 2015), with one additional scale point: 1 = Drowsy or asleep, 2 = Quiet, 3 = Alert, 4 = Startled, 5 = Fussing, 6 = Distressed. When various states or responses were observed, the highest rating was assigned. Five raters were trained by one of the authors (AL). Interrater reliability was assessed based on a total of 20 videos. A first set of 10 videos was scored directly after the training. A second set of ten videos was coded after all raters coded several videos independently. All raters obtained good interrater reliability, ICC (single measure, absolute agreement) for Paternal protective behavior >.84, Baby state before noise ICC >.67, Baby state during noise ICC >.75).

Parental Protection Questionnaire

To measure father's protective behavior toward the child in daily life, fathers were asked to fill in the PPQ (see Van 't Veer et al. (2019) for a prenatal version of this questionnaire) after the assessment. The partners reported on the participants' protective behavior as well. The questionnaire contains 12 items that were scored on the prevalence during the past month (1 = never, 2 = seldom, 3 = sometimes, 4 = often or always, 5 = not applicable). Examples of questions are: "I made sure that the baby's diaper was changed in time," or "I made sure that the living room was safe for the baby." Scores on items that were coded 5 were replaced with individual mean scores calculated over all other items. A one-factor exploratory factor analysis on participants' data revealed that the first factor explained 24.4% of the variance and seven items loaded> 0.40. These seven items ($\alpha_{participant}$ = .67 and $\alpha_{partner}$ = .80) were used to calculate mean scores for self-reported and partner-reported protective behavior.

Latent construct "Protective behavior"

Associations between participants' observed protective behavior and reported protective behavior (partner- and self-report) were examined in the full cases dataset, and indicated no underling latent construct for 'Protective behavior (range r = −.04 to .19, p > .16). Additionally, we explored a possible multi-informant component for self-reported protective behavior (Kraemer et al., 2003), combining both partner and participant PPQ scores on the seven EFA items. A PCA with varimax rotation did not reveal the presence of a multi-informant component. Observed and self-reported protective behavior were, therefore, separately examined in subsequent analyses.

2.3.2 | Neural responses to infant threat

fMRI infant threat paradigm

To examine the possible neural basis for paternal protection, fathers participated in an fMRI task adapted from Van 't Veer et al. (2019). In the current protection task, as described in Witte et al. (2019), participants watched videos while lying in an MRI scanner. The videos depicted either a scenario in which an infant was in a threatening situation or a matched neutral video in which there was no threat to the infant. During the task, participants watched 12 different video pairs (see Table S2 for a description of the videos, videos available upon request). The video fragments were filmed using a gender-neutral lifelike baby doll. Fathers were instructed to imagine that the infant in the videos was their own infant. To ease the task of imagining their own infant in the videos, a picture of their infant was shown before the task. Moreover, the visibility of the doll's face and the faces of the actors in the videos was minimized.

Each video, 12 threatening video fragments and 12 matched neutral video fragments, was shown twice, resulting in a total of 48 videos. The duration of each video fragment was 6 s. The videos were shown in one of four preprogrammed semi-random orders. For each order, it was ensured that videos were equally distributed across the task: 12 neutral and 12 threatening videos were presented during the first half of the task as well as the second half of the task. To maximize the power of the design, interstimulus interval (ISI) between videos was sampled using the web-based tool Neurodesign (Durnez et al., 2017). Videos were separated by an ISI of variable length ranging from 3.0 to 8.0 s, with a mean ISI of 4.5 s. The task was programmed in E-Prime software (version 2.0; Psychology Software Tools, Inc.).

Prior to the assessment, participants were asked to send a neutral picture of their own child. To ensure that each photo contained only the facial features of the infant, each photo was edited in Adobe Photoshop CS by adding a black-face contour to remove ears and most of the hair, and a black background. To make sure that all images contained approximately the same pixel ratio between the black background and face, images were resized so that face length was set on 10 cm. Finally, a selection of the picture containing the face was copied to a new black image of 640 × 480 pixels. Participants were familiarized with the edited image and task design during a practice task prior to the MRI scan session with two pairs of neutral and threatening videos that were not included in the real task. At the onset of the real task, an edited picture of participants' own infant was shown with a written instruction to imagine that their own infant is displayed in the succeeding videos. The instruction and edited picture were shown again after each eight videos (thus six times in total). After a stimulus interval of 250 ms, the instruction screen advanced to one of the four preprogrammed semi-random order of 48 videos as described above.

fMRI Parameters

MRI scanning was performed on a 3 T Philips Achieva TXMRI system (Philips Medical Systems, Best). For registration purposes, a T1-weighted anatomical scan was acquired (repetition time (TR) = 7.90 ms, echo time (TE) = 3.50 ms, flip angle = 8°, 155 transverse slices, ACQ voxel size 1.1 × 1.1 × 1.1 mm). The fMRI-task utilized a gradient-echo blood oxygen level-dependent (BOLD) echo-planar imaging sequence with: TR = 2,200 ms, TE = 30 ms, flip angle = 80°, 38 transverse slices, and ACQ voxel size of 2.75 × 2.75 × 2.75 mm (including a default interslice gap). The duration of the fMRI paradigm was 10 min 53 s (290 volumes). Fieldmap corrections were performed using a multi-acquisition B0 map, performed directly after the end of the fMRI protection task.

