



Parent-Reports of Sex-Typed Play Preference in Preschool Children: Relationships to 2D:4D Digit Ratio and Older Siblings' Sex

Lisa M. Körner¹ · Marie Luisa Schaper¹ · Bettina M. Pause¹ · Martin Heil¹

Received: 24 August 2018 / Revised: 11 February 2020 / Accepted: 13 February 2020 / Published online: 28 March 2020
© The Author(s) 2020

Abstract

Sex-typed play behavior shows large sex differences and seems to be affected by prenatal sex hormones. For example, a smaller, more male-typical ratio between the second and fourth digit length (2D:4D), a proposed marker for prenatal testosterone exposure, has been shown to be related to sex-typed play preference in childhood. Nevertheless, it is still being debated whether 2D:4D displays a stable sex difference throughout childhood, as there are few longitudinal studies. In the present study, children's 2D:4D was measured on both hands on four occasions from early infancy to early childhood (T1: 5 months, T2: 9 months, T3: 20 months, and T4: 40 months) providing the rare possibility to test the temporal stability of the sex difference. Parents completed the Preschool Activities Inventory at T4 and reported on the number of older brothers and sisters as a measure for socialization influences. Parents described boys as playing more masculine and less feminine than girls. Boys had smaller 2D:4D than girls at all measurements (T1–T4) and on both hands (right/left). Nevertheless, 2D:4D increased significantly from T3 to T4 in both sexes. Girls, but not boys, who were described as playing more masculine and less feminine had more masculine 2D:4D ratios at T1–T4 on both hands (except for right 2D:4D at T2 and T3) and had more older brothers and fewer older sisters. These data underline the stability of the sex difference in 2D:4D and show the importance of both biological and social influences on sex-typed play behavior.

Keywords Sex differences · 2D:4D · Prenatal testosterone · Sex-typed play preference · Preschool Activities Inventory

Introduction

Sex-typed toy preference is one of the earliest observed sex differences in behavior, becoming apparent in children as young as 12 months (Servin, Bohlin, & Berlin, 1999; Todd, Barry, & Thommessen, 2017; Todd et al., 2018; van de Beek, van Goozen, Buitelaar, & Cohen-Kettenis, 2009). Whereas such sex differences in toy preference are only modest in size around one year of age (for a review, see Zosuls & Ruble, 2018), they increase with age (Golombok et al., 2008; for a review, see Todd et al., 2018) with very large effect sizes of about Cohen's $d=3$ for preschool and primary school children (for a review and meta-analysis, see Davis & Hines, 2020; Hines, 2010). However, the size of these sex differences depends on the method used to determine sex-typed play preference. Observational

studies typically find sex differences for playing with cars/trucks and dolls (for reviews, see Davis & Hines, 2020; Zosuls & Ruble, 2018) that are smaller than for parents' reports on sex-typed play preferences in a questionnaire (for a review, see Hines, 2010). Noticeably, not only children, but also non-human primates seem to prefer to play with sex-specific toys (Alexander & Hines, 2002; Hassett, Siebert, & Wallen, 2008). This is notable because non-human primates are obviously less subject to social influences. Both the early appearance of sex-specific play preferences in children, as well as the evidence from non-human primates, suggest a biological component in the emergence of sex-typed play preferences in addition to socialization influences.

Sex hormones that are involved in the development of primary and secondary sexual characteristics in ontogenesis are a likely candidate for such a biological component. However, sex-typed play behavior appears long before puberty in a time frame in which there are no differences in sex hormone levels between boys and girls (Hines, 2010). Following from this, early prenatal effects of sex hormones have been suggested to impact sex-typed preferences in children (for a review, see Berenbaum & Beltz, 2016). For example, girls with congenital

✉ Lisa M. Körner
Lisa.Koerner@hhu.de

¹ Department of Experimental Psychology, Institute of Experimental Psychology, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

adrenal hyperplasia (CAH) who are exposed to high levels of androgens prenatally and in the early postnatal period show more male-typical play behavior than unaffected controls (Berenbaum & Hines, 1992; Nordenström, Servin, Bohlin, Larsson, & Wedell, 2002; Pasterski et al., 2005) and parents describe their daughters behavior as more masculine in comparison with unaffected female relatives (Hines, 2003). Additionally, women with CAH retrospectively describe their play behavior as more masculine (Hines, Brook, & Conway, 2004). This is further corroborated by experimental studies on female rhesus macaques that displayed a strong increase in rough-and-tumble play after they had been prenatally treated with androgens (for a review, see Thornton, Zehr, & Loose, 2009).

