Check for updates

Developmental Psychobiology

WILEY

Developmental Psychobiology

RESEARCH ARTICLE OPEN ACCESS

Do Child Gender Stereotypes Mediate Associations Between Human Fathers' Testosterone Levels and Paternal Involvement? Two Studies

Annemieke M. Witte[1] | Noor de Waal[2] | Martine W. F. T. Verhees[3] | Anna M. Lotz[4] | Marinus H. van IJzendoorn[5,6] | Marian J. Bakermans-Kranenburg[7,8,9]

[1]Clinical Child & Family Studies, Faculty of Behavioral and Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands | [2]Department of Developmental Psychology, Tilburg University, Tilburg, The Netherlands | [3]Quantitative Psychology and Individual Differences, KU Leuven, Leuven, Belgium | [4]Institute of Education and Child Studies, Faculty of Social and Behavioral Sciences, Leiden University, Leiden, The Netherlands | [5]Research Department of Clinical, Education and Health Psychology, Faculty of Brain Sciences, UCL, London, UK | [6]Department of Psychiatry, Monash University, Melbourne, Victoria, Australia | [7]William James Centre for Research, ISPA University Institute of Psychological, Social and Life Sciences, Lisbon, Portugal | [8]Center for Attachment Research, The New School for Social Research, New York, New York, USA | [9]Facultad de Psicología y Humanidades, Universidad San Sebastián, Sede Valdivia, Chile

Correspondence: Annemieke M. Witte (a.m.witte@vu.nl)

Received: 26 February 2025 | **Revised:** 20 June 2025 | **Accepted:** 18 August 2025

Funding: This study was supported by a European Research Council grant (ERC-AdG 669249) awarded to M.J.B-K and a Spinoza prize (Netherlands Organization for Scientific Research) awarded to M.H.v.I.J.

Keywords: fathers | gender stereotypes | hair testosterone levels | household chore | involvement

ABSTRACT

This preregistered study investigated whether human paternal testosterone levels predicted fathers' time spent with their child and their involvement in childcare and household tasks. Additionally, we examined whether associations were mediated by fathers' attitudes toward child gender stereotypes. Associations were tested in an exploratory sample ($n = 70$, $M = 6.73$ months postpartum, $SD = 2.13$) and a test sample ($n = 80$, $M = 2.67$ months postpartum, $SD = 0.77$). Testosterone was quantified from the 1-cm hair segments closest to the scalp. Fathers reported on the amount of time spent with their child and on their attitudes toward child gender stereotypes. Mothers and fathers reported on fathers' involvement in childcare and household tasks. In the exploratory sample, testosterone levels were positively associated with involvement in childcare and household tasks but unrelated to the time spent with the child. In the test sample, testosterone levels did not predict the time spent with the child or involvement in childcare and household tasks. In both samples, no evidence was found for mediation by attitudes toward child gender stereotypes. In conclusion, testosterone levels are differently but not systematically associated with involvement in childcare and household tasks during the first year of fatherhood.

1 | Introduction

Over the past few decades, fathers have become increasingly involved in childcare and household tasks, although mothers still spend more time with their children and devote more time to childcare and household tasks than fathers (Bakermans-Kranenburg et al. 2019; Craig 2006; Craig and Mullan 2011; Milkie et al. 2002; Yavorsky et al. 2015). While empirical evidence strongly supports the importance of father involvement for positive child development (e.g., Cano et al. 2019; Sarkadi et al.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2025 The Author(s). *Developmental Psychobiology* published by Wiley Periodicals LLC.

2008), the underlying mechanisms shaping fathers' involvement are largely unknown. According to the biobehavioral model of emergent fatherhood, reciprocal relations exist among the sociocultural context, hormonal processes, and paternal behavior in the prenatal, perinatal, and postnatal phases of fatherhood (Bakermans-Kranenburg et al. 2019). The present study investigated whether paternal testosterone levels measured in hair were associated with the amount of time fathers spent with their child and their involvement in childcare and household tasks. Furthermore, we examined whether associations were mediated by fathers' attitudes toward child gender stereotypes. Associations were tested in two studies including fathers in the first year of fatherhood, with one sample serving as an exploratory sample ($M = 6.73$ months postpartum, $SD = 2.13$) and one sample as a test sample ($M = 2.67$ months postpartum, $SD = 0.77$).

Testosterone is the primary male sex hormone that is produced by the hypothalamic–pituitary–gonadal (HPG) axis and promotes secondary sexual characteristics (e.g., body hair, muscle, and bone mass) as well as competitive, aggressive, antisocial, and mate-seeking behaviors (Grebe et al. 2019). During the course of pregnancy, fathers showed a decline in their testosterone levels (Saxbe et al. 2017). Furthermore, in several studies, testosterone decreased from the third trimester of pregnancy to 3 weeks (Storey et al. 2000), 5 weeks (Bakermans-Kranenburg et al. 2022), or 3 months postpartum (Corpuz et al. 2021). However, other studies reported that testosterone levels increased from the third trimester of pregnancy to 6 months postpartum (Witte et al. 2020) and from 3 to 9–10 months postpartum (Corpuz et al. 2021). These findings suggest that the timing of measurements of testosterone may impact the observed patterns in paternal hormonal trajectories. In addition, meta-analytic evidence indicates that fathers have lower testosterone levels than nonfathers, albeit combined effect sizes are relatively small ($r = 0.19$ in Grebe et al. [2019]; $r = 0.11$ in Meijer et al. [2019]), indicating that being a father cannot fully account for lower testosterone levels in men (see also Muller et al. 2009).

It has been suggested that a decrease in testosterone levels indicates a shift away from mating opportunities to investment in offspring and the partner relationship (Gettler et al. 2014; Gray and Campbell 2009; Wingfield et al. 1990). Indeed, partnered men with lower testosterone levels reported to have more satisfied and committed romantic relationships (Edelstein et al. 2014; Julian and McKenry 1989) and lower interest in sexual activity outside their romantic relationship (Edelstein et al. 2011; McIntyre et al. 2006). Furthermore, fathers who showed a greater decline in testosterone levels over the course of pregnancy reported higher postpartum investment, commitment, and relationship satisfaction (Saxbe et al. 2017). With respect to caregiving, meta-analytic studies have shown that lower paternal testosterone levels are related to more positive fathering behaviors (Grebe et al. 2019; Meijer et al. 2019; see also Luberti and Carré 2024). In line with these findings, higher paternal testosterone levels were associated with more negative parenting toward 18-month-old children and a higher likelihood for engaging in physical abusive behaviors (Rodriguez et al. 2021). Another study showed that in fathers with higher levels of estradiol, higher levels of testosterone were related to lower levels of sensitive caregiving (Bakermans-Kranenburg et al. 2022). Nevertheless, higher testosterone levels may not always predict more negative parenting behavior—in fact, higher testosterone levels may be adaptive under specific circumstances, such as challenging or threatening caregiving situations (Bos et al. 2010; Rilling and Mascaro 2017; van Anders et al. 2012).

Fathers' testosterone levels have also been related to paternal involvement (e.g., Gettler et al. 2011). Here, the focus is on two indicators of paternal involvement, namely the amount of time spent with the child and involvement in childcare and household activities. Research showed that fathers with lower testosterone levels were more involved in direct childcare (e.g., bathing, attending to the baby at night) and displayed more interest in taking an active role in caring for their toddler (Mascaro et al. 2013). In another study, Kuo et al. (2018) found that fathers who had lower testosterone levels on the second day after the birth of their baby reported to be more involved in direct and indirect parenting tasks when the infant was 2–4 months old. Moreover, fathers whose testosterone levels showed a greater decline during pregnancy were rated by their partners (but not by themselves) as more supportive and involved in household tasks (Edelstein et al. 2017). In addition, Corpuz et al. (2021) found that fathers who demonstrated a relative steeper increase in their testosterone levels from 3 to 9–10 months postpartum reported to have spent less time in active childcare.

Some studies have related paternal testosterone levels to the amount of time fathers spent with their child. De Vries et al. (2019) reported that testosterone levels were lower on home visit days (i.e., when fathers spent more time with their 12- to 30-month-old child) as compared to a working day, suggesting that testosterone levels may vary in response to the day-to-day amount of time fathers spent with the child. Other studies, however, showed no significant differences in fathers' testosterone levels between days when fathers spent more time with their child and days when fathers spent less time with their child (Gray et al. 2004; Storey et al. 2011). Taken together, most studies suggest that lower paternal testosterone levels are associated with more involvement in childcare and household tasks, while evidence for an association between testosterone levels and the amount of time fathers spent with their child has been inconsistent.