fMRI Data Analyses

Data were structured following the brain imaging data structure (BIDS) (Gorgolewski et al., 2016) and preprocessed using *fMRIPrep* (version 1.5.2; Esteban et al. (2018); RRID:SCR_016216)), which is based on *Nipype* 1.3.1 (Gorgolewski et al. 2011;RRID:SCR_002502), see supplementary material for details. Furthermore, in the FEAT module (Smith et al., 2004) of FMRI Expert Analysis Tool (FSL), spatial smoothing was performed using a Gaussian kernel with a full-width-at-half-maximum of 5 mm. High-pass filter cutoff was set at 90 s. First-level analyses of the participant for the contrast-threatening videos > neutral videos and the reverse contrast-threatening videos < neutral videos were performed using FEAT. Threatening and neutral videos as well as the instruction picture of own child were modeled separately as a square-wave function. Each predictor was convolved with a double gamma hemodynamic response function and temporal derivatives were added to the model, resulting in six regressors in each model. All first-level contrast images and the corresponding variance images were submitted to second level mixed-effects group whole-brain analyses. Group-level analysis was performed using FEAT to detect average activation for the contrast threat > neutral and the reverse contrast. Statistical maps were thresholded using clusters determined by Z> 3.1 and a cluster corrected significance threshold of $p < .05$. Masks were created for each significant cluster resulting from the threat > neutral contrast and used as input for a featquery extracting the individual mean *z*-value values for each significant cluster. These individual *z*-values were used for further confirmatory analyses. Visual inspection of motion parameters flagged one participant with head motion between 1.4 and 3.0 mm, and sensitivity analysis for fMRI higher level analysis was performed.

Salivary testosterone and vasopressin

Salivary samples were obtained at the beginning of each assessment. Most assessments took place in the late afternoon, that is, after 15:00, or early evening (67%), however, visits were also scheduled during the morning (18%) or early afternoon (15%) (M = 15:38 hr, SD = 3.05, range = 09:15–19:39). Participants were instructed not to consume any alcoholic drinks or have excessive physical exercise during the last 24 hr prior to each assessment. Furthermore, participants were asked not to drink any caffeine-containing drinks on the day of the assessment, and not to smoke, chew gum, eat or drink (other than water) 30 min before the start of the appointment. Just before sampling participants rinsed their mouth with water. All samples were stored as soon as possible at −20°C. Saliva samples collected at home were placed on ice for transportation. Samples of one participant were excluded from analyses since they were not frozen after collection.

Testosterone

Participants drool approximately 1.5 ml saliva into a 2 ml cryogenic vial (SalivaBio, Salimetrics), either directly into the vial or indirectly using a saliva collection aid (Saliva Bio, Salimetrics). Before drooling, participants were asked to swallow once and to bent slightly forward. To stimulate saliva production, participants were recommended to think about something sweet or sour, to move their jaws up and down, or to look at pictures of food that were provided. On average, it took 9.16 (SD = 5.64) min to complete the collection. T was quantified at Dresden LabService GmbH (Germany) by Luminscence enzyme immunoassay (IBL International GMBH), see Alvergne et al. (2009) for a description of the procedure. Fifty microliter of each sample

was used for the immunoassay. The limit of detection was 1.8 pg/ml. Samples used for this article were quantified in the same assay run (12 plates in total). Duplicate analysis was performed for a random selection of 30% of the samples. For these samples, an average value was calculated and used in further analyses. Inter-assay variability was 10% and intra-assay variability was 5.22%. Cross-reactivity with other substances tested ≤0.01%. T values are reported in pg/ml.

Vasopressin

Participants chewed on a cotton swab (Salivette, Sarstedt) for 60 s. Saliva for the quantification of AVP was collected using salivettes (Sarstedt). Participants were instructed to chew lightly on a cotton swab for 60 s and move the swab around in their mouth every now and then to boost saliva collection. Researchers kept track of the time using a stopwatch. AVP levels were quantified by radioimunnoassay at RIAgnosis (Sinzing, Germany), as previously described in Frijling et al. (2015) and Kagerbauer et al. (2019). Salivettes were centrifuged at 4°C for 30 min with ca. 5,000 ***g*** centrifugal force, after which 0.3 ml of saliva was pipetted into a vial for the analysis of AVP. The detection limit for AVP was 0.1 pg/ml. All samples were analyzed simultaneously in the same assay run. Intra-assay and inter-assay variability was <10%. There was no significant cross-reactivity with other neuropeptides (<0.7%). AVP values are reported in pg/mL.

2.4 | Data analyses

2.4.1 | Data Distribution

Distribution of the full cases data was visually inspected and normality was tested using the Shapiro Wilk test in SPSS version 23. fMRI mean *z*-values were normally distributed. Basal AVP levels approached normal distribution. T levels were right-skewed and were therefore log-transformed for further analyses. The distributions of self-reported and partner-reported protective behavior were left-skewed and were therefore transformed with a reflective log to approach a normal distribution. The distribution of observed protective behavior was right-skewed, however, transformations did not improve the data distribution, thus nontransformed scores of observed protective behavior were used in further analyses. Sensitivity analyses with Spearman correlations were performed for variables with a non-normal distribution. Visual inspection of boxplots in combination with corresponding *z*-values (outlier > 3.29) indicated one outlier for the variable Age of the participant. The outlier was winsorized (Tab achnick & Fidell, 2007). Further analyses were conducted in SPSS version 23, unless stated otherwise.

