

Patty X. Kuo<sup>1</sup>  
 Ekjyot K. Saini<sup>2</sup>  
 Elizabeth Thomason<sup>1</sup>  
 Oliver C. Schultheiss<sup>3</sup>  
 Richard Gonzalez<sup>1,4</sup>  
 Brenda L. Volling<sup>1,4</sup>

<sup>1</sup>Department of Psychology  
 University of Michigan  
 Ann Arbor, MI, 48109  
 E-mail: pkuo@umich.edu

<sup>2</sup>Human Development and Family Studies  
 Auburn University  
 Auburn, AL

<sup>3</sup>Department of Psychology and Sport  
 Sciences  
 Friedrich-Alexander University  
 Erlangen, Germany

<sup>4</sup>Center for Human Growth and  
 Development  
 University of Michigan  
 Ann Arbor, MI

# Individual Variation in Fathers' Testosterone Reactivity to Infant Distress Predicts Parenting Behaviors With Their 1-Year-Old Infants

**ABSTRACT:** Positive father involvement is associated with positive child outcomes. There is great variation in fathers' involvement and fathering behaviors, and men's testosterone (T) has been proposed as a potential biological contributor to paternal involvement. Previous studies investigating testosterone changes in response to father–infant interactions or exposure to infant cues were unclear as to whether individual variation in T is predictive of fathering behavior. We show that individual variation in fathers' T reactivity to their infants during a challenging laboratory paradigm (Strange Situation) uniquely predicted fathers' positive parenting behaviors during a subsequent father–infant interaction, in addition to other psychosocial determinants of paternal involvement, such as dispositional empathy and marital quality. The findings have implications for understanding fathering behaviors and how fathers can contribute to their children's socioemotional development. © 2015 Wiley Periodicals, Inc. *Dev Psychobiol* 58:303–314, 2016.

**Keywords:** hormones; parental care; human

## INTRODUCTION

Fathering behavior in humans has typically been understood from an ontogenetic perspective. According to Belsky's (1984) Multiple Determinants of Parenting Model, multiple psychological and socio-contextual forces shape fathering behavior, including parent personality and marital quality (Belsky & Jaffee, 2006; Cummings, 2010; Parke, 1996). At the same time, considerable empirical evidence demonstrates that biological bases of parenting effort exist across mammalian and avian species (Rilling, 2013), suggesting that fathering behavior can also be examined from a phylogenetic perspective. According to the Challenge Hypothesis (Wingfield, Hegner, Dufty, & Ball, 1990), decreased

testosterone (T) is believed to specifically facilitate paternal behavior.

An underlying assumption of existing research on human fathering and T is that decreases in men's T facilitates infant care—as proposed by the Challenge Hypothesis (Wingfield et al., 1990)—but so far no study has directly tested this assumption. Instead, studies have only focused on whether interactions with infants produce long (Gettler, McDade, Feranil, & Kuzawa, 2011b) or short-term changes in fathers' T (Gettler, McDade, Agustin, & Kuzawa, 2011; Gray, Parkin, & Samms-Vaughan, 2007; Storey, Noseworthy, Delahunty, Halfyard, & McKay, 2011). Moreover, these studies on short-term changes in T as a result of infant interaction have only used play-based interaction paradigms, and findings have been inconsistent across studies—play with infants rarely elicits significant changes in T.

Because crying is the primary modality of infant communication, situations that elicit men's empathic concern for their infants rather than a playful father–infant

Manuscript Received: 20 February 2015

Manuscript Accepted: 5 October 2015

Correspondence to: Patty X. Kuo or Brenda L. Volling

Article first published online in Wiley Online Library

(wileyonlinelibrary.com): 24 October 2015

DOI 10.1002/dev.21370 • © 2015 Wiley Periodicals, Inc.

interaction may more readily elicit T reactivity. Understanding the social context is crucial for understanding T reactivity as T should decrease to facilitate sensitive and nurturing parenting behavior (van Anders, Goldey, & Kuo, 2011). Indeed, fathers show greater neural activation when hearing infant cries compared to laughter, whereas non-parents show the opposite response (Seifritz et al., 2003), suggesting that fathers experience stronger physiological changes when seeing their infants in distress compared to playing with their infants. Infant cries specifically activate regions important for attention and emotion that should support a father's ability to react appropriately to the situational context (Seifritz et al., 2003; Swain, Dayton, Kim, Tolman, & Volling, 2014; Swain, Lorberbaum, Kose, & Strathearn, 2007). Thus, understanding the social context surrounding father–infant interaction is paramount.

For parents, infant cries are strong stimuli, and can often elicit multiple types of emotional responses, including empathy, annoyance, or aggravation. Individual differences in T reactivity to infant distress appears to be modulated by empathetic feelings and behavior. Men's T is lower when men feel empathetic when hearing recorded infant cries and engage in nurturing behavior with infant dolls, but men's T increases in the absence of empathetic feelings and behavior (Fleming, Corter, Stallings, & Steiner, 2002; Storey, Walsh, Quinton, & Wynne-Edwards, 2000; van Anders, Tolman, & Volling, 2012). The absence of empathetic feelings and behavior in response to infant cries is highly problematic because lack of empathy leads to thoughts of infant abuse (Fairbrother, Barr, Pauwels, Brant, & Green, 2015). Because increases in T facilitate aggressive behaviors (Carré, McCormick, & Hariri, 2011), and are associated with decreased empathy (Fleming et al., 2002; Hermans, Putman, & van Honk, 2006), increased T in response to infant distress could, at best, interfere with sensitive parenting behavior and, at worst, lead to infant abuse. Taken together, infant distress may trigger certain emotional responses that are accompanied by the corresponding hormonal response (empathy with decreased T, aggravation with increased T) that could facilitate matching behaviors (empathy with sensitive behavior, aggravation with negative, intrusive behavior). But, whether T reactivity to infant cries actually predicts parenting behavior is unknown.

All previous T-infant distress studies used stimulated infant cries instead of observing fathers with their own infants (Fleming et al., 2002; Storey et al., 2000; van Anders et al., 2012). Given that fathers activate more strongly to their own infants compared to other infants in emotion regulation and empathy-related neural circuits (Kuo, Carp, Light, & Grewen, 2012; Swain et al., 2014), men may be more motivated to soothe, protect, and care for their own infant given the intense nature of

the parental attachment bond. Hence, the current study examined fathers' T reactivity to observing their own infants in a distress-evoking laboratory paradigm.

Another obvious gap within the father-T literature is that in adopting a phylogenetic approach, it ignores crucial ontogenetic determinants of fathering, as proposed by the Multiple Determinants of Parenting Model (Belsky, 1984). Fathering is a multiply determined process that operates from a proximal to distal manner (Belsky, 1984). As such, fathering behavior is likely to be first determined by the social context at hand, then the father's empathy-related traits, and finally broader patterns of family functioning: the father's relationship with the infant's mother, and his overall quantity of involvement with the infant.