Associations between fathers' testosterone levels and paternal involvement may be mediated by their attitudes toward child gender stereotypes. Child gender stereotypes can be defined as attitudes, beliefs, and expectancies toward the characteristics and behaviors of boys (e.g., only boys should engage in competitive sports) and girls (e.g., quiet girls will have a happier life than assertive girls). Gender Schema Theory (Bem 1981, 1983) suggests that individuals' behavior is guided by internalized gender schemas, which are shaped by gender-typed experiences. It is plausible that gender schemas influence not only how people view and enact adult gender roles but also how parents perceive children's gender roles and engage in parenting (related) behavior. Currently, there are no studies that have examined relations between father's attitudes toward child gender stereotypes and paternal involvement. Research in Western countries showed that fathers with more egalitarian gender role attitudes were more involved in childcare (Atkinson et al. 2021; Bulanda 2004). In a similar vein, Dutch men with more egalitarian gender role attitudes spent more time on recreational activities with their 0- to 4-year-old children (e.g., playing board

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

games), provided more physical and logistical caregiving (e.g., bathing the child, helping the child get dressed), and were more involved in responsible childcare tasks (e.g., staying at home when the child is sick). However, men's gender role attitudes were less predictive of the amount of time that fathers spent on recreational activities when children were older (Keizer 2015). Other research showed that more traditional gender role attitudes were associated with a more traditional division of household tasks among spouses (McGovern and Meyers 2002). It remains unknown whether the above findings can be generalized to associations of fathers' attitudes toward child gender stereotypes and paternal involvement.

There are also no studies available examining relations between paternal testosterone levels and fathers' attitudes toward child gender stereotypes. Research outside the fathering context has shown that higher testosterone levels in boys predicted more male-typical play behavior (Lamminmäki et al. 2012). Furthermore, girls diagnosed with classical congenital adrenal hyperplasia (CAH), a genetic disorder that causes abnormally high levels of testosterone, displayed greater interest in male-typical toys, male activities, and male playmates (Hines 2011). Similarly, mothers who took androgenic progestins during pregnancy gave birth to children who displayed greater preferences for male-typical toys and activities, whereas mothers who took antiandrogenic progestins during pregnancy gave birth to children who showed an opposite preference (Ehrhardt et al. 1968; Ehrhardt et al. 1977). Although these findings demonstrate a link between testosterone levels and gender-related behaviors and preferences of children, it remains unclear whether testosterone levels are associated with attitudes toward child gender stereotypes in adult men. It may, however, be speculated that fathers with higher testosterone levels hold beliefs that boys and girls should act and behave in certain ways that are consistent with traditional gender stereotypes (e.g., girls should not be competitive). Yet, the question whether associations between fathers' testosterone levels and paternal involvement are mediated by fathers' attitudes toward child gender stereotypes is clearly a research question for exploratory studies in the context of discovery (Van IJzendoorn and Bakermans-Kranenburg 2024).

Previous studies investigating associations between paternal testosterone levels and parenting behavior have typically focused on testosterone levels in saliva or blood. Here, we measure fathers' testosterone levels in hair samples, extending the limited body of research that has applied this method to study fatherhood (Verhees et al. 2021) or self-reported aggression or risk taking in males (Grotzinger et al. 2018; Ronay et al. 2018). An advantage of measuring testosterone in hair is that it provides a more long-term indication of fathers' testosterone levels as compared to saliva and blood samples (Dettenborn et al. 2012; Grotzinger et al. 2018). That is, the 1-cm hair segment closest to the scalp reflects testosterone secretion in the preceding month as hair typically grows approximately 1 cm per month (Stalder et al. 2017). Of note, testosterone levels in hair have been moderately correlated with testosterone levels in saliva (Chen et al. 2019; Stern et al. 2022).

In sum, the aims of the present preregistered study were to examine whether fathers' testosterone levels were associated with (1) the amount of time fathers spent with their child and (2) their involvement in childcare and household tasks; and whether these associations were mediated by fathers' attitudes toward child gender stereotypes. It was hypothesized that fathers with lower testosterone would spend more time with their child and would be more involved in childcare and household tasks. Furthermore, we hypothesized that fathers with lower levels of testosterone would have less specific child gender stereotypes, which in turn would predict more time spent with their child and higher levels of involvement in childcare and household tasks. These hypotheses are important to test across different phases of fatherhood, but perhaps particularly in the first year of fatherhood as this is an important developmental milestone in men's lives, in which they adjust to their new role and responsibilities as a parent. In the present study, we examined our hypotheses in two studies, with one sample serving as an exploratory sample (Study 1) and the other as a test sample (Study 2).

2 | Study 1

2.1 | Method

2.1.1 | Procedure

Participants were enrolled in a research program named "Father Trials," which consisted of three randomized controlled trials (RCTs) for which three independent samples were recruited. Participants in Study 1 were enrolled in a randomized double-blind, placebo-controlled within-subject trial investigating the effects of oxytocin and vasopressin administration on behavioral and neural responses to infant signals. Participants could enroll when infants were between 2 and 12 months old, as hormonal effects were expected during this period (Witte, de Moor, et al. 2022; Witte, Riem, et al. 2022). For this reason, the present study used data collected during this postnatal period. Participants took part in three research visits with intervening periods of 1–2 weeks. During each visit, fathers self-administered a nasal spray containing oxytocin, vasopressin, or a placebo (not relevant to the current study). During the first visit, a strand of hair was cut to measure fathers' testosterone levels. After completion of the first visit, fathers and their partners were requested to fill in a set of online questionnaires at home. For this study, data used were all from the first research visit.

2.1.2 | Sample

Participants included 70 first-time fathers with a healthy single full-term born infant (47% girls) between 2 and 12 months old (M = 6.73, SD = 2.13). Most fathers (99%) were born in the Netherlands, and they were on average 35.48 years old (SD = 4.31, range 27–49). Fathers were fluent in speaking and writing Dutch and cohabited with the biological mother (M = 33.3 years, SD = 2.86) and infant. Both parents had parental authority. Fathers' educational level was as follows: 67% completed higher university education (e.g., master's degree), 24% completed higher professional education (e.g., bachelor's degree), and 9% completed higher vocational education. Two fathers reported no working hours during the time of the research visit (n = 1, medical reason; n = 1, no reason provided). Fathers worked on average 34.86 h (SD = 8.49) per week during the course of the

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

study. Fathers were recruited via the municipality database of the city of Amsterdam in the Netherlands. Fathers were excluded from participation in case of neurological disorders, endocrine diseases, psychiatric disorders, cardiovascular diseases, use of psychoactive medications, excessive alcohol or drugs intake, use of hard drugs within the past 3 months, or MRI contraindications (e.g., metallic foreign objects). Travelling expenses were covered, and fathers received a financial compensation for participation in the study. The study was approved by the Ethics Committees of the Leiden University Medical Centre and registered in the Overview of Medical Research in the Netherlands (ID: NL-OMON25106). Both parents provided written informed consent prior to the start of the study.

2.1.3 | Measures

2.1.3.1 | Hair Testosterone Levels. Hair strands with a diameter of approximately 3–5 mm were cut as close as possible to the inion of the occipital protuberance (i.e., the center of the back of the head). Hair strands were wrapped in aluminum foil and stored at room temperature prior to shipment to Dresden Labservices GMBH (http://www.labservice-dresden.de) for further analysis. The 1-cm hair segments closest to the scalp were quantified for testosterone. These segments reflect testosterone secretion in the month preceding hair sampling as hair typically grows approximately 1 cm per month (Stalder et al. 2017). The quantification of hair testosterone concentrations was achieved through the use of liquid chromatography–tandem mass spectrometry with online solid-phase extraction (LC–MS/MS) (Gao et al. 2013). In similar samples analyzed in the same lab, the inter- and intraassay variability coefficients of variation were less than 10% for testosterone concentrations in hair (Gao et al. 2013). A random selection of 36% of hair segments were assayed in duplicate. Following Salimetrics guidelines (www.salimetrics.com), the average intraassay coefficient of variation was calculated by first determining the standard deviation of both hair samples. This standard deviation was then divided by the mean of the two measurements and consequently multiplied by 100. The average intraassay coefficient of variation was 6%. For hair segments that were assayed in duplicate, the average value of the two assays was included in further analyses. Fathers' testosterone levels were not related to age of the father or child. The distribution was positively skewed, and therefore, a log transformation was performed to approach a normal distribution.