Notably, testosterone levels differ between boys and girls from around Week 8 of gestation (Judd, Robinson, Young, & Jones, 1976) and are presumed to have organizational effects on the brain. In turn, they should affect sex differences in behavior later in life (for a review, see Cohen-Bendahan, van de Beek, & Berenbaum, 2005). In this context, one study indicated that girls and boys exposed to higher levels of testosterone in amniotic fluid showed more male-typical play behavior later in life (Auyeung et al., 2009). By contrast, two other amniocentesis studies were not able to find this relationship (Knickmeyer et al., 2005; van de Beek et al., 2009). However, all three studies used different measures to quantify sex-typed play preferences and the children differed in age.

Thus, amniocentesis studies, which are very rare due to the high effort involved, show inconsistent results. However, studies on the relationship between sex-typed play behavior and the ratio between the second and fourth digit length (a presumed marker for the prenatal testosterone level) show more consistent results (Hönekopp & Thierfelder, 2009; Mitsui et al., 2016; Wong & Hines, 2016). Following the observation that females have larger ratios between the second and fourth digit (2D:4D) than males, Manning, Scutt, Wilson, and Lewis-Jones (1998) suggested the 2D:4D as an easily accessible marker for prenatal testosterone exposure.

Two meta-analyses confirmed a sex difference in 2D:4D (Grimbos, Dawood, Burriss, Zucker, & Puts, 2010; Hönekopp & Watson, 2010). This sex difference is already established prenatally as suggested by two studies on aborted fetuses (Galis, Ten Broek, Van Dongen, & Wijnaendts, 2010; Malas, Dogan, Hilal Evcil, & Desdicioglu, 2006). In favor of 2D:4D as a marker for prenatal testosterone exposure, females with CAH tend to have masculinized (i.e., smaller) 2D:4D (Brown, Hines, Fane, & Breedlove, 2002; Ökten, Kalyoncu, & Yaris, 2002). In mice, increasing androgens and reducing estrogens in utero decreased 2D:4D (Zheng & Cohn, 2011).

Even though there is convincing evidence that 2D:4D is influenced prenatally by sex hormones, the specific relationship with prenatal sex hormone levels from amniotic fluid and umbilical cord blood in humans is still not fully understood. There are only two published studies that link 2D:4D to prenatal

sex hormones from amniotic fluid with inconsistent findings. One study found that more masculine 2D:4D on the right hands of 29 2-year-olds (not separated for sex) were associated with a higher ratio of testosterone to estradiol levels (Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004), while another study found that the right and left 2D:4D in newborn girls (but not boys) to be correlated with amniotic fluid testosterone (Ventura, Gomes, Pita, Neto, & Taylor, 2013). Only one study using umbilical cord blood (sampled at birth) to measure sex hormones found an expected negative relationship between testosterone and left 2D:4D in girls (Whitehouse et al., 2015), while others failed to show this negative relationship (Çetin, Can, & Özcan, 2016; Hickey et al., 2010; Hollier et al., 2015). Additionally, in case testosterone effects on 2D:4D are mediated by androgen receptor type, there should be a positive correlation between functional androgen receptor gene variation (CAG stretches) and 2D:4D. Two meta-analyses have failed to show a relationship between 2D:4D and CAG stretches (Hönekopp, 2013; Voracek, 2014). Therefore, 2D:4D as a marker for prenatal hormone exposure should be interpreted with caution (for an overview of existing evidence, see Richards, 2017).

Assuming that 2D:4D is determined (at least partially) by prenatal hormones, the temporal stability of the sex difference in 2D:4D should be high, and sex differences should be present early in life. While a cross-sectional study on 2- to 25-year-old individuals found no age differences in 2D:4D (Manning et al., 1998), another cross-sectional study showed an increase in 2D:4D with age in 2- to 5-year-olds (Williams, Greenhalgh, & Manning, 2003). Rare longitudinal studies, however, are better suited to evaluate the stability of 2D:4D controlling for inter-individual differences. Such studies have shown a slight increase in 2D:4D with age (McIntyre, Cohn, & Ellison, 2006: T1: 6–7 years, T2: 8–9 years; McIntyre, Ellison, Lieberman, Demerath, & Towne, 2005: T1: 1 year, T2: 5 years, T3: 9 years, T4: 13 years, T5: 17 years; Trivers, Manning, & Jacobson, 2006: T1: 7–13 years, T2: 11–17 years; Wong & Hines, 2016: T1: 20–40 months, T2: 26–47 months). However, high correlations between the measurements (Pearson's $r = .71-.88$) suggest a high temporal stability of 2D:4D. By contrast, one longitudinal study on 0- to 2-year-olds showed a decrease of 2D:4D in the first year and an increase in the second year of life and low correlations (Pearson's $r = .35-.53$) between measurements (Knickmeyer, Woolson, Hamer, Konneker, & Gilmore, 2011: T1: 2 weeks, T2: 12 months, T3: 24 months).