In this paper, we rely on two theoretical traditions, one from developmental psychology (Belsky, 1984), the other from evolutionary biology (Wingfield et al., 1990), to bridge ontogenetic and phylogenetic approaches to fathering research. We contend that fathering behavior should be multiply determined by psychological and socio-contextual factors, as Belsky (1984) postulates, but that testosterone should also factor into fathering behavior, as Wingfield et al. (1990), contend.

In this study we explored two questions. The first question focuses on situational context within father–infant interactions. We explored whether fathers' testosterone changes in response to observing his own infant in distress compared to interacting in a play-based interaction. Based on previous experimental work on fathers' responses to playing with their children (Gettler et al., 2011; Gray et al., 2007; Storey et al., 2011) compared to men's responses to hearing audio-taped infant cries (Fleming et al., 2002; Storey et al., 2000; van Anders et al., 2012), we hypothesized that fathers would experience significant changes in T in response to seeing their infants in distress. Our second question explored whether there were unique contributions of fathers' testosterone, the level of infant distress within the situational context, fathers' empathetic traits, marital relations, and quantity of childcare involvement on fathers' parenting behaviors. Because fathering is multiply determined by biological, psychological, and socio-contextual factors, we hypothesized that all of these factors would uniquely contribute to fathers' parenting behaviors.

## MATERIALS AND METHODS

### Recruitment

Participants were part of a longitudinal study investigating changes in family functioning after the birth of a second child across five time points: prenatal (during mother's third trimester of pregnancy), and 1, 4, 8, and 12 months following

the infant's birth. Women pregnant with their second child were recruited using advertisements and flyers posted in child care centers, local hospitals, pediatricians' offices, child-birth education classes, and obstetric clinics. Once families agreed to participate ( $N=241$  out of 408 eligible), the first prenatal home visit was scheduled and the study was explained in greater detail with an opportunity for the parents to ask questions prior to consenting. Families were compensated \$300 for completing all five time points of the study. Parents were asked to participate in the hormone substudy at the 12-month home visit. If parents consented to the hormone substudy, they provided three saliva samples during the 12-month laboratory visit. Each individual parent (mother or father) had the option to opt in or out of the substudy. A total of 175 fathers participated in the substudy.

### Participants

The 175 participating father–infant dyads did not differ significantly from the recruited sample on father's age, years of marriage, infant's gender, or father's race/ethnicity, but had significantly higher household incomes ( $\chi^2(3)=20.96$ ,  $p < .001$ ) and levels of father education ( $\chi^2(3)=9.33$ ,  $p < .05$ ). See Table 1 for a summary of participant characteristics.

### Study Summary

Data for the current report were drawn from a larger longitudinal study examining changes in the family after the birth of a second child using multiple methods (e.g., interviews, questionnaires, home observations). The hormone substudy was conducted at the 12-month time point and was designed to investigate how different hormones (e.g., T) were related to parenting and the quality of the infant's attachments to mother and father. Information obtained for the hormone substudy was used to address the questions of the current investigation.

At 12 months, two laboratory visits were conducted to assess the security of parent–infant attachment, parent–infant interaction, and variation in hormonal reactivity. The 12-month laboratory visit was completed at either 12 or 13 months of age and counterbalanced across mothers and fathers. In the current study, we used the Strange Situation Procedure (SSP) (Ainsworth, Blehar, Waters, & Wall, 1978) as a means of assessing whether men's T levels changed in response to their infant's distress. The SSP was designed to assess individual differences in infant–parent attachment relationships through a series of separations and reunions with parents. Separation from parents is a developmentally challenging stressor for 1-year-old infants which reliably increases infant stress (Gunnar, 2005) and other attachment behaviors (e.g., searching for the parent, clinging upon reunion). Parents often attempt to soothe and comfort their children upon reunion. We anticipated that fathers' observations, as well as their active involvement in managing their own infant's distress during the SSP, would elicit T changes comparable to those documented in earlier studies examining men's T reactivity to infant distress cues (Fleming et al., 2002; van Anders et al., 2012). The 15 min Teaching Task interaction session that followed involved a

**Table 1. Sample Characteristics**

	Mean	SD
Mother's age (in years)	31.94	3.86
Father's age (in years)	33.38	4.64
Years of marriage	5.80	2.62
	<i>N</i>	%
Infant's gender		
Girl	78	44.6
Boy	97	55.4
Mother's education		
High school degree or some college	21	12.0
Bachelor's degree	68	38.9
Professional degree	86	49.1
Father's education		
High school degree or some college	29	16.6
Bachelor's degree	66	37.7
Professional degree	80	45.7
Mother's race/ethnicity		
European American	154	88
African American	8	4.6
Asian/Asian American	4	2.3
Hispanic	6	3.4
Other	3	1.7
Father's race/ethnicity		
European American	154	88
African American	8	4.6
Asian/Asian American	5	2.9
Hispanic	5	2.9
Other	3	1.7
Family income		
20,000–59,999	41	23.4
60,000–99,999	67	38.3
>100,000	67	38.3

series of teaching tasks (Vondra, Shaw, & Kevenides, 1995) during which parent–infant interaction quality was assessed. Men were asked to teach their 1-year-old infants how to use a series of toys (e.g., hit keys in order on a xylophone) that were beyond their child's developmental age to accomplish alone during the Teaching Task.

The current report uses infants' distress levels during the father–infant SSP, parenting behaviors from the teaching task, saliva collected during the laboratory visit to assay T, fathers' reports of marital relationship quality at 12 months, division of infant care at 12 months based on couple report, and fathers' prenatal report of dispositional empathy and personal distress.

### Study Protocol

Prior to the laboratory visit, fathers were given a consent form to participate in the hormone substudy at the 12-month home visit. Upon arriving in the laboratory between 7:57 and 19:27 hr, fathers were informed by a trained researcher about the procedures for the Strange Situation and the father–infant

teaching task, as well as how to provide saliva samples for hormone analysis. Although we attempted to minimize time variation by scheduling visits between 13:00 and 18:00 hr (34% of visits), we prioritized the families' scheduling availability due to the large number of visits (~400 across mothers and fathers) that needed to be coordinated. To take diurnal variation into consideration, the time of the visit was used as a control in analyses. The procedure for the lab visit was as follows: First, fathers were instructed to provide a baseline saliva sample while their infants were with them in the waiting room (e.g., being held, playing on the floor with toys). Second, father–infant dyads participated in a video-taped SSP and following the SSP, fathers provided a second saliva sample (approximately 20 min after the first sample). Third, father–infant dyads participated in a 15 min video-taped teaching task (Volling, McElwain, Notaro, & Herrera, 2002), after which fathers provided the third and final saliva sample.