2.1.3.2 | Time Spent With child. The amount of time fathers spent with their child was measured with an online questionnaire (Horstman et al. 2022; Verhees et al. 2023). Fathers were asked to report the number of hours they had spent with their child for each day of the week, excluding the hours during which either the father, the child, or both were sleeping. Inspection of z-values indicated one outlier ($z > 3.29$), which was winsorized. A mean score was computed to indicate the daily number of hours fathers spent with their child.

2.1.3.3 | Involvement in Childcare and Household Tasks. Fathers' involvement in childcare and household tasks was measured via the Task Division Questionnaire (TDQ), which was developed for the purpose of the current study (see Table S1 for more details). Fathers and mothers reported on a total of 18 items measuring the division of household tasks during the past week (e.g., "washing dishes or loading/unloading the dishwasher" and "making dinner"). In addition, five items were administered measuring division of childcare tasks (e.g., "bathing and dressing the child in the morning" and "changing diapers"). Items were scored on a 6-point rating scale: 1 = *Neither of us*, 2 = *(Almost) exclusively myself*, 3 = *More often me than my partner*, 4 = *Both equally*, 5 = *More often my partner than me*, and 6 = *(Almost) exclusively my partner*. Six items were excluded because many parents (>19%) indicated that neither of them engaged in the task (see Table S2 for more details). For consistency purposes, one item was excluded because more than 45% of fathers and mothers in Study 2 indicated that neither of them engaged in the activity (i.e., "dropping off and picking up the baby"), probably because of the young age of the baby. Subsequently, response options were recoded into a 5-point rating scale, ranging from 1 = *(Almost) exclusively myself* to 5 = *(Almost) exclusively my partner*. Fathers' scores were mirrored so that higher scores reflected greater father involvement in childcare and household tasks for both mothers and fathers. Fathers' mean self-report scores were significantly higher than mothers' scores for them, $t(59) = 7.99$, $p < 0.001$, $d = 1.03$, but scores of fathers and mothers were strongly correlated ($r = 0.61$, $p < 0.001$). Therefore, z-scores of both mother and father reports were averaged into one score indicating paternal involvement in childcare and household tasks. Cronbach's alpha was 0.75 (combined father and mother reports).

2.1.3.4 | Attitudes Toward Child Gender Stereotypes. Fathers' attitudes toward child gender stereotypes were measured via the Gender Specific Orientation Questionnaire (GSOQ), which was developed for the purpose of the current study (see Table S3 for more details). The 19-item questionnaire included items such as "Only boys should engage in competitive sports" and "Quiet girls will have a happier life than assertive girls." Fathers rated items on a 5-point rating scale: 1 = *Totally agree*, 2 = *Agree*, 3 = *Neutral*, 4 = *Disagree*, and 5 = *Totally disagree*. In total, 14 items were recoded such that higher scores indicated more specific child gender stereotypes. Cronbach's alpha was 0.79. A mean score was calculated and included in further analysis.

2.1.3.5 | Missing Data. Rates of missing data per variable were as follows: 38.6% hair testosterone levels (due to baldness, insufficient hair material or testosterone levels that could not be detected), 7.1% time spent with child, 2.9% involvement in childcare and household tasks, and 7.1% child gender stereotypes. Missing value analysis showed that Little MCAR (Little 1988) was not significant, $\chi^2(86) = 89.79$, $p = 0.369$, indicating that data were missing completely at random. Prior to conducting the statistical analyses, missing values were imputed through predictive mean matching using the MICE (Multivariate Imputation by Chained Equations) package (Van Buuren and Groothuis-Oudshoorn 2011). Autocorrelation function plots (Azur et al. 2011) revealed convergence of 50 imputed datasets.

2.1.4 | Data Analysis Plan

2.1.4.1 | Main Analysis. Preliminary analyses were performed in SPSS version 27.0, and mediation analyses were

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

performed in R using the package lavaan, which yield pooled estimates for direct and indirect effects (Rosseel 2012). Model fit was evaluated using the comparative fit index (CFI) and the root mean square error of approximation (RMSEA). Acceptable and good model fit is indicated by CFI values above 0.90 and 0.95, respectively, and RSMEA values below 0.08 and 0.06, respectively (Hu and Bentler 1999). Standardized regression coefficients were examined to evaluate the direct and indirect effects. Separate mediation models were performed with amount of time spent with child and involvement in childcare household tasks as dependent variables. As Study 1 is explorative in nature, we did not conduct a power analysis. We did not deviate from the preregistered analytical plan, except that we also combined the reported associations in Study 1 and Study 2 meta-analytically.

2.1.4.2 | Sensitivity Analyses. To test the robustness of the results, sensitivity analyses were conducted with involvement in childcare and household tasks reported by mothers and fathers separately.

2.2 | Results

Table 1 presents the descriptive statistics and Pearson correlations based on complete data and imputed data. Fathers' testosterone levels were positively associated with fathers' involvement in childcare and household tasks ($r_{pooled} = 0.30$, $p = 0.041$). No other significant correlations were observed in the complete cases or imputed datasets. Child sex and child age were not related to the variables of interest.

2.2.1 | Testosterone Levels, Child Stereotypes, and Amount of Time Spent With Child

The model testing whether the association between fathers' testosterone levels and the amount of time spent with their child was mediated by child gender stereotypes showed a good fit, $\chi^2(3) = 2.39$, $p = 0.496$, CFI = 1.00, RMSEA < 0.001. Fathers' testosterone levels were not directly associated with the amount of time spent with their child, $\beta = -0.14$, $p = 0.308$. Fathers' testosterone levels were not significantly related to attitudes toward child gender stereotypes, $\beta = 0.13$, $p = 0.340$. Fathers' attitudes toward child stereotypes were not significantly associated with the amount of time spent with their child, $\beta = 0.14$, $p = 0.272$. Finally, testosterone levels were not indirectly related to the amount of time spent with their child via child gender stereotypes, $\beta = -0.02$, $p = 0.506$ (see Figure 1).

2.2.2 | Testosterone Levels, Child Stereotypes, and Involvement in Childcare and Household Tasks

The model testing whether the association between fathers' testosterone levels and involvement in childcare and household tasks was mediated by child gender stereotypes showed a good fit, $\chi^2(3) = 4.11$, $p = 0.250$, CFI = 1.00, RMSEA < 0.001. Fathers' testosterone levels were directly and positively associated with involvement in childcare and household tasks, $\beta = 0.29$, $p = 0.033$, indicating that fathers with higher testosterone levels were more involved in household and childcare tasks. Fathers'

TABLE 1 | Descriptive statistics and correlations between variables based on complete cases and imputed data.

Variables	*M* (*SD*)	Min	Max	*N*	1.	2.	3.	4.
1. Testosterone levels (in hair)	1.06 (0.84)	0.23	3.81	43	—	0.29	−0.12	0.13
2. Involvement in childcare and household tasks	−0.00 (0.89)	−2.47	2.17	68	0.30*	—	−0.17	−0.03
3. Amount of time spent with child (hours)	5.90 (1.68)	2.32	10.71	65	−0.11	−0.18	—	0.13
4. Attitudes toward child gender stereotypes	1.59 (0.38)	1.05	2.74	65	0.14	−0.04	0.13	—

Note: Involvement in childcare and household tasks is the average of the standardized scores of mother and father reports. Hair testosterone concentrations are presented inpg/mL. Correlations based on complete cases are presented above the diagonal, and pooled correlations for the 50 imputed datasets are presented below the diagonal.
*$p < 0.05$.

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

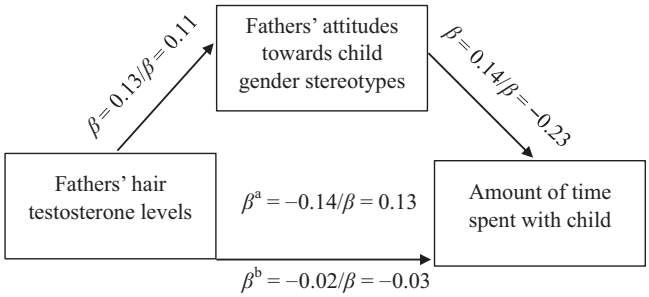

FIGURE 1 | Mediation model testing whether fathers' hair testosterone levels predict fathers' time spent with their child via fathers' attitudes toward child gender stereotypes. Standardized path coefficients in Study 1 and Study 2 are presented on the left and right side of the slash, respectively. [a]Direct effect. [b]Indirect effect.