2.4.2 | Multiple Imputation

Nine percent of the data was missing (range: 0%–25%). Little MCAR test (Little, 1988b) was not significant ($\chi^2(194) = 191.98$, $p = .53$), indicating that data were missing completely at random. Missing data were multiply imputed with the R package "mice" (Van Buuren & Groothuis-Oudshoorn, 2011) in R (R Development Core Team, 2008). The missing data were imputed 50 times with 100 iterations, using predictive mean matching (PMM; Little, 1988a). For the construction of the prediction model, all variables of interest (observed—and self-reported protective behavior, basal T and AVP levels, individual mean *z*-values for fMRI clusters 1–4), all possible covariates and auxiliary variables (i.e., variables that are not part of the model but that are correlated with the variables in the model: age of the infant at the time of the first assessment and scored baby state before AST noise) were taken into account. Autocorrelation function (ACF) plots (Azur et al., 2011) indicated that all imputations converged. Moreover, correlations between imputed variables (Table 2) were similar to the correlations between nonimputed variables (see Table S3). The imputed datasets were used for further confirmatory analyses.

2.4.3 | Covariates

For the full-cases and pooled multiply imputed data, it was tested whether time of saliva collection, age of participant at time of first assessment, excessive physical exercise, recent alcohol, and caffeine use were significantly correlated with basal hormonal levels. In addition, baby state during AST noise was examined as a significant covariate for observed protective response during the AST and number of days between the start date of the intervention and the date the PPQ was completed was examined as a significant covariate for the PPQ. Only a significant correlation between time of saliva collection and T was observed ($r = -.23$, $p = .048$; $r_{pooled} = -0.23$, $p = .046$), therefore, *T* values were corrected for time of collection via residualizing. Spearman correlations revealed approximately similar outcomes.

2.4.4 | Confirmatory Analyses

Pearson bivariate correlations were calculated to examine the relation between the neural, behavioral, and hormonal measurements. Only neural clusters that included brain areas involved in the parental care network, visual processing areas, or threat detection (as reported in Van 't Veer et al., 2019) were included in the current confirmatory analyses. Significance was set at $p < .05$. Equivalence tests using a two one-sided tests (TOST) procedure, were performed to explore whether observed effects that failed to reach statistical significance were not caused by insufficient statistical power (Lakens et al., 2018). Lower and upper bounds were set to an effect size of $r = 0.08$, which was considered the smallest effect size of interest based on previous literature (Alyousefi-Van Dijk et al., 2019; Meijer et al., 2019).

2.4.5 | Exploratory Analyses

Bivariate correlations between neural clusters activating on the reverse contrast neutral > threat, and behavioral and

TABLE 2 Correlations between neural, behavioral and hormonal measurements based on multiply imputed data

	1.	2	3.	4.	5.	6.	7.
1. Testosterone	.						
2. Vasopressin	−0.06	.					
3. AST	−0.07	−0.20					
4. PPQ	−0.06	−0.06	0.16				
5. fMRI cluster 1	0.16	−0.01	0.02	0.09			
6. fMRI cluster 2	0.22†	0.07	−0.02	0.08	0.85**		
7. fMRI cluster 3	0.14	−0.08	0.00	0.03	0.74**	0.70**	.
8. fMRI cluster 4	0.16	−0.00	−0.05	0.01	0.81**	0.77**	0.90**

Note.: Bivariate Pearson correlations are calculated based on pooled multiply imputed data. AST: protective behavior observed during the Auditory Startling task. PPQ: reflective log transformed self-reported protective behavior measured with the Paternal Protection Questionnaire. fMRI cluster 1–4 are individual mean *z*-values based on the contrast threat > neutral. Cluster 1: left cuneal cortex, left lateral occipital cortex and left occipital pole. Cluster 2: right cuneal cortex, right lateral occipital cortex and right occipital pole. Cluster 3: left supramarginal gyrus, planum temporale, lateral occipital cortex, central opercular cortex, middle temporal cortex, parietal operculum cortex. Cluster 4: insular cortex, frontal orbital cortex, temporal pole, superior frontal gyrus, frontal operculum cortex, central operculum cortex, thalamus, putamen, precentral gyrus, right planum temporale, right supramarginal gyrus, right middle temporal gyrus, right inferior temporal cortex, right superior temporal gyrus, anterior cingulate gyrus, amygdala. Testosterone values were log transformed and corrected for time of collection. Spearman correlations did not differ significantly from Pearson correlations shown above.

**$p < .01$

†$p < 0.1$.