### Infant Distress and Strange Situation Procedure

The SSP (Ainsworth et al., 1978) assesses the quality of infant–parent attachment relationships through a series of separations and reunions between parents and infants which become increasingly stressful over the course of seven 3 min episodes (following a 1 min introduction to the room). Infants often become visibly upset during the procedure, searching for their parents during separations, and seeking comfort and contact upon reunions. Each separation episode was rated by two trained coders for infant distress using a 5-point rating scale, with 1 = *no distress*, to 5 = *immediate full distress*. Reliability between coders was determined using intraclass correlation coefficients (ICC) for each separation episode (Bartko, 1976). Reliability cases were randomly chosen and comprised of 20% of the sample (ICC average across three episodes = .94). The mean of the distress ratings was calculated to create an average distress score across the two separation episodes ( $M = 3.04$ ,  $SD = 1.28$ ,  $Range = 1-5$ ,  $\alpha = .83$ ). While separated from their infants, parents were allowed to observe their infants through a one-way observation window (and allowed to curtail the separation at their request) before returning to comfort their children during the reunion. Experimenters also curtailed separation episodes if they judged the infant's distress had reached extreme levels, as is standard practice when conducting the SSP.

### Teaching Task and Coding

During the teaching task, the fathers and infants were presented with three different toys, each in separate boxes with an instruction card specific to that toy (Vondra et al., 1995). Fathers were asked to teach the infant to hit each key on a xylophone with a mallet, push all of the levers on an activity box, and hit the shapes on a toy turtle's back. Fathers were told that all of the tasks were beyond the ability of a 12-month-old infant to complete alone and were asked to help their children do the best they could. Fathers were given 5 min for each toy. The teaching task session is challenging for both parents and infants, and requires more active parent participation than a standard free-play situation (Volling et al., 2002). Five trained

coders rated each 5 min episode separately for the father's behaviors using a coding system adapted from the NICHD Study of Early Child Care Research Network (2000), with 1 = *Not at all Characteristic* to 7 = *Very Characteristic*. Global codes were used to assess each 5 min period in its entirety for each of the behavioral codes described below, and then averaged across the entire 15 min paradigm (3 periods). Reliability between coders was determined using intraclass correlation coefficients for each 5 min epoch (Bartko, 1976). Reliability cases were randomly chosen and comprised 19% of the entire sample.

*Sensitivity/Responsiveness* (ICC = .86) focused on how the father responded to and observed the child's expressions, gestures, and signals, including whether the interaction was child-centered. *Intrusiveness* (ICC = .88) measured how overcontrolling and intrusive fathers were toward their child (e.g., not allowing the child to influence the pace or focus of the interaction and failing to modulate behavior that elicited distress from the child). *Detachment* (ICC = .88) measured how emotionally uninvolved or disengaged the father was toward the child (e.g., failing to respond to the child's approaches, vocalizations, smiles, or other social bids). *Positive regard* (ICC = .85) measured the father's positive feelings toward the child (e.g., smiling, warm tone of voice). *Negative regard* (ICC = .85) rated the father's negative feelings (e.g., disapproval, sarcasm) toward the child. *Stimulation of cognitive development* (ICC = .85) measured the degree to which the father tried to foster the child's cognitive development (e.g., attempting to focus child on task).

Means of each parenting code were calculated across the three 5 min teaching task sessions based on previous research to form a parenting composite (Vandell, 1996). The parenting composite was calculated as: Sensitivity + Positive Regard + Stimulation of Development – Intrusiveness – Negative Regard – Detachment. A higher score indicated more sensitively engaged fathering whereas a lower score indicated more intrusive and detached fathering. The first principal component of the six subscales confirmed the signs assigned to the subscales of the composite.

### Father's Empathy and Personal Distress

Fathers completed the Empathic concern (7 items,  $\alpha = .77$ , e.g., "I often have tender, concerned feelings of people less fortunate than me") and Personal distress (7 items,  $\alpha = .71$ , e.g., "In emergency situations, I feel apprehensive and ill-at-ease") subscales of the Interpersonal reactivity index (Davis, 1980). Each item was rated from 1 = *does not describe me well* to 5 = *describes me very well*. Scores were calculated by taking the mean across items.

### Marital Quality

At 12 months, fathers completed the 25-item Intimate Relations Questionnaire (Braiker & Kelley, 1979), which measured four dimensions of the marital relationship: *love* ( $\alpha = .86$ , e.g., "To what extent do you love your partner at this stage?"), *ambivalence* ( $\alpha = .76$ , e.g., "To what extent do you feel 'trapped' or pressured to continue in this relationship?"), *maintenance* ( $\alpha = .69$ , e.g., "To what extent do you

try to change your behavior to help solve certain problems between you and your spouse?"), and *conflict* ( $\alpha = .72$ , e.g., "When you and your spouse argue, how serious are the problems or arguments?"). Each item was rated on a 9-point scale ranging from 1 = *not at all/never* to 9 = *very/extremely*. Subscale scores were calculated by averaging across items.

### Division of Infant Care

Division of infant care was assessed during the 12-month home visit during a joint couple interview. Husbands and wives had to jointly agree on their involvement in nine tasks over the past month (e.g., "changing poopy diapers",  $\alpha = .84$ ) and each task was rated from 1 = *Almost Always Mother* to 5 = *Almost Always Father*. The mean across items was calculated, with a score of 3 indicating shared involvement.

### Saliva Collection and Assays

Saliva samples were collected in 50 ml polypropylene tubes (United Lab Plastics) three times during the laboratory visit. Participants provided 10 ml of saliva per sample. Saliva collection was stimulated by chewing sugar-free Trident Original gum, which was found to leave testosterone assay results unaffected (Dabbs Jr, 1991). The first sample was collected as a baseline after the fathers arrived in the laboratory (T1). The second sample (T2) was taken after the Strange Situation, which was approximately 15 min after the first separation of the SSP. The third sample (T3) was taken after the teaching task, approximately 20 min after the second sample. All samples were frozen at  $-20^{\circ}\text{C}$  until assayed. Samples were analyzed by radioimmunoassay using a commercial kit from Siemens Healthcare. The assay was modified for use with saliva according to published protocol (Campbell, Schultheiss, & McClelland, 1999). Water-based dilutions of all standards and controls were prepared to determine salivary testosterone concentrations. Samples were assayed in duplicate and the mean levels for each sample were utilized for analysis. Controls were used to assess assay reliability. The intra-assay CV was 10.17%, and the inter-assay CV was 21.22%. We created a composite of T reactivity, defined as the percent change  $((T2 - T1) / T1) \times 100$ ; a commonly used method of assessing short-term T reactivity (Carré, Iselin, Welker, Hariri, & Dodge, 2014; Fleming et al., 2002; Gray et al., 2007; Storey et al., 2011; van Anders et al., 2012; Weisman, Zagoory-Sharon, & Feldman, 2014). This composite of T reactivity is not subject to the usual interpretational difficulties when difference scores are used as predictors in a regression because T reactivity is normalized by T1.<sup>1</sup>