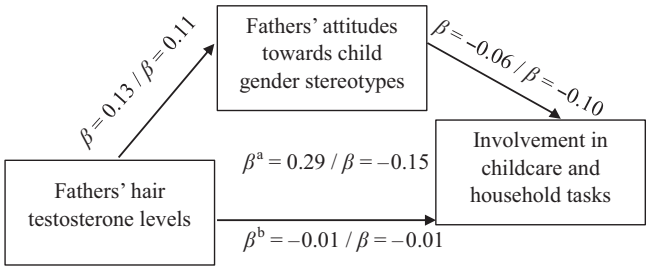

FIGURE 2 | Mediation model testing whether fathers' hair testosterone levels predict involvement in childcare and household tasks via fathers' attitudes toward child gender stereotypes. Standardized path coefficients in Study 1 and Study 2 are presented on the left and right side of the slash, respectively. [a]Direct effect. [b]Indirect effect.

testosterone levels were not significantly associated with attitudes toward child gender stereotypes, $\beta = 0.13$, $p = 0.347$. Fathers' attitudes toward child stereotypes were not significantly related to involvement in childcare and household tasks, $\beta = -0.06$, $p = 0.623$. Fathers' testosterone levels were thus not indirectly related to involvement in childcare and household tasks via child gender stereotypes, $\beta = -0.01$, $p = 0.681$ (see Figure 2).

2.2.3 | Sensitivity Analyses

Preregistered sensitivity analyses were conducted with involvement in childcare and household tasks reported by mothers and fathers separately. Results indicated that the association between fathers' testosterone levels and involvement in childcare and household tasks as reported by fathers fell just short of significance ($\beta = 0.24$, $p = 0.077$). The association between fathers' testosterone levels and involvement in childcare and household tasks as reported by mothers was significant ($\beta = 0.28$, $p = 0.044$). In response to a reviewer's suggestion, we tested correlations between fathers' testosterone levels and their self-reported depressive symptoms, as measured with the Edinburgh Postnatal Depression Scale (EPDS; Cox et al. 1987), and their self-reported family functioning, as measured with the General Functioning 12-item subscale (GF12) of the McMaster Family Assessment Device (FAD; Epstein et al. 1983). No significant associations were found. These additional analyses suggest that the unexpected positive association between fathers' testosterone levels and involvement in childcare and household tasks cannot be accounted for by differences in depressive symptoms or family functioning.

3 | Study 2

3.1 | Method

3.1.1 | Procedure

Participants were enrolled in the research program "Father Trials" and took part in an RCT investigating the effects of a baby carrier or baby seat on fathers' behavioral, hormonal, and neural functioning (e.g., Riem et al. 2021; Verhees et al. 2023). Participants could enroll when infants were between 1 and 3 months old, as this age period was considered relevant for baby carrier usage. As a result, the examination of associations was restricted to this particular postnatal window. This study used data from the first research visit. During the first research visit (before any intervention took place), a strand of hair was cut to measure fathers' testosterone levels. After completion of the first research visit, fathers and their partners were requested to fill in a set of online questionnaires at home, including the questionnaires used in the present study. In total, six participants filled in the questionnaires after completion of the intervention period ($n = 3$ baby carrier condition, $n = 3$ baby seat condition).

3.1.2 | Sample

Participants included 80 first-time fathers with a healthy single full-term born infant (53% boys) between 1 and 5 months old ($M = 2.67$ months, $SD = 0.77$). Fathers were mainly born in the Netherlands (92%) and were on average 33.20 years old ($SD = 5.36$, range 25–56). Fathers were fluent in speaking and writing Dutch and cohabited with the biological mother ($M = 30.92$ years, $SD = 4.43$) and infant. One participant was not biologically related to the infant but cohabited with the mother from mid-pregnancy onward. Fathers' educational level was as follows: 29% completed higher university education (e.g., master's degree), 37% completed higher professional education (e.g., bachelor's degree), 31% completed higher vocational education, and 3% completed lower vocational education. Two fathers reported no working hours during the time of the research visit ($n = 2$, no reason provided). Fathers worked on average 38.77 h ($SD = 5.26$) per week during the time of the research visit. Fathers were excluded from participation in case of neurological disorders, endocrine diseases, psychiatric disorders, cardiovascular diseases, use of psychoactive medications, excessive alcohol or drugs intake, use of hard drugs within the past 3 months, MRI contraindications (e.g., metallic foreign objects), using a baby carrier for more than 5 h per week at the time of inclusion (relevant for other research questions), or having an injury that could hinder use of a baby carrier. Recruitment difficulties resulted in the inclusion of seven participants who met an exclusion criterion (MRI contraindications, $n = 3$; diabetes and use of medication potentially interfering with the endocrine system [metformin], $n = 1$; ADHD, $n = 1$; cardiovascular disease, $n = 1$; healthy infant born after 36 weeks and 6 days of gestation, $n = 1$) (see also Lotz et al. 2021). Fathers were recruited via municipal databases of cities in the Western

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

part of the Netherlands, infant welfare centers, midwife practices, and social media. Travelling expenses were covered, and fathers received financial compensation for participation in the study. The study was approved by the Ethics Committees of the Leiden University Medical Centre and the Department of Education and Child studies at Leiden University. Both parents provided written informed consent prior to the start of the study.

3.1.3 | Measures

3.1.3.1 | Hair Testosterone Levels. To measure fathers' testosterone levels in hair, the same procedures were followed as described in Study 1. Hair strands were shipped to Dresden Labservices GMBH for further analysis, and quantification of hair testosterone concentrations was achieved using LC–MS/MS (Gao et al. 2013). A random selection of 41% of hair segments was assayed in duplicate, with an average intraassay coefficient of variation of 19%. Removal of the highest coefficient resulted in an average intraassay coefficient of 16%. For hair segments that were assayed in duplicate, the average value of the two assays was included in further analyses. Fathers' testosterone levels were not related to age of the father or child. As in Study 1, the distribution was positively skewed, and therefore, a log transformation was performed to approach a normal distribution.

3.1.3.2 | Time Spent With Child. The amount of time fathers spent with their child was measured using the same online questionnaire and procedures as in Study 1. Inspection of z-values indicated one outlier ($z > 3.29$) that was winsorized. As in Study 1, a mean score was computed to indicate the daily number of hours fathers spent with their child.

3.1.3.3 | Involvement in Childcare and Household Tasks. Fathers' involvement in childcare and household tasks was measured with the TDQ, following the same procedures as in Study 1. Consistent with Study 1, fathers' mean self-report scores were significantly higher than mothers' scores for them, $t(69) = 8.29$, $p < 0.001$, $d = 0.99$, but scores of fathers and mothers were strongly correlated ($r = 0.67$, $p < 0.001$). Therefore, z-scores of both mother and father reports were averaged into one score indicating paternal involvement in childcare and household tasks. Cronbach's alpha was 0.84 (combined father and mother reports).

3.1.3.4 | Attitudes Toward Child Gender Stereotypes. Fathers' attitudes toward child gender stereotypes were measured with the GSOQ, following the same procedures as in Study 1. Cronbach's alpha was 0.84. As in Study 1, a mean score was calculated and included in further analysis.

3.1.3.5 | Missing Data. Rates of missing data per variable were as follows: 31.6% hair testosterone levels (due to baldness, insufficient hair material or testosterone levels that could not be detected), 6.3% time spent with child, 1.3% involvement in childcare and household tasks, and 7.6% child gender stereotypes. Missing value analysis showed that Little MCAR (Little 1988) was not significant, $\chi^2(52) = 63.75$, $p = 0.127$, indicating that data were missing completely at random. Prior to conducting the statistical analyses, missing values were imputed through predictive mean matching using the MICE (Multivariate Imputation by Chained Equations) package (Van Buuren and Groothuis-Oudshoorn 2011), which resulted in 50 imputed datasets.

3.1.4 | Data Analysis Plan

3.1.4.1 | Main Analysis. Preliminary analyses were performed in SPSS version 27.0, and mediation analyses were performed in R using the package lavaan, following the same approach as described in Study 1.

3.1.4.2 | Power Analysis. The software program G*Power was used to calculate power for the first three paths of the mediation models (see Figures 1 and 2). As described in the preregistration, we conducted a power analysis for Study 2 using the effect sizes obtained in Study 1. Alpha was set at 0.05, and sample size was set at 80. In Study 1, the pooled correlation between fathers' testosterone levels and paternal involvement in childcare and household tasks was $r = 0.30$. In Study 2, an effect size of $r = 0.30$ would result in a power of 0.78. Based on the results of Study 1, Study 2 only had enough power to test the association between fathers' testosterone levels and their involvement in childcare and household tasks. Nevertheless, we tested all associations to provide a more complete picture of the data.