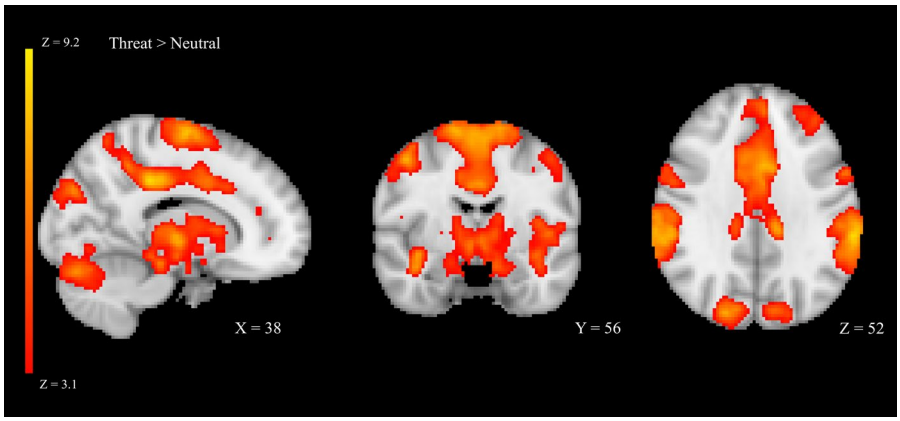

FIGURE 2 Neural activation for the contrast threat > neutral. Note. Activation is thresholded at $Z > 3.2$, $p < .05$

hormonal measurements were explored in the complete cases dataset. Significance was set at $p < .05$. A Benjamini–Hochberg correction was applied to control for multiple testing. When the confirmatory analyses revealed significant correlations between the neural, behavioral, and hormonal measurements, mediating effects of neural responses on the relation between salivary hormone levels and observed and self-reported protective behavior were explored.

2.4.6 | Sensitivity Analyses

We conducted three sensitivity analyses. First, a sensitivity analysis was performed for the fMRI higher-level data analysis, excluding one participant with head motion between 1.4 and 3.0 mm. Second, Spearman correlations were performed when analyses included variables that were not normally distributed. Third, sensitivity analyses were performed excluding nine participants based on the a priori stated exclusion criteria. To control for multiple testing, a Benjamini–Hochberg correction was applied to these sensitivity analyses.

3 | Results

3.1 | fMRI Results

Second level mixed-effect group whole-brain analyses revealed four significant clusters for the contrast threat > neutral (see Figure 2 and Table 3 for cluster information). Descriptive statistics reported below are based on pooled values from the multiple imputed dataset. The first cluster corresponded to the left cuneal cortex, left lateral occipital

TABLE 3 Brain coordinates of the peak average z-value for the contrast threat > neutral

Cluster index	Voxels	Region	*x*	*y*	*z*	Peak *z*	*p*
4	50,297	L Insular cortex	−42	16	−8	8.89	0.00
3	6,255	L Supramarginal gyrus	−64	−36	24	9.24	1.23e-36
2	828	R Lateral occipital cortex	20	−84	32	7.18	4.87e-9
1	556	L Cuneal cortex	−8	−82	30	5.76	8.94e-7

Note: L = left, R = right.

cortex, and the left occipital pole. The average mean *z*-value for activation was 0.69 (*SD* = 0.95, range = −2.15 to 2.53). The second cluster consisted of the right cuneal cortex, right lateral occipital cortex, and the right occipital pole. The average mean *z*-value for activation was 0.77 (*SD* = 0.95, range = −1.98 to 2.90). The third cluster included the left supramarginal gyrus, planum temporale, lateral occipital cortex, central opercular cortex, middle temporal cortex, and parietal operculum cortex. The average mean *z*-value for activation was 1.03 (*SD* = 0.95, range = −1.04 to 2.94). The fourth cluster corresponded to the bilateral insular cortex, bilateral frontal orbital cortex, bilateral temporal pole, bilateral superior frontal gyrus, bilateral frontal operculum cortex, bilateral central operculum cortex, bilateral thalamus, bilateral putamen, bilateral precentral gyrus, right planum temporale, right supramarginal gyrus, right middle temporal gyrus, right inferior temporal cortex, right superior temporal gyrus, bilateral anterior cingulate gyrus, and bilateral amygdala. The average mean *z*-value for activation was 0.89 (*SD* = 0.77, range = −0.79 to 2.62).

Second-level mixed-effect group whole-brain analyses revealed 14 significant clusters for the contrast neutral > threat (see Table S4 and Figure S2 for cluster information).

3.2 | Associations between behavioral, neural, and hormonal measurements

As shown in Table 2, analyses revealed no significant associations between observed and self-reported protective behavior, neural reactivity, and T and AVP levels. Significant correlations were observed between the four fMRI clusters (range $r = 0.70–0.90$, $p < .01$), indicating that neural reactivity in the four clusters were strongly correlated. Equivalence tests were nonsignificant, indicating that the effect sizes were statistically equivalent (i.e., the observed effect sizes were not significantly different from the smallest effect size of interest, which was set at $r = .08$).

3.3 | Exploratory Analyses

As shown in Table S5, analyses revealed no significant associations between the neutral > threat neural reactivity in the 14 clusters and basal hormone levels, observed- and self-reported protective behavior. Since the confirmatory analyses revealed no significant correlations between the neural, behavioral, and hormonal measurements, mediating effects of neural responses on the relation between salivary hormone levels and observed and self-reported protective behavior were not examined.