### Statistical Analyses

T levels at T1 had no outliers ( $\pm 3$  SD), T2 had one outlier, and T3 had four outliers. These values were excluded from further analyses. T percent change values were considered to be outliers if they were  $\pm 3$  SD from the mean. T change from T1 to T2 had two outliers, from T1 to T3 had five outliers, and from T2 to T3 had four outliers, and these values were excluded from further analyses, which is standard practice in prior T research (Carré et al., 2014; van Anders et al., 2012). Eight values were missing for T1, four for T2, and nine for T3 due to error in sample collection or assay, and blood contamination in samples. Results from Little's Chi-Square Test of MCAR (Little, 1988) revealed that T level values were missing completely at random ( $\chi^2(7) = 8.378$ ,  $p = .30$ ). Thus missing data cases were excluded from analyses via pairwise deletion for the zero-order correlations, an acceptable procedure for data missing completely at random (Allison, 2001). To maximize power, missing data were dealt with by using full information restricted maximum likelihood (FIRML) in the linear mixed model and regression analyses.

To address our first aim, mean testosterone levels were examined using linear mixed models to assess change in fathers' T over the three times of saliva collection. Where relevant, degrees of freedom for the linear mixed models were computed using the Kenward-Roger correction. To address our second aim, which focused on examining whether variation in T reactivity, along with other psychosocial determinants, predicted fathering behavior, we conducted a series of five hierarchical regression models.

## RESULTS

Our preliminary analyses included testing for potential covariates of T, T reactivity, and fathering behaviors. Potential known confounds of T levels (last time brushed teeth, father's age, time of day, BMI, and seasonality) were evaluated as potential covariates in Pearson correlations. BMI, time of Day and last time brushed teeth were significant covariates of Time 1 (baseline) testosterone (T1) and Time 2 (after SSP) testosterone (T2). Time of Day and last time brushed teeth were significant covariates of Time 3 (after Teaching Task) testosterone (T3). The same T level

---

variances and covariances from the variables that created the composite score; however, this critique is true for all regression betas, which are always combinations of variances and covariances from all variables in the model. Some advocate a residualized change score by including both T1 and T2 as separate predictors. Our approach for dealing with these concerns is to use percent change instead of the raw difference score because percent change normalizes the difference by the individual's own baseline; we report statistical inference for models both with and without testosterone at T1 as a covariate.

<sup>1</sup>The use of difference scores in regression, both as predictor and as dependent variable, continues to be debated (Allison, 1990; Shadish, Cook, & Campbell, 2002). The key concern when a difference score is used as predictor is that the regression slope, beta, associated with a composite predictor is difficult to interpret because the beta will be a function of

confounds were evaluated as potential covariates for T reactivity. There were no significant covariates for T1 to T2 percent change. Covariates were only included in analyses when they were significantly associated with the relevant T variable; therefore, none of these covariates (last time brushed teeth, father's age, time of day, BMI, and seasonality) were included in analyses involving T1 to T2 percent change. We also tested whether demographic variables (father age, years married, father's education, father's race/ethnicity and infant's gender) covaried with fathers' parenting behaviors that may need to be controlled in the hierarchical regressions, using correlations (fathers age and years married) and one-way ANOVAs (fathers' education, race/ethnicity, household income, and infant gender). Infant gender was the only significant covariate and was included in the regression analyses predicting fathers' parenting behavior.

We first examined the raw T scores at the three time points to assess the fathers' T trajectories across the laboratory sessions. We conducted a linear mixed model, using significant covariates identified in our preliminary analyses (BMI, time of day, time since last brushed teeth). There was a significant omnibus time effect,  $F(2, 293.1) = 29.60$ ,  $p < .001$ , indicating that, on average, fathers' T declined significantly over time. Pairwise comparisons of mean differences between T levels showed that baseline T was significantly different from T2 after SSP,  $t(143.93) = -5.69$ ,  $p < .001$ , and T3 after the teaching interaction,  $t(143.93) = 7.34$ ,  $p < .001$ , but T2 and T3 were not significantly different from each other ( $p = .099$ ). In sum, fathers, on average, displayed a significant decline

in T while observing and interacting with their infants as they participated in the stressful SSP, but showed no additional significant change in T during the time they interacted with their infants in the Teaching Task.

Having established the basic trajectory of T patterns across the three time points, we conducted Pearson correlations to examine the associations between T levels at baseline (T1), after the SSP (T2) and after the Teaching Task (T3), and percent change scores across the three segments with the psychosocial determinants of fathering in line with Belsky's (1984) determinants of parenting model where father characteristics (e.g., dispositional empathy and personal distress), infant characteristics (e.g., infant's distress during the SSP, infant gender), and social-contextual characteristics (e.g., the division of childcare, marital relationship quality) predict fathers' parenting behaviors during the Teaching Task (see Table 2). T1 was positively associated with fathers' dispositional empathic concern. Only the correlation between the T percent change score from baseline (T1) to after SSP (T2) was significantly, negatively correlated with fathers' parenting behaviors during the Teaching Task ( $p < .05$ ) indicating that greater declines in T during the SSP were associated with fathers' more sensitive, cognitively stimulating and positive interactions with their infants. All other T levels and percent change scores, including the percent change across T2 and T3 while fathers interacted with their infants in the Teaching Task, were not significantly correlated with fathers' parenting behaviors ( $p$ 's  $> .239$ ). Individual differences in fathers' T were highly stable across the three time samples as indicated by the significant

**Table 2. Correlation Matrix and Descriptives of Study Variables, Excluding all T and T Percent Change Outliers (N = 142–149)**

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Parenting behavior	—											
2. T baseline (T1)	.02	—										
3. T SSP (T2)	-.09	.77**	—									
4. T TT (T3)	-.06	.78**	.80**	—								
5. Empathic concern	.02	.20*	.11	.11	—							
6. Personal distress	.22**	.06	.08	.05	.16*	—						
7. Division of infant care	.04	-.08	-.02	-.05	-.04	.01	—					
8. Marital love	.18*	-.03	-.02	.01	.01	-.17*	-.21*	—				
9. Infant distress	.17*	.01	.06	-.06	-.03	.08	-.01	.12	—			
10. T1–T2 reactivity	-.17*	-.28**	.37**	.10	-.11	.03	.05	.02	.09	—		
11. T1–T3 reactivity	-.10	-.24**	.11	.40**	-.10	-.02	.01	.10	-.08	.58**	—	
12. T2–T3 reactivity	.09	.06	-.22**	.38**	.02	-.06	-.03	.13	-.19*	-.37**	.52**	—
Mean	5.85	66.05	59.05	57.28	3.68	2.04	2.25	7.33	3.04	-8.98	-11.84	-1.64
Standard deviation	3.61	20.10	18.99	18.74	.64	.55	.60	.96	1.28	18.96	19.92	19.66

\* $p < .05$ .