3.1.4.3 | Sensitivity Analyses. Sensitivity analyses were conducted to assess the robustness of the results. First, sensitivity analyses were conducted with involvement in childcare and household tasks reported by mothers and fathers separately. Second, sensitivity analyses were performed excluding participants who completed the questionnaires after completion of the intervention ($n = 6$). Third, sensitivity analyses were performed excluding participants who met any exclusion criterion ($n = 7$).

3.2 | Results

Table 2 presents the descriptive statistics and Pearson correlations based on complete data and imputed data. Involvement in childcare and household tasks was positively correlated with the amount of time spent with the child ($r_{pooled} = 0.31$, $p = 0.008$), indicating that fathers who were more involved in household and childcare tasks spent more time with their child. No other significant correlations were observed in the complete or imputed dataset. Child sex and child age were not related to the variables of interest.

3.2.1 | Testosterone Levels, Child Stereotypes, and Amount of Time Spent With Child

The model testing whether the association between fathers' testosterone levels and the amount of time spent with their child was mediated by child gender stereotypes showed a good fit, $\chi^2(3) = 5.06$, $p = 0.168$, CFI = 1.00, RMSEA < 0.001. Fathers' testosterone levels were not directly related to the amount of time spent with their child, $\beta = 0.13$, $p = 0.300$. Fathers' testosterone levels were not significantly associated with attitudes toward child gender stereotypes, $\beta = 0.11$, $p = 0.362$. Fathers' attitudes

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Descriptive statistics and correlations between variables based on complete cases and imputed data.

Variables	*M* (*SD*)	Min	Max	*N*	1.	2.	3.	4.
1. Testosterone levels (in hair)	0.66 (0.72)	0.01	3.85	54	—	−0.14	0.06	0.19
2. Involvement in childcare and household tasks	−0.00 (0.94)	−2.29	2.64	77	−0.15	—	0.31*	−0.13
3. Amount of time spent with child (hours)	5.62 (2.62)	1.57	14.24	74	0.09	0.30*	—	−0.21
4. Attitudes toward child gender stereotypes	1.76 (0.46)	1.05	3.47	73	0.13	−0.11	−0.22	—

Note: Involvement in childcare and household tasks is the average of the standardized scores of mother and father reports. Hair testosterone concentrations are presented inpg/mL. Correlations based on complete cases are presented above the diagonal, and pooled correlations for the 50 imputed datasets are presented below the diagonal.
*$p < 0.01$.

toward child stereotypes were not significantly associated with the amount of time spent with their child, $\beta = -0.23$, $p = 0.064$. Fathers' testosterone levels were not indirectly related to involvement via child gender stereotypes, $\beta = -0.03$, $p = 0.527$ (see Figure 1).

3.2.2 | Testosterone Levels, Child Stereotypes, and Involvement in Childcare and Household Tasks

The model testing whether the association between fathers' testosterone levels and involvement in childcare and household tasks was mediated by child gender stereotypes showed a good fit, $\chi^2(3) = 3.22$, $p = 0.359$, CFI = 1.00, RMSEA < 0.001. Fathers' testosterone levels were not directly related to involvement in childcare and household tasks, $\beta = -0.15$, $p = 0.222$. Fathers' testosterone levels were not significantly associated with attitudes toward child gender stereotypes, $\beta = 0.11$, $p = 0.351$. Fathers' attitudes toward child stereotypes were not significantly associated with involvement in childcare and household tasks, $\beta = -0.10$, $p = 0.427$. Fathers' testosterone levels were not indirectly related to involvement in childcare and household tasks via child gender stereotypes, $\beta = -0.01$, $p = 0.723$ (see Figure 2).

3.2.3 | Sensitivity Analyses

Preregistered sensitivity analyses were conducted with involvement in childcare and household tasks reported by mothers and fathers separately. Furthermore, we conducted sensitivity analyses excluding participants ($n = 6$) who filled in the questionnaires after the intervention (i.e., not during the time period included in the present study). Results did not lead to different conclusions. Finally, we ran sensitivity analyses excluding participants who met one of the exclusion criteria ($n = 7$). Results indicated a negative association between fathers' attitudes toward child gender stereotypes and the amount of time fathers spent with their child, $\beta = -0.29$, $p = 0.018$.

3.2.4 | Explorative Meta-Analytic Associations

Although not preregistered, we also combined the reported associations in Study 1 and Study 2 meta-analytically. Combined effect sizes indicated that fathers' testosterone levels were not meta-analytically associated with involvement in childcare or household tasks ($r = 0.08$, $p = 0.734$, 95% CI [−0.37, 0.53]) or the amount of time spent with the child ($r = -0.01$, $p = 0.954$, 95% CI [−0.20, 0.19]). Furthermore, fathers' testosterone levels were not meta-analytically associated with fathers' attitudes toward child gender stereotypes ($r = 0.14$, $p = 0.107$, 95% CI [−0.29, 0.30]), and neither were attitudes toward child gender stereotypes meta-analytically associated with involvement in childcare and household tasks ($r = -0.08$, $p = 0.357$, 95% CI [−0.24, 0.09]) or the amount of time spent with the child ($r = -0.05$, $p = 0.786$, 95% CI [−0.39, 0.30]).

4 | Discussion

The present preregistered study investigated whether paternal testosterone levels in hair predicted the amount of time fathers

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

spent with their child and fathers' involvement in childcare and household tasks. Furthermore, we investigated whether associations were mediated by fathers' attitudes toward child gender stereotypes. Associations were tested in two studies, with one sample serving as an exploratory sample and one sample as a test sample. Unexpectedly, findings from the exploratory sample showed that fathers with higher testosterone levels were more involved in childcare and household tasks. However, this finding did not emerge in the test sample. Furthermore, in both samples, fathers' testosterone levels were not related to time spent with the child. Finally, fathers' testosterone levels were not related to fathers' attitudes toward child gender stereotypes, and neither were associations between testosterone levels and time spent with the child or involvement in childcare and household tasks mediated by attitudes toward child gender stereotypes.

We found that higher testosterone levels among fathers in Study 1 (the explorative study) were related to greater involvement in childcare and household tasks, which is in contrast with the idea that lower testosterone levels facilitate investment in offspring and the partner relationship (Gettler et al. 2014; Gray and Campbell 2009; Wingfield et al. 1990). Previous studies reported that fathers with lower testosterone levels were more involved in household tasks (Edelstein et al. 2017) and child caregiving (Kuo et al. 2018; Mascaro et al. 2013). A possible explanation for the divergent findings is that in previous studies testosterone levels were measured during pregnancy (Edelstein et al. 2017), the first days after the infant was born (Kuo et al. 2018), or when children were 1- to 2-year old (Mascaro et al. 2013). In contrast, in Study 1, fathers' testosterone levels were measured when children were on average 7 months old. Associations between fathers' testosterone levels and involvement in childcare and household tasks may vary depending on the stage of fatherhood, although this suggestion awaits further investigation. Having said that, in Study 2 (the test study), in which testosterone levels were measured when children were on average 3 months old, no significant associations were found between fathers' testosterone levels and involvement in childcare and household tasks. Future research may use longitudinal design with testosterone measurements that begin in the prenatal period and continue during the first year postpartum. This would allow for a better understanding of how fathers' testosterone trajectories relate to fathering outcomes over time.

In both samples, we found no evidence that fathers' testosterone levels were related to the amount of time spent with the child. Of note, the present study measured testosterone levels in hair, which provides a more stable and long-term indication of fathers' testosterone levels. This approach differs from previous studies (e.g., Corpuz et al. 2021; Edelstein et al. 2017; Gray et al. 2004; Kuo et al. 2018), which measured testosterone levels in saliva. Although testosterone levels in hair were found to be moderately associated with testosterone levels in saliva (Chen et al. 2019; Stern et al. 2022), saliva testosterone samples also capture state fluctuations, including reactivity to environmental stimuli (Grotzinger et al. 2018; De Vries et al. 2019). Multiple samples over several days are needed to increase the reliability of salivary testosterone measures, although this also increases participation burden. Future studies may examine whether testosterone levels measured in hair and based on multiple saliva samples are differentially related to paternal involvement.