3.4 | Sensitivity Analyses

Analyses excluding one participant with excessive head motion revealed two significant clusters for the contrast threat > neutral. These two clusters contained the same brain regions as the four clusters in the group analysis reported above. One of these two clusters was very large (64,886 voxels) meaning that the calculated mean *z*-value for this cluster would be less specific for further interpretation. For this reason, it was decided to perform further analyses based on the four clusters. Second, Spearman correlations revealed approximately similar outcomes for the correlational analyses including variables that were not normally distributed. Third, analyses excluding nine participants who did not meet a priori stated inclusion criteria (see Table S6) revealed a negative association between AVP and observed protective behavior ($r = -.28$, $p = .03$) and a positive association between T and fMRI cluster 2 ($r = 0.26$, $p < .05$). These findings did not survive corrections for multiple testing.

4 | DISCUSSION

This study, preregistered on https://osf.io/2acxd, explored the psychobiological correlates of paternal protection in the early postnatal period. To this aim, protective behavior, neural responses to infant threat, and T and AVP levels were assessed in new fathers. Our main analyses revealed no significant associations between the behavioral, neural, and hormonal measures. Because significant associations were absent, the neural responses could not mediate the relation between salivary hormone levels and protective behavior.

Although protection is a crucial aspect of parenting, measures for parental protective behavior are scarce. This study used a postnatal version of the PPQ (Van 't Veer et al., 2019) as an assessment of self-reported and partner-reported protective behavior. In addition, we introduced a new behavioral task to observe paternal protective behavior (the AST). Both measures aim to study the level of paternal precautionary behavior in real-life (Hahn-Holbrook et al., 2011) and obtained good reliability, but they were not significantly related. The measures may assess different aspects of paternal protection. Specifically, the PPQ examines behavior aimed at preventing possible harm to the infant, for example, focusing on hygiene and a safe environment, whereas in the AST an actual stressor, that is, a loud sound, is presented, calling for an immediate protective response. An alternative explanation could be that different response processes are involved in the two measures (Dang et al., 2020). The PPQ asks for participants' own perception of fathers' protective behavior in

daily life situations during the past month, whereas in the AST, participants' actual behavior is assessed in a controlled unknown laboratory setting. Research indicates that reported and observational measures are often weakly correlated (Dang et al., 2020), especially in the realm of parenting (e.g., Voorthuis et al., 2013). In the AST, the majority of participants and infants showed a startle response in response to the sound, and a range of behaviors were observed, suggesting that the manipulation had the intended effect and inter-individual differences in protective behavior were observed. However, one could question whether the loud sound was experienced as a real threat for the infant. Due to ethical reasons, no stronger stressors could be used to trigger a paternal protective response. With these limitations in mind, we suggest that both the PPQ and the AST may be promising measures to assess different dimensions of protective behavior.

We did not find significant correlations between T and AVP levels and self-reported and observed protective behavior in our main analyses. These results did not support the hypotheses that T and AVP would be positively associated with protective behavior. Our hypotheses were based on the Steroid/Peptide Theory of Social Bonds (Van Anders et al., 2011), in which high levels of T and AVP are linked to parental contexts that involve a need for protective responses, particularly in the context of protective aggression. It should be noted that protective aggression is not directly measured in the PPQ or AST; the PPQ and AST are more likely to reflect child-focused behavior instead of parental aggression against the source of the danger the child is exposed to. Measuring child-focused behavior instead of source-focused protective aggression might explain the absence of an association between the hormonal and behavioral measurements. Furthermore, we focused on endogenous basal levels of T and AVP, rather than on reactivity or exogenous administration. This approach differs from that of several other studies (Cohen-Bendahan et al., 2015; Goetz et al., 2014; Thijssen et al., 2018; Van Anders et al., 2012), and might explain the diverging results.

Neural activation in response to infant threat was measured via an fMRI protection paradigm, adapted from Van 't Veer et al. (2019). In this paradigm, situations in which an infant is in immediate danger are shown; these are potential accidents that are likely to provoke a protective behavioral action when occurring in real life. In the current sample, whole-brain analyses revealed four significant clusters for the contrast threat > neutral, and these clusters were strongly related to each other. The clusters comprised brain areas associated with visual processing (Grill-Spector & Malach, 2004), impulse control (Hu et al., 2016), empathy (Feldman, 2015; Rilling & Mascaro, 2017), mentalizing (Feldman, 2015), emotion regulation (Feldman, 2015) and threat detection, and reaction (Hahn-Holbrook et al., 2011; Swain & Ho, 2017). Brain areas involved in the latter four processes are associated with the parental caregiving network (Feldman, 2015; Rilling & Mascaro, 2017). These results align with those by Van't Veer et al. (2019), who also found neural activation in brain networks associated with visual processing, threat detection, and the parental caregiving network, in first-time fathers (to-be) using a similar paradigm in an independent and smaller sample. This study thus validated the role of several brain networks in the processing of infant-threatening situations, and presents replicable findings with the infant-threat paradigm. Future fMRI studies may examine these replicated brain regions using a Regions of Interest (ROI) approach.

We did not find significant associations between neural responses to infant threat and basal T and AVP levels in our main analyses. The lack of associations between T, AVP, and neural responses to infant threat was not in line with our predictions based on limited previous research that explored the relation between basal salivary T, AVP, and neural responses (i.e., in areas involved in social cognition) to positive and negative infant stimuli (Atzil et al., 2012; Khoddam et al., 2020). For example, Khoddam et al. (2020) observed a positive correlation between T and neural responses to infant cry. Although infant cry can be an indicator of an infant-threatening situation, our infant-threat fMRI paradigm may represent a more direct measure for infant-threatening situations as infants in immediate danger are shown. Since no previous studies specifically focused on neural responses to the observation of infant-threatening situations, future studies should elaborate on (the lack of) associations between basal levels and reactive T and AVP, and neural responses to infant threat.