\*\* $p < .01$ .

**Table 3. Hierarchical Regression Predicting Fathering Behavior During Teaching Task ( $N = 162$ ), Removing All T Level Outliers and T1–T2 Percent Change Outliers**

Variable	Step 1			Step 2			Step 3			Step 4			Step 5		
	B	SE	$\beta$	B	SE	$\beta$	B	SE	$\beta$	B	SE	$\beta$	B	SE	$\beta$
Time of day	-.00	.00	-.04	-.00	.00	-.02	.00	.00	.00	.00	.00	.01	-.00	.00	-.03
Infant gender <sup>a</sup>	-2.04**	.53	-.29	-1.87**	.54	-.26	-1.86**	.53	-.26	-1.71**	.53	-.24	-1.60**	.52	-.23
T1–T2 reactivity				-.03*	.01	-.15	-.03*	.01	-.17	-.04*	.01	-.19	-.04*	.01	-.19
Infant distress							.46*	.21	.16	.41	.21	.15	.30	.21	.11
Empathic concern										-.23	.41	-.04	-.27	.39	-.05
Personal distress										.89	.46	.14	1.19*	.46	.19
Marital love													.81**	.28	.23
Division of childcare													.36	.43	.06
R <sup>2</sup>	.09			.11			.13			.16			.20		
Wald test	14.89***			3.97*			4.73*			3.81			8.46*		

Note: Similar pattern of significant results for T1–T2 reactivity emerged when including baseline T1 in all models.

<sup>a</sup>(infant gender 0 = girl, 1 = boy).

\* $p < .05$ .

\*\* $p < .01$ .

\*\*\* $p < .001$ .

positive correlations in men's T across T1, T2, and T3 (see Table 2).

Fathers' parenting behaviors were also positively associated with fathers' dispositional personal distress ( $p < .01$ ), their reports of marital love ( $p < .05$ ), and infant distress during SSP ( $p < .05$ ). Marital love was the only marital quality dimension significantly associated with fathers' parenting behavior and was used in subsequent analyses as an indicator of marital relationship quality (marital maintenance, conflict, ambivalence:  $p$ 's  $> .250$ ).

We used hierarchical multiple regression with full information restricted maximum likelihood to test whether T reactivity to the SSP, infant distress, fathers' empathy, marital relationship quality, and division of infant care each uniquely predicted fathers' parenting behavior in the Teaching Task. We included the variables in steps guided by Belsky's (1984) Multiple Determinants of Parenting model, in which variables were added from proximal to distal influence. We first included variables that were immediate to the situation (father's T reactivity and infant distress in SSP), then fathers' individual characteristics (empathic concern, personal distress), and finally, family level constructs (marital relationship quality, division of infant care). We were particularly interested in fathers' declines in T during the SSP because the procedure requires that fathers observe their infant's distress, as well as comfort and soothe their infants, and because it was the only T change score significantly correlated with fathering behavior. We hypothesized situational T reactivity in this emotionally eliciting context would be predictive of sensitive parenting behaviors during

the Teaching Task. We added time of Day and infant gender in the first step as controls. T reactivity to the SSP was then added in the second step to test whether T reactivity predicted fathering behavior, and this model fit significantly better than the first model that only included time of Day and infant gender ( $p < .05$ ). In the third step, infant distress was added to assess whether T reactivity was indirectly affecting fathering via levels of infant distress, and this addition significantly improved the model fit ( $p < .05$ ). In the fourth step, fathers' dispositional tendencies (empathic concern and personal distress) were added to assess whether fathers' empathy-related characteristics may account for the role of T reactivity and infant distress in fathering behavior, but this did not significantly improve model fit ( $p = .149$ ). In the fifth step, marital love and division of infant care were added as contextual predictors of fathering and significantly improved model fit ( $p < .05$ ). Results presented in Table 3 revealed that T reactivity from T1 to T2 uniquely predicted ( $p < .05$ ) fathers' parenting behavior across models at Step 2 through Step 5, even when controlling for time of day. The pattern of results and statistical significance for T reactivity remained the same when baseline T1 was added as a covariate in each of the models. Consistent with Belsky's model, multiple psychological and socio-contextual factors also uniquely predicted fathers' parenting behavior in the final model: infant gender, with fathers being more positively engaged with daughters compared to sons, fathers' trait personal distress and marital love, which both positively predicted fathering behavior. When we tested alternate models that examined whether levels of

T (T1, T2, T3) or concurrent changes in T (T2–T3) during the Teaching Task were predictive of parenting behaviors, all T predictors in these models were nonsignificant (all  $p$ 's > .12).

## CONCLUSIONS

We explored whether changes in fathers' T in response to their own infants in distress predicted sensitive and positively engaging fathering behavior during a teaching task. We found, on average, that fathers' T declined in response to the SSP, quite possibly in response to witnessing their infants in distress, and not during their time interacting with the infant in the Teaching Task. We also explored whether fathers' parenting behavior could be predicted by biological, psychological, and socio-contextual factors. We found that infant characteristics, fathers' T reactivity to infant distress, fathers' empathy related traits and relationship quality with the infant's mother each uniquely predicted fathers' parenting behavior. The remainder of this discussion will evaluate and contextualize our findings.

### Average Patterns of Testosterone Reactivity

On average, fathers' T declined significantly during the SSP and not the Teaching Task. We hypothesized that these T declines were elicited because fathers were not only observing their infants in distress during the separations of the SSP, but also because of their active participation in comforting and soothing their infants upon reunion. We cannot, however, rule out the possibility that men's T would have declined initially during the first 20 min of the lab visit regardless of whether it was the SSP or the Teaching Task that occurred first because we did not counterbalance the two sessions. Future studies that wish to elucidate how fathers' testosterone changes in response to infant distress compared to playful interaction should counterbalance the SSP and the Teaching Task. Although the lack of counterbalancing within our study is a methodological issue, our results are consistent with previous work that has found that infant cries elicit significant changes in T (van Anders, Tolman, & Jainagaraj, 2014; van Anders et al., 2012) compared to playful interactions, which often do not (Gettler et al., 2011; Storey et al., 2011).