In both Study 1 and Study 2, there were no significant associations between paternal testosterone levels and fathers' attitudes toward child gender stereotypes, and fathers' attitudes toward child gender stereotypes were unrelated to time spent with their child and involvement in childcare and household tasks. This was the first study to test these associations. Our sample included relatively highly educated first-time fathers living in the Netherlands. In 2022, the Netherlands was ranked third in the European Union on the Gender Equality Index (European Institute for Gender Equality 2022). Likewise, fathers in the present study scored relatively low on gender-stereotype attitudes (i.e., mean scores were 1.76 and 1.59 on a 5-point rating scale). There were no fathers with scores in the high end of the rating scale (i.e., highest scores were 2.74 and 3.47 in Study 1 and Study 2, respectively). Thus, significant associations may be absent due to a floor effect. Samples of fathers with different educational and cultural backgrounds and samples with fathers who report more gender-stereotype attitudes may show different results. Furthermore, future studies may build on our findings by exploring how different dimensions of gender attitudes—such as those related to adults versus children—are associated with paternal testosterone levels and fathers' time spent with their child and their involvement in childcare and household tasks.

Strengths of the present study comprise the inclusion of both an exploration and test sample, as this enhanced the scientific rigor. The inability to replicate the relation between fathers' hair testosterone levels and involvement in childcare and household tasks signifies that this might have been a mere chance finding and highlights the importance of replication before strong conclusions are drawn. Furthermore, we preregistered our study on the Open Science Framework (OSF), which maximizes the transparency, reproducibility, and credibility of our findings. A final strength is that we focused on two indicators of paternal involvement, namely fathers' involvement in childcare and household activities (self-report and partner-report) and fathers' time spent with the child. There are also limitations that should be noted. First, we used self-reports, which may be prone to social desirability bias. Second, the TDQ and GSOQ were developed for the purpose of the present study. These questionnaires showed acceptable to good internal consistency but await validation. Third, fathers were in their first year of parenthood, and we could not examine whether trajectories of testosterone levels are related to involvement in childcare and household tasks over the years. Related to this issue, a notable and unfortunate shortcoming of the field is the lack of a standard measure to assess paternal involvement. As a result, the current and previous studies used different methods to assess paternal involvement. Fourth, there was a difference in timing between the first study (mean of 6.73 months postpartum) and the second study (mean of 2.67 months postpartum), related to the RCT designs of the studies in which fathers participated. Although this may limit the direct comparability of study results, it allows for a broader understanding of how testosterone levels vary and may be differentially related to fathering outcomes across different phases in the postnatal period. A final limitation is that results should be interpreted with caution, also considering the relatively high amount of missing testosterone data, which we addressed through imputation. Furthermore, it remains unclear what caused the relatively high average intraassay coefficient in Study 2. This may have obscured potential associations between

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

paternal testosterone levels and the time fathers spent with their child.

In conclusion, results from Study 1 showed a positive relationship between fathers' testosterone levels and self- and partner-reported involvement in childcare and household tasks, but this finding was not replicated in Study 2. In both studies, associations between fathers' testosterone levels and the time spent with their child were absent. Furthermore, in both samples, we found no evidence for mediation by fathers' attitudes toward child gender stereotypes. Our findings should be considered preliminary, and further insight into the validity and generalizability of our findings should be achieved through replication in samples that are more inclusive in terms of socioeconomic status and ethnicity. Yet, we are aware that recruitment of fathers for research on parenting and child development is complicated and often results in an overrepresentation of fathers from higher socioeconomic strata. Hence, there is a need to use effective recruitment strategies that promote the inclusive involvement of fathers in parenting research (Deneault et al., 2025).

Author Contributions

Annemieke M. Witte: methodology, formal analysis, investigation, data curation, writing – original draft, visualization project administration. **Noor de Waal**: formal analysis, data curation, writing – review and editing. **Martine W. F. T. Verhees**: data curation, writing – review and editing. **Anna M. Lotz**: investigation, writing – review and editing, project administration. **Marinus H. van Ijzendoorn**: methodology, writing – review and editing, funding acquisition. **Marian J. Bakermans-Kranenburg**: conceptualization, methodology, validation, writing – review and editing, supervision, funding acquisition.

Acknowledgments

We would like to thank all families who participated in this study. We also thank the Father Trials research team for their contribution to data collection.

Ethics Statement

Study 1 was approved by the Ethics Committees of the Leiden University Medical Centre. Study 2 was approved by the Ethics Committees of the Leiden University Medical Centre and the Department of Education and Child studies at Leiden University.

Consent

Prior to the start of the study, all parents provided written informed consent.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

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

References

van Anders, S. M., R. M. Tolman, and B. L. Volling. 2012. "Baby Cries and Nurturance Affect Testosterone in Men." *Hormones and Behavior* 61, no. 1: 31–36. https://doi.org/10.1016/j.yhbeh.2011.09.012.

Atkinson, L., S. A. Silverio, D. Bick, and V. Fallon. 2021. "Relationships Between Paternal Attitudes, Paternal Involvement, and Infant-Feeding Outcomes: Mixed-Methods Findings From a Global On-Line Survey of English-Speaking Fathers." *Maternal & Child Nutrition* 17, no. S1: e13147. https://doi.org/10.1111/mcn.13147.

Azur, M. J., E. A. Stuart, C. Frangakis, and P. J. Leaf. 2011. "Multiple Imputation by Chained Equations: What Is It and How Does It Work?" *International Journal of Methods in Psychiatric Research* 20, no. 1: 40–49. https://doi.org/10.1002/mpr.329.

Bakermans-Kranenburg, M. J., A. Lotz, K. Alyousefi-van Dijk, and M. van IJzendoorn. 2019. "Birth of a Father: Fathering in the First 1,000 Days." *Child Development Perspectives* 13, no. 4: 247–253. https://doi.org/10.1111/cdep.12347.

Bakermans-Kranenburg, M. J., M. W. F. T. Verhees, A. M. Lotz, K. Alyousefi-van Dijk, and M. H. van IJzendoorn. 2022. "Is Paternal Oxytocin an Oxymoron? Oxytocin, Vasopressin, Testosterone, Oestradiol and Cortisol in Emerging Fatherhood." *Philosophical Transactions of the Royal Society B: Biological Sciences* 377, no. 1858: 20210060. https://doi.org/10.1098/rstb.2021.0060.

Bem, S. L. 1981. "Gender Schema Theory: A Cognitive Account of Sex Typing." *Psychological Review* 88: 354–364. https://doi.org/10.1037/0033-295x.88.4.354.

Bem, S. L. 1983. "Gender Schema Theory and Its Implications for Child Development: Raising Gender-Aschematic Children in a Gender-Schematic Society." *Signs* 8: 598–616.

Bos, P. A., E. J. Hermans, E. R. Montoya, N. F. Ramsey, and J. van Honk. 2010. "Testosterone Administration Modulates Neural Responses to Crying Infants in Young Females." *Psychoneuroendocrinology* 35, no. 1: 114–121. https://doi.org/10.1016/j.psyneuen.2009.09.013.

Bulanda, R. E. 2004. "Paternal Involvement With Children: The Influence of Gender Ideologies." *Journal of Marriage and Family* 66, no. 1: 40–45. https://doi.org/10.1111/j.0022-2455.2004.00003.x.

Cano, T., F. Perales, and J. Baxter. 2019. "A Matter of Time: Father Involvement and Child CognitiveOutcomes." *Journal of Marriage and Family* 81, no. 1: 164–184. https://doi.org/10.1111/jomf.12532.

Chen, Z., Q. Zhang, S. Chen, W. Wang, G. Liu, and H. Deng. 2019. "Determination, Intercorrelation and Intraindividual Stability of Five Steroids in Hair, Saliva and Urine Among Chinese College Students." *Steroids* 149: 108418. https://doi.org/10.1016/j.steroids.2019.05.010.

Corpuz, R., S. D'Alessandro, and G. K. Collom. 2021. "The Postnatal Testosterone Rebound in First-time Fathers and the Quality and Quantity of Paternal Care." *Developmental Psychobiology* 63, no. 5: 1415–1427. https://doi.org/10.1002/dev.22064.

Cox, J. L., J. M. Holden, and R. Sagovsky. 1987. "Detection of Postnatal Depression: Development of the 10-Item Edinburgh Postnatal Depression Scale." *The British Journal of Psychiatry* 150, no. 6: 782–786. https://doi.org/10.1192/bjp.150.6.782.

Craig, L. 2006. "Does Father Care Mean Fathers Share? A Comparison of How Mothers and Fathers in Intact Families Spend Time With Children." *Gender and Society* 20, no. 2: 259–281. http://www.jstor.org/stable/27640883.

Craig, L., and K. Mullan. 2011. "How Mothers and Fathers Share Childcare: A Cross-National Time-Use Comparison." *American Sociological Review* 76, no. 6: 834–861. https://doi.org/10.1177/0003122411427673.