Additionally, no significant associations were observed between neural responses to infant-threatening situations and self-reported and observed protective behavior. This finding is not in line with those of a previous study (Van 't Veer et al., 2019), in which prenatally a positive relation was found between combined partner- and self-reported paternal protective behavior towards their unborn child in daily life and brain activation while watching threatening (versus neutral) situations concerning their own (versus an unknown) infant. The extent to which differences in study design, that is, inclusion of a child familiarity factor and a focus on protective behavior towards the unborn child (and thereby their pregnant partner), might explain the discrepancy in findings with this study remains to be explored by future research.

The following limitations of this study should be mentioned. First, our sample size ($N = 77$), although relatively large compared to previous studies on indicators or correlates of paternal protection (with Ns ranging from 16 to 55), was insufficient to detect small effect sizes. Equivalence tests based on a priori determined smallest effects size of interest were nonsignificant, indicating that the null effects are as yet undetermined (Lakens et al., 2018) and replication studies with larger sample sizes are needed. This highlights the need for more research into paternal protective behavior. Second, a single hormone measurement was used to assess basal T and AVP levels. As a single assessment might not provide a sufficiently reliable estimate of basal hormone levels, results should be replicated. Future studies might consider using repeated measurements. Third, we did not incorporate the measurement of salivary cortisol levels in the current analyses. The hormone cortisol has been associated with parenting behaviors, especially in contexts of parental arousal (Bos, 2017). Previous studies revealed that the relation between T

levels and caregiving behavior was dependent on cortisol levels (e.g., Bos et al., 2018; Voorthuis et al., 2019). Future research studies into the psychobiology of paternal protection are advised to incorporate cortisol as well.

In sum, this study was the first to look at the psychobiology of paternal protective behavior. We used observational as well as self-reported measures of protective parenting. Overall, this study replicated previously reported brain areas to be associated with paternal protection. However, we did not find robust associations between protective behavior, neural activation in response to infant-directed threat and T and AVP levels. Sensitivity analyses indicated a possible link between basal T levels and neural responses to infant-threatening videos, and suggested a negative association between basal AVP levels and observed protective behavior. Taking into account that these findings did not survive the correction for multiple testing and emerged in sensitivity analyses with a somewhat smaller sample size, these results should be interpreted carefully and await replication. Our understanding of the psychobiology of paternal protective behavior is thus still rather incomplete and needs more attention in future research on paternal behavior. Since accumulating knowledge points to fathers' important role in early child development (Bakermans-Kranenburg et al., 2019), further examination of the psychobiology of paternal protective behavior may contribute to a better understanding of the development of both paternal and child behavior.

CREediT author statement

AML; Methodology, Formal analysis, Investigation, Data Curation, Writing—Original Draft, Visualization, Project administration. MWFTV; Methodology, Formal analysis, Data Curation, Writing—Review & Editing, Supervision. RSMB; Methodology, Software, Formal analysis, Data Curation, Writing—Review & Editing, Supervision. LH; Software, Formal analysis, Investigation, Data Curation, Writing—Review & Editing, Project administration. MMEH-R; Methodology, Formal analysis, Data Curation, Writing—Review & Editing, Visualization. MHvIJ; Conceptualization, Methodology, Writing—Review & Editing, Supervision, Funding acquisition. MJB-K; Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding acquisition.

ACKNOWLEDGMENT

The authors thank Dr. Anna van 't Veer and Kim Alyousefi-van Dijk, MSc., for their role in task design, task development, and data collection. Furthermore, the authors thank dr. Dana Shai for kindly sharing the protocol of the Auditory Startling Task, Dr. Tomas Knapen and Noa van der Knaap, BSc., for their assistance in the fMRI data analysis, as well as prof. Carlo Schuengel for his assistance in the conceptualization of this study.

CONFLICT OF INTEREST

None.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Anna M. Lotz https://orcid.org/0000-0001-7868-9870
Martine W. F. T. Verhees https://orcid.org/0000-0001-7471-6341
Marinus H. van IJzendoorn https://orcid.org/0000-0003-1144-454X
Renate S. M. Buisman https://orcid.org/0000-0002-0394-1418

REFERENCES

Abraham, E., & Feldman, R. (2018). The neurobiology of human allomaternal care: Implications for fathering, coparenting, and children's social development. *Physiology and Behavior*, *193*, 25–34. https://doi.org/10.1016/j.physbeh.2017.12.034

Alvergne, A., Faurie, C., & Raymond, M. (2009). Variation in testosterone levels and male reproductive effort: Insight from a polygynous human population. *Hormones and Behavior*, *56*(5), 491–497. https://doi.org/10.1016/j.yhbeh.2009.07.013