### Individual Differences in Testosterone Reactivity

T reactivity during the SSP predicted subsequent sensitive, responsive fathering during the father-infant teaching task, whereas concurrent changes in T during the

teaching task were not associated with fathers' behaviors, nor were absolute levels of T. Our results suggest that individual variation in fathers' T reactivity during a stressful laboratory paradigm in which fathers interacted and comforted their own distressed infants predicted individual differences in sensitive fathering. These findings also support prior research that men's T declined in response to infant distress when men were allowed to nurture and respond to infant distress (van Anders et al., 2012). We hypothesize, based on the current findings and previous work on men's empathy, that fathers' T reactivity to infant distress is potentially modulated by cognitive appraisals of their infant's distress, and that fathers' dispositional empathic characteristics may shape how they respond to their distressed infant. For example, if fathers interpret infant crying as a means of communicating the infant's internal emotional state and empathize with their infant's distress, they may experience a decline in T, which may, in turn, facilitate a nurturant response. Alternatively, when fathers interpret their infant's crying as aggravating, they may experience increases in T which, in turn, facilitates an intrusive or negative response. We acknowledge that a limitation of this study is that we did not measure fathers' actual feelings of empathy and personal distress before, during or after the SSP, and instead, used a measure of dispositional empathy. Future research examining fathers' responses to infant distress may need to examine situational empathy directly.

Although fathers' parenting behaviors during the Teaching Task were not concurrently related to T change during the Teaching Task (i.e., T2–T3 percent change), infant distress during the SSP was related to further declines in fathers' T during the Teaching Task (see correlations in Table 3), providing additional evidence for the link between men's T reactivity in response to infant distress. We cannot determine, however, whether there is a mechanism that drives both fathers' sensitivity and their T decreases. Future research is needed to replicate our findings. T production is regulated by the hypothalamic-pituitary-gonadal axis (Swerdlow, Wang, & Bhasin, 1992), but the underlying mechanisms by which T is implicated in parental responsiveness in humans is largely unknown (Bos, Panksepp, Bluthé, & Honk, 2012). There may be other hormones implicated, such as prolactin, which facilitates paternal behavior and is inversely related to testosterone, (Reburn & Wynne-Edwards, 1999), progesterone, which increases affiliation and acts as a suppressor of T (Brown et al., 2009), estradiol, which is converted from testosterone in the brain (Trainor & Marler, 2002), or cortisol, which also suppresses T (Sapolsky, 1985). Therefore, decreases in testosterone may be associated with increases in cortisol, prolactin, progesterone, and estradiol, and these in turn,

may facilitate sensitive, nurturing behavior in fathers. Given that infant cries elicit activation in men's brain regions associated with motivation, emotion regulation and social cognition (Swain et al., 2014), the activation of these areas during fathers' exposure to their infant's cries may elicit changes in hormone levels, which, in turn, facilitate sensitive fathering behaviors. Future studies would benefit by assessing these additional hormones and consider them simultaneously in the prediction of fathering behavior to determine if changes in T still remain significant in predicting human paternal behavior.

Given the situational specificity of our current results, we must also acknowledge that other contexts (such as infant defense) may result in different T responses for men that may or may not be related to sensitive parenting (van Anders et al., 2011). It may be more useful to consider men's androgen levels as part of a flexible regulatory system that responds to environmental demands. For instance, situations in which the infant may be at-risk for harm or injury inflicted by others might elicit increases in men's T and require that fathers react quickly and aggressively to protect their infant (van Anders et al., 2011). Thus, we are not arguing that universal declines in T will always be associated with "good fathering" but that the contextual demands of the caregiving context (i.e., protection from harm vs. nurturance) as well as the anticipated functional outcome for the infant need to be considered when interpreting relations between men's T responses and paternal behavior.

### **The Role of Psychosocial Characteristics in Predicting Fathering Behaviors**

Our results also indicated that sensitive responding during father–infant interaction was multiply determined by infant, father, and social contextual characteristics (Belsky, 1984). In addition to T reactivity, our final regression model revealed that the infant's gender, fathers' reports of marital love, and fathers' dispositional personal distress were unique predictors of fathering behavior. Specifically, fathers of daughters were more sensitive than fathers of boys, whereas men's personal distress reactions and marital love positively predicted sensitive fathering behaviors. Observer-rated levels of infant distress, empathic concern, and quantity of father involvement in physical care did not uniquely predict fathering behaviors during the Teaching Task.

### **Infant Characteristics**

Fathers of daughters were more sensitively engaged than fathers of sons, consistent with previous research comparing fathers' sensitivity during interactions with

toddler girls and boys (Lovas, 2005). Fathers may behave differently with girls and boys based on their own beliefs about gender (Bem, 1983). For example, fathers may believe girls are more delicate and, in turn, behave more sensitively with daughters than sons. Although we did not examine fathers' gender beliefs, sex-typed behavior toward infants tends to operate subconsciously (Culp, Cook, & Housley, 1983), so fathers may be unaware that they are less sensitive toward boys.

Observer-rated levels of infant distress during the SSP was a unique predictor of fathers' parenting behavior when initially entered into our hierarchical regression models, but were no longer significant when fathers' empathic traits were entered into the model at step 4, indicating that fathers' empathy-related psychological state could be explaining the relationship between situational context (infant distress) and fathers' parenting behavior.

### **Fathers' Personality Traits**

Fathers' empathic concern did not predict parenting behavior, but personal distress did. Our findings are likely guided by the operational definition of these variables. Our measure of empathic concern assesses the tendency to feel concern for those who are less fortunate, whereas the personal distress scale measures the tendency to feel discomfort and distress when observing others in distress (Davis, 1980). Concern for others who are less fortunate may not apply to fathers' interactions with their infants, because it is unlikely that fathers believe their children are less fortunate than them. In contrast, fathers who have greater trait personal distress may be more reactive, both physiologically and psychologically, when they observe their own infants in distress (Ho, Konrath, Brown, & Swain, 2014). These reactions may, in turn, mobilize the father to soothe his infant to reduce both the infant's distress and his own.

### **Contextual Family Characteristics**

Fathers' quantity of involvement in the physical care of their infant did not uniquely predict their quality of behavior during the Teaching Task. We offer two explanations for this finding. First, there are very few primary caregiving fathers in our sample. On average, couples reported that mothers usually did most of the physical care for the infant (e.g., feeding, changing diapers). Therefore, there may not have been enough variation in the quantity of fathers' involvement to predict the quality of their fathering behavior. Alternatively, quantity may not predict quality because they are two discrete constructs: there may be less variation

in quantity, but more variation in quality. Except in extreme cases of abuse or neglect, most parents attend to an infant's physical needs. But, not all parents behave in a sensitive manner with their infants.