De Vries, E. E., L. D. van der Pol, H. J. Vermeer, M. G. Groeneveld, T. Fiers, and J. Mesman. 2019. "Testosterone and Fathers' Parenting Unraveled: Links With the Quantity and Quality of Father-Child Interactions." *Adaptive Human Behavior and Physiology* 5: 297–316. https://doi.org/10.1007/s40750-019-00118-z.

Deneault, A. A., J. S. Feldman, A. Aytuglu, et al. 2025. "Best Practices for Recruiting and Retaining Fathers in Parenting Research: Insights From Fathering Researchers." *Parenting* 1–27. https://doi.org/10.1080/15295192.2025.2509515.

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Dettenborn, L., C. Muhtz, N. Skoluda, et al. 2012. "Introducing a Novel Method to Assess Cumulative Steroid Concentrations: Increased Hair Cortisol Concentrations Over 6 Months in Medicated Patients With Depression." *Stress* 15, no. 3: 348–353. https://doi.org/10.3109/10253890.2011.619239.

Edelstein, R. S., W. J. Chopik, and E. L. Kean. 2011. "Sociosexuality Moderates the Association Between Testosterone and Relationship Status in Men and Women." *Hormones and Behavior* 60, no. 3: 248–255. https://doi.org/10.1016/j.yhbeh.2011.05.007.

Edelstein, R. S., W. J. Chopik, D. E. Saxbe, B. M. Wardecker, A. C. Moors, and O. P. LaBelle. 2017. "Prospectiveand Dyadic Associations Between Expectant Parents' prenatal Hormone Changes Andpostpartum Parenting Outcomes." *Developmental Psychobiology* 59, no. 1: 77–90. https://doi.org/10.1002/dev.21469.

Edelstein, R. S., S. M. van Anders, W. J. Chopik, K. L. Goldey, and B. M. Wardecker. 2014. "Dyadicassociations Between Testosterone and Relationship Quality in Couples." *Hormones and Behavior* 65, no. 4: 401–407. https://doi.org/10.1016/j.yhbeh.2014.03.003.

Ehrhardt, A. A., R. Epstein, and J. Money. 1968. "Fetal Androgens and Female Gender Identity in the Early-Treated Adrenogenital Syndrome." *Johns Hopkins Medical Journal* 122, no. 3: 160–167.

Ehrhardt, A. A., G. C. Grisanti, and H. F. Meyer-Bahlburg. 1977. "Prenatal Exposure to Medroxyprogesterone Acetate (MPA) in Girls." *Psychoneuroendocrinology* 2, no. 4: 391–398. https://doi.org/10.1016/0306-4530(77)90010-5.

Epstein, N. B., L. M. Baldwin, and D. S. Bishop. 1983. "The McMaster Family Assessment Device." *Journal of Marital and Family Therapy* 9, no. 2: 171–180. https://doi.org/10.1111/j.1752-0606.1983.tb01497.x.

European Institute for Gender Equality. 2022. "Gender Equality Index." https://eige.europa.eu/gender-equality-index/2022/country/NL#:~:text=Progress%20in%20gender%20equality,ranking%20has%20remained%20the%20same.

Gao, W., T. Stalder, P. Foley, M. Rauh, H. Deng, and C. Kirschbaum. 2013. "Quantitative Analysis of Steroid Hormones in Human Hair Using a Column-Switching LC–APCI–MS/MS Assay." *Journal of Chromatography B* 928: 1–8. https://doi.org/10.1016/j.jchromb.2013.03.008.

Gettler, L. T., T. W. McDade, A. B. Feranil, and C. W. Kuzawa. 2011. "Longitudinal Evidence That Fatherhood Decreases Testosterone in Human Males." *Proceedings of the National Academy of Sciences of the United States of America* 108, no. 39: 16194–16199. https://doi.org/10.1073/pnas.1105403108.

Gettler, L. T. 2014. "Applyingsocioendocrinology to Evolutionary Models: Fatherhood and Physiology." *Evolutionary Anthropology: Issues, News, and Reviews* 23, no. 4: 146–160. https://doi.org/10.1002/even.21412.

Gray, P. B., and B. C. Campbell. 2009. "Human Male Testosterone, Pair-Bonding, and Fatherhood." In *Endocrinology of Social Relationships*, edited by P. T. Ellison and P. B. Gray, 270–293. Harvard University Press.

Gray, P. B., B. C. Campbell, F. W. Marlowe, S. F. Lipson, and P. T. Ellison. 2004. "Social Variables Predict Between-Subject but Not Day-to-Day Variation in the Testosterone of US Men." *Psychoneuroendocrinology* 29, no. 9: 1153–1162. https://doi.org/10.1016/j.psyneuen.2004.01.008.

Grebe, N. M., R. E. Sarafin, C. R. Strenth, and S. Zilioli. 2019. "Pair-Bonding, Fatherhood, and the Role of Testosterone: A Meta-Analytic Review." *Neuroscience & Biobehavioral Reviews* 98: 221–233. https://doi.org/10.1016/j.neubiorev.2019.01.010.

Grotzinger, A. D., F. D. Mann, M. W. Patterson, J. L. Tackett, E. M. Tucker-Drob, and K. P. Harden. 2018. "Hair and Salivary Testosterone, Hair Cortisol, and Externalizing Behaviors in Adolescents." *Psychological Science* 29, no. 5: 688–699. https://doi.org/10.1177/0956797617742981.

Hines, M. 2011. "Prenatal Endocrine Influences on Sexual Orientation and on Sexually Differentiated Childhood Behavior." *Frontiers in Neuroendocrinology* 32, no. 2: 170–182. https://doi.org/10.1016/j.yfrne.2011.02.006.

Horstman, L. I., M. M. Riem, K. Alyousefi-vanDijk, A. M. Lotz, and M. J. Bakermans-Kranenburg. 2022. "Fathers'involvement in Early Childcare Is Associated With Amygdala Resting-stateconnectivity." *Social Cognitive and Affective Neuroscience* 17, no. 2: 198–205. https://doi.org/10.1093/scan/nsab086.

Hu, L., and P. M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6, no. 1: 1–55. https://doi.org/10.1080/10705519909540118.

Julian, T., and P. C. McKenry. 1989. "Relationship of Testosterone to Men's family Functioning at Mid-life: a Research Note." *Aggressive Behavior* 15, no. 4: 281–289. https://doi.org/10.1002/ab.2480150403.

Keizer, R. 2015. "Which Men Become Involved Fathers? The Impact of Men's Own Attitudes on Paternal Involvement in the Netherlands." *International Review of Sociology* 25, no. 3: 359–372. https://doi.org/10.1080/03906701.2015.1078531.

Kuo, P. X., J. M. Braungart-Rieker, J. E. B. Lefever, M. S. Sarma, M. O'Neill, and L. T. Gettler. 2018. "Fathers' cortisol and Testosterone in the Days Around Infants' births Predictlater Paternal Involvement." *Hormones and Behavior* 106: 28–34. https://doi.org/10.1016/j.yhbeh.2018.08.011.

Lamminmäki, A., M. Hines, T. Kuiri-Hänninen, L. Kilpeläinen, L. Dunkel, and U. Sankilampi. 2012. "Testosterone Measured in Infancy Predicts Subsequent Sex-Typed Behavior in Boys and in Girls." *Hormones and Behavior* 61, no. 4: 611–616. https://doi.org/10.1016/j.yhbeh.2012.02.013.

Little, R. J. 1988. "A Test of Missingcompletely at Random for Multivariate Data With Missing Values." *Journal of the American Statistical Association* 83, no. 404: 1198–1202. https://doi.org/10.1080/01621459.1988.10478722.

Lotz, A. M., M. W. Verhees, L. I. Horstman, et al. 2021. "Exploring Thehormonal and Neural Correlates of Paternal Protective Behavior to Theirinfants." *Developmental psychobiology* 63, no. 5: 1358–1369. https://doi.org/10.1002/dev.22055.

Luberti, F. R., and J. M. Carré. 2024. "Testosterone's Role in Modulating Human Behaviors Relevant to Mating and Parenting." *Frontiers in Neuroendocrinology* 72: 101112. https://doi.org/10.1016/j.yfrne.2023.101112.

Mascaro, J. S., P. D. Hackett, and J. K. Rilling. 2013. "Testicular Volume Is Inversely Correlated With Nurturing-Related Brain Activity in Human Fathers." *Proceedings of the National Academy of Sciences of the United States of America* 110, no. 39: 15746–15751. https://doi.org/10.1073/pnas.1305579110.