Alyousefi-van Dijk, K., van 't Veer, A. E., Meijer, W. M., Lotz, A. M., Rijlaarsdam, J., Witteman, J., & Bakermans-Kranenburg, M. J. (2019). Vasopressin differentially affects handgrip force of expectant fathers in reaction to own and unknown infant faces. *Frontiers in Behavioral Neuroscience*, *13*, 1–10. https://doi.org/10.3389/fnbeh.2019.00105

Atzil, S., Hendler, T., Zagoory-Sharon, O., Winetraub, Y., & Feldman, R. (2012). Synchrony and specificity in the maternal and the paternal brain: Relations to oxytocin and vasopressin. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(8), 798–811. https://doi.org/10.1016/j.jaac.2012.06.008

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. (2011). Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research*, *20*, 40–49. https://doi.org/10.1002/mpr.329

Bakermans-Kranenburg, M. J., Lotz, A., Alyousefi-Van Dijk, K., & Van IJzendoorn, M.(2019). Birth of a father: Fathering in the first 1,000 Days. *Child Development Perspectives*, *13*(4), 247–253. https://doi.org/10.1111/cdep.12347

Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H.(2017). Protective parenting: Neurobiological and behavioral dimensions. *Current Opinion in Psychology*, *15*, 45–49. https://doi.org/10.1016/j.copsyc.2017.02.001

Bales, K. L., & Saltzman, W. (2016). Fathering in rodents: Neurobiological substrates and consequences for offspring. *Hormones and Behavior*, *77*, 249–259. https://doi.org/10.1016/j.yhbeh.2015.05.021

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

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

Bowbly, J. (1982). Attachment and loss: Retrospect and prospect. *American Journal of Orthopsychiatry*, *52*(4), 664–678. https://doi.org/10.1111/j.1939-0025.1982.tb01456.x

Cohen-Bendahan, C. C. C., Beijers, R., Van Doornen, L. J. P., & de Weerth, C. (2015). Explicit and implicit caregiving interests in expectant fathers: Do endogenous and exogenous oxytocin and vasopressin matter? *Infant Behavior and Development*, *41*, 26–37. https://doi.org/10.1016/j.infbeh.2015.06.007

Dang, J., King, K. M., & Inzlicht, M. (2020). Why are self-report and behavioral measures weakly correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269. https://doi.org/10.1016/j.tics.2020.01.007

Durnez, J., Blair, R., & Poldrack, R. (2017). Neurodesign: Optimal experimental designs for task fMRI. *BioRxiv* 119594 [preprint], 1–21. https://doi.org/10.1101/119594

Edelstein, R. S., Chopik, W. J., Saxbe, D. E., Wardecker, B. M., Moors, A. C., & LaBelle, O. P. (2017). Prospective and dyadic associations between expectant parents' prenatal hormone changes and postpartum parenting outcomes. *Developmental Psychobiology*, *59*, 77–90. https://doi.org/10.1002/dev.21469

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2018). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*, 111–116. https://doi.org/10.1038/s41592-018-0235-4

Feldman, R. (2015). The adaptive human parental brain: Implications for children's social development. *Trends in Neurosciences*, *38*(6), 387–399. https://doi.org/10.1016/j.tins.2015.04.004

Frijling, J. L., van Zuiden, M., Nawijn, L., Koch, S. B. J., Neumann, I. D., Veltman, D. J., & Olff, M. (2015). Salivary oxytocin and vasopressin levels in police officers with and without post-traumatic stress disorder. *Journal of Neuroendocrinology*, *27*, 743–751. https://doi.org/10.1111/jne.12300

Gettler, L. T., McDade, T. W., Feranil, A. B., & Kuzawa, C. W. (2011). Longitudinal evidence that fatherhood decreases testosterone in human males. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(39), 16194–16199. https://doi.org/10.1073/pnas.1105403108

Goetz, S. M. M., Tang, L., Thomason, M. E., Diamond, M. P., Hariri, A. R., & Carré, J. M. (2014). Testosterone rapidly increases neural reactivity to threat in healthy men: A novel two-step pharmacological challenge paradigm. *Biological Psychiatry*, *76*, 324–331. https://doi.org/10.1016/j.biopsych.2014.01.016

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*, 1–9. https://doi.org/10.1038/sdata.2016.44

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, *5*, 13. https://doi.org/10.3389/fninf.2011.00013

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–677. https://doi.org/10.1146/annurev.neuro.27.070203.144220

Hahn-Holbrook, J., Holbrook, C., & Haselton, M. G. (2011). Parental precaution: Neurobiological means and adaptive ends. *Neuroscience and Biobehavioral Reviews*, *35*, 1052–1066. https://doi.org/10.1016/j.neubiorev.2010.09.015

Hu, S., Ide, J. S., Zhang, S., & Li, C. S. R. (2016). The right superior frontal gyrus and individual variation in proactive control of impulsive response. *Journal of Neuroscience*, *36*(50), 12688–12696. https://doi.org/10.1523/JNEUROSCI.1175-16.2016

Kagerbauer, S. M., Debus, J. M., Martin, J., Gempt, J., Jungwirth, B., Hapfelmeier, A., & Podtschaske, A. H. (2019). Absence of a diurnal rhythm of oxytocin and arginine-vasopressin in human cerebrospinal fluid, blood and saliva. *Neuropeptides*, *78*, 1–6. https://doi.org/10.1016/j.npep.2019.101977