Fathers' love for their spouses positively predicted their sensitive parenting behaviors. These findings are consistent with a vast literature on the spillover between marital relationships on father-child relationships (Belsky & Jaffee, 2006; Cummings, 2010; Parke, 1996). More love between spouses translates into greater warmth in the family and thus, more sensitive parenting behaviors. Although we did not measure fathers' typical patterns of parenting behavior, it is likely that fathers who are more sensitive with their infants at home were more sensitive with their infants in the laboratory during the Teaching Task.

## CONCLUSIONS

Fathering involves a complex combination of nurturance, protection, discipline, teaching, and mentoring that varies across the infancy, childhood, and adolescence of their offspring. Undoubtedly, the relations between fathering behaviors, hormonal variation, and neural networks will be equally complex. This will require that future research move beyond an examination of between-group differences (e.g., fathers vs. non-fathers) and begin to examine individual differences (i.e., within-group) and how fathers respond to different caregiving demands. In this paper, we bridged ontogenetic and phylogenetic approaches to understand individual differences in fathers' parenting behavior. Sensitive and responsive fathering has been linked to young children's social, emotional, and cognitive development and our study provides some of the first empirical evidence that declines in fathers' T, in addition to fathers' empathy-related traits and family context characteristics may benefit infant socioemotional development by enhancing nurturant fathering behaviors.

## NOTES

This research was supported by grants (R01HD042607, K02HD047423) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to Brenda L. Volling, and a faculty grant to Volling and Schultheiss from the Office of Research at the University of Michigan. We are grateful to the fathers of the Family Transitions Study and to the research assistants who provided invaluable help with data collection. Portions of this research were presented at the Biennial Meeting of the Society for Research in Child Development, Seattle, WA, April 2013.

## REFERENCES

- Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillsdale, N.J.: New York: Lawrence Erlbaum Associates; distributed by Halsted Press Division of Wiley.
- Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological methodology*, 20(1), 93–114.
- Allison, P. D. (2001). *Missing data*: Sage.
- Bartko, J. J. (1976). On various intraclass correlation reliability coefficients. *Psychological bulletin*, 83(5), 762–765.
- Belsky, J. (1984). The determinants of parenting: A process model. *Child Development*, 83–96.
- Belsky, J., & Jaffee, S. R. (2006). The multiple determinants of parenting. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol 3: Risk, disorder, and adaptation (2nd ed.)*. (pp. 38–85). Hoboken, NJ US: John Wiley & Sons Inc.
- Bem, S. L. (1983). Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. *Signs*, 8(4), 598–616. DOI: 10.2307/3173685
- Bos, P. A., Panksepp, J., Bluthé, R.-M., & Honk, J. V. (2012). Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: A review of single administration studies. *Frontiers in Neuroendocrinology*, 33(1), 17–35. DOI: <http://dx.doi.org/10.1016/j.yfme.2011.01.002>
- Braiker, H. B., & Kelley, H. H. (1979). Conflict in the development of close relationships. In R. L. Burgess & T. L. Huston (Eds.), *Social exchange in developing relationships* (pp. 135–168). New York: Academic.
- Brown, S. L., Fredrickson, B. L., Wirth, M. M., Poulin, M. J., Meier, E. A., Heaphy, E. D., ... Schultheiss, O. C. (2009). Social closeness increases salivary progesterone in humans. *Hormones and Behavior*, 56(1), 108–111. DOI: <http://dx.doi.org/10.1016/j.yhbeh.2009.03.022>
- Campbell, K. L., Schultheiss, O. C., & McClelland, D. C. (1999). A necessary adjustment of protocol for use of DPC Coat-a-Count total testosterone assay with saliva. *Clinical Biochemistry*, 32(1), 83–85. DOI: 10.1016/s0009-9120(98)00080-0
- Carré, J. M., Iselin, A. M., Welker, K. M., Hariri, A. R., & Dodge, K. A. (2014). Testosterone reactivity to provocation mediates the effect of early intervention on aggressive behavior. *Psychological Science*. DOI: 10.1177/0956797614525642
- Carré, J. M., McCormick, C. M., & Hariri, A. R. (2011). The social neuroendocrinology of human aggression. *Psychoneuroendocrinology*, 36(7), 935–944. DOI: 10.1016/j.psychneuen.2011.02.001
- Culp, R. E., Cook, A. S., & Housley, P. C. (1983). A comparison of observed and reported adult-infant interactions: Effects of perceived sex. *Sex Roles*, 9(4), 475–479. DOI: 10.1007/BF00289787
- Cummings, E. M. (2010). Fathers, marriages, and families: Revisiting and updating the framework for fathering in