McGovern, J. M., and S. A. Meyers. 2002. "Relationships Between Sex-Role Attitudes, Division of Household Tasks, and Marital Adjustment." *Contemporary Family Therapy* 24: 601–618. https://doi.org/10.1023/A:1021225313735.

McIntyre, M., S. W. Gangestad, P. B. Gray, et al. 2006. "Romantic Involvement Often Reduces Men's Testosterone Levels—But Not Always: The Moderating Role of Extrapair Sexual Interest." *Journal of Personality and Social Psychology* 91: 642–651. https://doi.org/10.1037/0022-3514.91.4.642.

Meijer, W. M., M. H. van Ijzendoorn, and M. J. Bakermans-Kranenburg. 2019. "Challenging the Challenge Hypothesis on Testosterone in Fathers: Limited Meta-Analytic Support." *Psychoneuroendocrinology* 110: 104435. https://doi.org/10.1016/j.psyneuen.2019.104435.

Milkie, M. A., S. M. Bianchi, M. J. Mattingly, and J. P. Robinson. 2002. "Gendered Division of Childrearing: Ideals, Realities, and the Relationship to Parental Well-Being." *Sex Roles* 47, no. 1: 21–38. https://doi.org/10.1023/A:1020627602889.

Muller, M. N., F. W. Marlowe, R. Bugumba, and P. T. Ellison. 2009. "Testosterone and Paternal Care in East Africanforagers and Pastoralists." *Proceedings of the Royal Society B:Biological Sciences* 276, no. 1655: 347–354. https://doi.org/10.1098/rspb.2008.1028.

Riem, M. M., A. M. Lotz, L. I. Horstman, et al. 2021. "A Soft Baby Carrier Intervention Enhancesamygdala Responses to Infant Crying in Fathers:

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

a Randomized Controlledtrial." *Psychoneuroendocrinology* 132: 105380. https://doi.org/10.1016/j.psyneuen.2021.105380.

Rilling, J. K., and J. S. Mascaro. 2017. "The Neurobiology of Fatherhood." *Current Opinion in Psychology* 15: 26–32. https://doi.org/10.1016/j.copsyc.2017.02.013.

Rodriguez, C. M., D. A. Granger, and E. M. Leerkes. 2021. "Testosterone Associations With Parents' Child Abuse Risk and At-Risk Parenting: A Multimethod Longitudinal Examination." *Child Maltreatment* 26, no. 1: 50–62. https://doi.org/10.1177/1077559520930819.

Ronay, R., L. van der Meij, J. K. Oostrom, and T. V. Pollet. 2018. "No Evidence for a Relationship Between Hair Testosterone Concentrations and 2D:4D Ratio or Risk Taking." *Frontiers in Behavioral Neuroscience* 12: 30. https://doi.org/10.3389/fnbeh.2018.00030.

Rosseel, Y. 2012. "lavaan: an R Packagefor Structural Equation Modeling." *Journal of Statistical Software* 48, no. 1: 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/. Accessed 22 Aug 2023.

Sarkadi, A., R. Kristiansson, F. Oberklaid, and S. Bremberg. 2008. "Fathers' involvement and Children'sdevelopmental Outcomes: a Systematic Review of Longitudinal Studies." *Acta Paediatrica* 97, no. 2: 153–158. https://doi.org/10.1111/j.1651-2227.2007.00572.x.

Saxbe, D. E., R. S. Edelstein, H. M. Lyden, B. M. Wardecker, W. J. Chopik, and A. C. Moors. 2017. "Fathers' Decline in Testosterone and Synchrony With Partner Testosterone During Pregnancy Predicts Greater Postpartum Relationship Investment." *Hormones and Behavior* 90: 39–47. https://doi.org/10.1016/j.yhbeh.2016.07.005.

Stalder, T., S. Steudte-Schmiedgen, N. Alexander, et al. 2017. "Stress-Related and Basic Determinants of Hair Cortisol in Humans: A Meta-Analysis." *Psychoneuroendocrinology* 77: 261–274. https://doi.org/10.1016/j.psyneuen.2016.12.017.

Stern, J., R. C. Arslan, and L. Penke. 2022. "Stability and Validity of Steroid Hormones in Hair and Saliva Across Two Ovulatory Cycles." *Comprehensive Psychoneuroendocrinology* 9: 100114. https://doi.org/10.1016/j.cpnec.2022.100114.

Storey, A. E., C. J. Walsh, R. L. Quinton, and K. E. Wynne-Edwards. 2000. "Hormonal Correlates of Paternal Responsiveness in New and Expectant Fathers." *Evolution and Human Behavior* 21, no. 2: 79–95. https://doi.org/10.1016/S1090-5138(99)00042-2.

Storey, A. E., D. E. Noseworthy, K. M. Delahunty, S. J. Halfyard, and D. W. McKay. 2011. "The Effects ofsocial Context on the Hormonal and Behavioral Responsiveness of Humanfathers." *Hormones and Behavior* 60, no. 4: 353–361. https://doi.org/10.1016/j.yhbeh.2011.07.001.

Van Buuren, S., and K. Groothuis-Oudshoorn. 2011. "mice: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45, no. 3: 1–67. https://doi.org/10.18637/jss.v045.i03.

Van IJzendoorn, M. H., and M. J. Bakermans-Kranenburg. 2024. *Matters of Significance. Replication, Translation, and Academic Freedom in Developmental Science*. UCL Press.

Verhees, M. W. F. T., A. M. Lotz, M. H. M. de Moor, et al. 2023. "Effects of a Soft Baby Carrier on Fathers' Behavior and Hormones: A Randomized Controlled Trial." *Journal of Child and Family Studies* 33: 1979–1994. https://doi.org/10.1007/s10826-023-02678-x.

Verhees, M. W., M. H. van IJzendoorn, K. Alyousefi-van Dijk, A. M. Lotz, N. de Waal, and M. J. Bakermans-Kranenburg. 2021. "Child Maltreatment Affects Fathers' response to Infant Crying,Not Mediated by Cortisol or Testosterone." *Comprehensive Psychoneuroendocrinology* 8: 100083. https://doi.org/10.1016/j.cpnec.2021.100083.

Wingfield, J. C., R. E. Hegner, M. A. Dufty Jr, and G. F. Ball. 1990. "The "Challenge Hypothesis":Theoretical Implications for Patterns of Testosterone Secretion, Matingsystems, and Breeding Strategies." *The American Naturalist* 136, no. 6: 829–846. https://doi.org/10.1086/285134.

Witte, A. M., M. H. M. de Moor, M. Majdandžić, M. W. F. T. Verhees, M. H. van Ijzendoorn, and M. J. Bakermans-Kranenburg. 2022. "Effects of Oxytocin and Vasopressin Administration on Human Fathers' Sensitive and Challenging Parenting: A Randomized Within-Subject Controlled Trial." *Hormones and Behavior* 142: 105175. https://doi.org/10.1016/j.yhbeh.2022.105175.

Witte, A. M., M. M. E. Riem, N. van der Knaap, M. H. M. de Moor, M. H. van Ijzendoorn, and M. J. Bakermans-Kranenburg. 2022. "The Effects of Oxytocin and Vasopressin Administration on Fathers' Neural Responses to Infant Crying: A Randomized Controlled Within-Subject Study." *Psychoneuroendocrinology* 140: 105731. https://doi.org/10.1016/j.psyneuen.2022.105731.

Witte, A. M., M. J. Bakermans-Kranenburg, M. H. van IJzendoorn, O. Szepsenwol, and D. Shai. 2020. "Predictinginfant–father Attachment: the Role of Pre-and Postnatal Triadic family Allianceand Paternal Testosterone Levels." *Attachment & Human Development* 22, no. 6: 653–667. https://doi.org/10.1080/14616734.2019.1680713.

Yavorsky, J. E., C. M. Kamp Dush, and S. J. Schoppe-Sullivan. 2015. "The Production of Inequality: The Gender Division of Labor across the Transition to Parenthood." *Journal of Marriage and Family* 77, no. 3: 662–679. https://doi.org/10.1111/jomf.12189.

Supporting Information

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

Supplementary Table 1. Items of Task Division Questionnaire—Parent Report (Original item description and English translation).

Supplementary Table 2. Percentage of parents reporting 'neither of us' for items of Task Division Questionnaire. **Supplementary Table 3**. Items of Gender Specific Orientation Questionnaire—Parent Report (Original item description and English translation).

10982302, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dev.70079 by Ku Leuven, Wiley Online Library on [03/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License