Khoddam, H., Goldenberg, D., Stoycos, S. A., Horton, K. T., Marshall, N., Cárdenas, S. I., Kaplan, J., & Saxbe, D. (2020). How do expectant fathers respond to infant cry? Examining brain and behavioral responses and the moderating role of testosterone. *Social Cognitive and Affective Neuroscience*, *15*(4), 437–446. https://doi.org/10.1093/scan/nsaa051

Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Thomas Boyce, W., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives background: Child risk and psychopathology. *American Journal of Psychiatry*, *160*, 1566–1577. https://doi.org/10.1176/appi.ajp.160.9.1566

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Li, T., Chen, X., Mascaro, J., Haroon, E., & Rilling, J. K. (2017). Intranasal oxytocin, but not vasopressin, augments neural responses to toddlers in human fathers. *Hormones and Behavior*, *93*, 193–202. https://doi.org/10.1016/j.yhbeh.2017.01.006

Little, R. J. A. (1988a). Missing-data adjustments in large surveys. *Journal of Business and Economic Statistics*, *6*(3), 287–296. https://doi.org/10.1080/07350015.1988.10509663

Little, R. J. (1988b). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Lotz, A. M., Rijlaarsdam, J., Witteman, J., Meijer, W., van Dijk, K., van Ijzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2020). Vasopressin and parental expressed emotion in the transition to fatherhood. *Attachment and Human Development*, 1–17. https://doi.org/10.1080/14616734.2020.1719427

Mah, B. L., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., & Smith, R. (2015) Oxytocin promotes protective behavior in depressed mothers: A pilot study with the enthusiastic stranger paradigm. *Depression and Anxiety*, *32*, 76–81. https://doi.org/10.1002/da.22245

Mascaro, J. S., Hackett, P. D., Gouzoules, H., Lori, A., & Rilling, J. K. (2014). Behavioral and genetic correlates of the neural response to infant crying among human fathers. *Social Cognitive and Affective Neuroscience*, *9*(11), 1704–1712. https://doi.org/10.1093/scan/nst166

Meijer, W. M., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2019). Challenging the challenge hypothesis on testosterone in fathers: Limited meta-analytic support. *Psychoneuroendocrinology*, *110*, 1–11. https://doi.org/10.1016/j.psyneuen.2019.104435

Rilling, J. K., & Mascaro, J. S. (2017). The neurobiology of fatherhood. *Current Opinion in Psychology*, *15*, 26–32. https://doi.org/10.1016/j.copsyc.2017.02.013

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Swain, J. E., & Ho, S. H. S. (2017). Neuroendocrine mechanisms for parental sensitivity: Overview, recent advances and future directions. *Current Opinion in Psychology*, *15*, 105–110. https://doi.org/10.1016/j.copsyc.2017.02.027

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Allyn & Bacon/Pearson Education.

Thijssen, S., Van 't Veer, A. E., Witteman, J., Meijer, W. M., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2018). Effects of vasopressin on neural processing of infant crying in expectant fathers. *Hormones and Behavior*, *103*, 19–27. https://doi.org/10.1016/j.yhbeh.2018.05.014

van 't Veer, A. E., Thijssen, S., Witteman, J., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2019). Exploring the neural basis for paternal protection: An investigation of the neural response to infants in danger. *Social Cognitive and Affective Neuroscience*, *14*(4), 447–457. https://doi.org/10.1093/scan/nsz018

Van Anders, S. M., Goldey, K. L., & Kuo, P. X. (2011). The steroid/peptide theory of social bonds: Integrating testosterone and peptide responses

for classifying social behavioral contexts. *Psychoneuroendocrinology*, *36*, 1265–1275. https://doi.org/10.1016/j.psyneuen.2011.06.001

Van Anders, S. M., Tolman, R. M., & Volling, B. L. (2012). Baby cries and nurturance affect testosterone in men. *Hormones and Behavior*, *61*, 31–36. https://doi.org/10.1016/j.yhbeh.2011.09.012

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

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 Behavior and Development*, *56*, 1–9. https://doi.org/10.1016/j.infbeh.2017.08.002

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). *Attachment and Human Development*, *15*(5–6), 603–617. https://doi.org/10.1080/14616734.2013.820897

Weisman, O., Zagoory-Sharon, O., & Feldman, R. (2014). Oxytocin administration, salivary testosterone, and father-infant social behavior. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *49*, 47–52. https://doi.org/10.1016/j.pnpbp.2013.11.006

Witte, A. M., de Moor, M. H. M., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2019). Fathers today: Design of a randomized controlled trial examining the role of oxytocin and vasopressin in behavioral and neural responses to infant signals. *BMC Psychology*, *7*(81), 1–11. https://doi.org/10.1186/s40359-019-0356-2

Wynne-Edwards, K. E., & Timonin, M. E. (2007). Paternal care in rodents: Weakening support for hormonal regulation of the transition to behavioral fatherhood in rodent animal models of biparental care. *Hormones and Behavior*, *52*(1), 114–121. https://doi.org/10.1016/j.yhbeh.2007.03.018

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Lotz AM, Verhees MWFT, Horstman LI, et al. Exploring the hormonal and neural correlates of paternal protective behavior to their infants. *Dev Psychobiol*. 2020;00:1–12. https://doi.org/10.1002/dev.22055