- family context. In M. E. Lamb (Ed.), *The Role of the Father in Child Development* (pp. 154–176). Hoboken, NJ: John Wiley and Sons.
- Dabbs, J. M., Jr. (1991). Salivary testosterone measurements: Collecting, storing, and mailing saliva samples. *Physiology & Behavior*, 49(4), 815–817. DOI: [http://dx.doi.org/10.1016/0031-9384\(91\)90323-G](http://dx.doi.org/10.1016/0031-9384(91)90323-G)
- Davis, M. H. (1980). Measuring individual differences in empathy: Evidence of a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113–126.
- Fairbrother, N., Barr, R. G., Pauwels, J., Brant, R., & Green, J. (2015). Maternal thoughts of harm in response to infant crying: An experimental analysis. *Archives of Women's Mental Health*, 18(3), 447–455. DOI: [10.1007/s00737-014-0471-2](https://doi.org/10.1007/s00737-014-0471-2)
- Fleming, A. S., Corter, C., Stallings, J., & Steiner, M. (2002). Testosterone and prolactin are associated with emotional responses to infant cries in new fathers. *Hormones and Behavior*, 42(4), 399–413. DOI: [10.1006/hbeh.2002.1840](https://doi.org/10.1006/hbeh.2002.1840)
- Gettler, L. T., McDade, T. W., Agustin, S. S., & Kuzawa, C. W. (2011). Short-term changes in fathers' hormones during father-child play: Impacts of paternal attitudes and experience. *Hormones and Behavior*, 60(5), 599–606. DOI: <http://dx.doi.org/10.1016/j.yhbeh.2011.08.009>
- Gettler, L. T., McDade, T. W., Feranil, A. B., & Kuzawa, C. W. (2011b). Longitudinal evidence that fatherhood decreases testosterone in human males. *Proceedings of the National Academy of Sciences*. DOI: [10.1073/pnas.1105403108](https://doi.org/10.1073/pnas.1105403108)
- Gray, P. B., Parkin, J. C., & Samms-Vaughan, M. E. (2007). Hormonal correlates of human paternal interactions: A hospital-based investigation in urban Jamaica. *Hormones and Behavior*, 52(4), 499–507. DOI: <http://dx.doi.org/10.1016/j.yhbeh.2007.07.005>
- Gunnar, M. R. (2005). Attachment and stress in early development: Does attachment add to the potency of social regulators of infant stress? In C. S. Carter, L. Ahnert, K. E. Grossmann, S. B. Hrdy, M. E. Lamb, S. W. Porges, & N. Sachser, (Eds.), *Attachment and bonding: A new synthesis*. (pp. 245–255). Cambridge, MA, US: MIT Press.
- Hermans, E. J., Putman, P., & van Honk, J. (2006). Testosterone administration reduces empathetic behavior: A facial mimicry study. *Psychoneuroendocrinology*, 31(7), 859–866. DOI: [10.1016/j.psychoneu.2006.04.002](https://doi.org/10.1016/j.psychoneu.2006.04.002)
- Ho, S. S., Konrath, S., Brown, S., & Swain, J. E. (2014). Empathy and stress related neural responses in maternal decision making. *Frontiers in Neuroscience*, 8, 152. DOI: [10.3389/fnins.2014.00152](https://doi.org/10.3389/fnins.2014.00152)
- Kuo, P. X., Carp, J., Light, K. C., & Grewen, K. M. (2012). Neural responses to infants linked with behavioral interactions and testosterone in fathers. *Biological Psychology*, 91(2), 302–306. DOI: <http://dx.doi.org/10.1016/j.biopsycho.2012.08.002>
- Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202.
- Lovas, G. S. (2005). Gender and patterns of emotional availability in mother-toddler and father-toddler dyads. *Infant Mental Health Journal*, 26(4), 327–353. DOI: [10.1002/imhj.20056](https://doi.org/10.1002/imhj.20056)
- Network, N. E. C. R. (2000). Factors associated with fathers' caregiving activities and sensitivity with young children. *Journal of Family Psychology*, 14(2), 200–219. DOI: [10.1037/0893-3200.14.2.200](https://doi.org/10.1037/0893-3200.14.2.200)
- Parke, R. D. (1996). *Fatherhood*. Cambridge, Mass: Harvard University Press.
- Reburn, C. J., & Wynne-Edwards, K. E. (1999). Hormonal changes in males of a naturally biparental and a uniparental mammal. *Hormones and Behavior*, 35(2), 163–176. DOI: <http://dx.doi.org/10.1006/hbeh.1998.1509>
- Rilling, J. K. (2013). The neural and hormonal bases of human parental care. *Neuropsychologia*, 51(4), 731–747.
- Sapolsky, R. M. (1985). Stress-induced suppression of testicular function in the wild baboon: Role of glucocorticoids. *Endocrinology*, 116(6), 2273–2278. DOI: [10.1210/endo-116-6-2273](https://doi.org/10.1210/endo-116-6-2273)
- Seifritz, E., Esposito, F., Neuhoﬀ, J. G., Lüthi, A., Mustovic, H., Dammann, G., . . . Di Salle, F. (2003). Differential sex-independent amygdala response to infant crying and laughing in parents versus nonparents. *Biological Psychiatry*, 54(12), 1367–1375. DOI: [10.1016/S0006-3223\(03\)00697-8](https://doi.org/10.1016/S0006-3223(03)00697-8)
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.
- Storey, A. E., Noseworthy, D. E., Delahunty, K. M., Halfyard, S. J., & McKay, D. W. (2011). The effects of social context on the hormonal and behavioral responsiveness of human fathers. *Hormones and Behavior*, 60(4), 353–361. DOI: <http://dx.doi.org/10.1016/j.yhbeh.2011.07.001>
- Storey, A. E., Walsh, C. J., Quinton, R. L., & Wynne-Edwards, K. E. (2000). Hormonal correlates of paternal responsiveness in new and expectant fathers. *Evolution and Human Behavior*, 21(2), 79–95. DOI: [10.1016/S1090-5138\(99\)00042-2](https://doi.org/10.1016/S1090-5138(99)00042-2)
- Swain, J. E., Dayton, C. J., Kim, P., Tolman, R. M., & Volling, B. L. (2014). Progress on the paternal brain: Theory, animal models, human brain research, and mental health implications. *Infant Mental Health Journal*, 35(5), 394–408. DOI: [10.1002/imhj.21471](https://doi.org/10.1002/imhj.21471)
- Swain, J. E., Lorberbaum, J. P., Kose, S., & Strathearn, L. (2007). Brain basis of early parent-infant interactions: psychology, physiology, and in vivo functional neuroimaging studies. *Journal of Child Psychology and Psychiatry*, 48(3–4), 262–287. DOI: [10.1111/j.1469-7610.2007.01731.x](https://doi.org/10.1111/j.1469-7610.2007.01731.x)
- Swerdlow, R. S., Wang, C., & Bhasin, S. (1992). 10 Developments in the control of testicular function. *Baillière's Clinical Endocrinology and Metabolism*, 6(2), 451–483.
- Trainor, B. C., & Marler, C. A. (2002). Testosterone promotes paternal behaviour in a monogamous mammal via conversion to oestrogen. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1493), 823–829.

- 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(9), 1265–1275.
- van Anders, S. M., Tolman, R. M., & Jainagaraj, G. (2014). Examining how infant interactions influence men's hormones, affect, and aggression using the Michigan infant nurturance simulation paradigm. *Fathering*, 12(2), 143–160. DOI: 10.3149/fth.1202.143
- van Anders, S. M., Tolman, R. M., & Volling, B. L. (2012). Baby cries and nurturance affect testosterone in men. *Hormones and Behavior*, 61(1), 31–36.
- Vandell, D. L. (1996). Characteristics of infant child care: Factors contributing to positive caregiving: NICHD early child care research network. *Early Childhood Research Quarterly*, 11(3), 269–306. DOI: 10.1016/s0885-2006(96)90009-5
- Volling, B. L., McElwain, N. L., Notaro, P. C., & Herrera, C. (2002). Parents' emotional availability and infant emotional competence: Predictors of parent-infant attachment and emerging self-regulation. *Journal of Family Psychology*, 16(4), 447–465. DOI: 10.1037/0893-3200.16.4.447
- Vondra, J. I., Shaw, D. S., & Kevenides, M. C. (1995). Predicting infant attachment classification from multiple, contemporaneous measures of maternal care. *Infant Behavior and Development*, 18(4), 415–425. DOI: 10.1016/0163-6383(95)90031-4
- Weisman, O., Zagoory-Sharon, O., & Feldman, R. (2014). Oxytocin administration, salivary testosterone, and father—infant social behavior. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 49, 47–52. DOI: 10.1016/j.pnpbp.2013.11.006
- Wingfield, J. C., Hegner, R. E., Dufty, A. M. J., & Ball, G. F. (1990). The “challenge hypothesis”: Theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. *American Naturalist*, 136(6), 829–846.