Supporting the validity of 2D:4D, three studies (Hönekopp & Thierfelder, 2009; Mitsui et al., 2016; Wong & Hines, 2016) have shown that children with lower 2D:4D (and, thus, supposedly higher testosterone exposure in utero) display more masculine play behavior (as described by the parents' answers on the Preschool Activities Inventory; PSAI; Golombok & Rust, 1993). Nevertheless, there is considerable discrepancy concerning the side of the hand (right/left) and the sex of the children in which the correlations were found (Hönekopp &

Thierfelder, 2009: left 2D:4D of boys; Mitsui et al., 2016: right and left 2D:4D of boys; Wong & Hines, 2016: right 2D:4D of boys and right and left 2D:4D of girls).

In addition to a likely biological effect on play behavior, socialization plays an important role. It is well known that children's (sex-specific) toy preferences are influenced by parents, teachers, (older) siblings, and peers. For example, children are often reinforced for sex-congruent behavior (for reviews, see Berenbaum, Blakemore, & Beltz, 2011; Hines, 2010). Whereas the reinforcement of sex-congruent behavior of parents and teachers and the influence of peers is difficult to quantify, studies that have recorded the number of older brothers and sisters have shown that both girls and boys with more older brothers displayed more male-typical and less female-typical behavior and with more older sisters more female-typical and less male-typical play behavior (Hines et al., 2002b; Mitsui et al., 2016; Rust et al., 2000). According to social cognitive theory, this effect of older siblings is explained by observational learning and thus as a socialization factor (Berenbaum et al., 2011; Rust et al., 2000), which seems plausible because children may play with older siblings and their toys. However, it cannot be ruled out that the influence of older siblings on the behavior of younger siblings may also (partly) be based on a genetic or hormonal and, thus, a biological component (Berenbaum et al., 2011). In this context, the fraternal birth-order effect is worth mentioning as it describes the increased probability for a younger brother to be gay with an increasing number of older brothers. This effect is explained as a consequence of a progressive immunization of some mothers against male antigens with each pregnancy with a male fetus and a simultaneous increase in antibodies that affect the sexual differentiation of the brain (for a meta-analysis, see Blanchard, 2018). Bogaert et al. (2018) found antibody levels against the Y-linked protein NLGN4Y that is important in brain development, to be higher in mothers of gay sons than in the control samples. Additionally, in a large community sample, later fraternal birth order was related to elevated gender variance in boys (Coomer, Skorska, van der Miesen, Peragine, & VanderLaan, 2018) which indicates that the progressive immunization hypothesis is not only valid for homosexuality but also for sex-typed behavior in childhood. However, this effect has only been shown for boys with older brothers, so that the observed relationship between the sex-typed play behavior of boys and the number of older sisters as well as the relationship between the sex-typed play behavior of girls and the number of older brothers and sisters (Hines et al., 2002b; Mitsui et al., 2016; Rust et al., 2000) cannot be attributed to the birth-order effect and therefore most likely indicates socializing effects. To date, there are no studies that investigated both the influence of older siblings and prenatal hormonal effects in order to be able to make a statement as to whether the effects are additive or interactive (Berenbaum et al., 2011).

The current study aims at clarifying the relationship between sex-typed play behavior and 2D:4D. Only one of the three studies examining this relationship (Hönekopp & Thierfelder, 2009; Mitsui et al., 2016; Wong & Hines, 2016) assessed digit ratio in a longitudinal design (with only two measurements on 2- to 3-year-olds over a 6- to 8-month period; Wong & Hines, 2016). By contrast, the present longitudinal study consisted of four measurements of the digit ratios from both hands at different ages starting in early infancy (T1: 5 months, T2: 9 months, T3: 20 months, and T4: 40 months). The sex differences in 2D:4D should be stable if they are indeed influenced by prenatal testosterone. Therefore, we need multiple measurements at different ages during early infancy and childhood to draw conclusions about 2D:4D as a potential marker for prenatal testosterone. Our longitudinal design has a clear advantage over a cross-sectional design because it reduces inter-individual variance and increases the signal-to-noise ratio. At T4, parents completed the PSAI (Golombok & Rust, 1993) to record sex-typed play behavior of their children. The PSAI is a standardized, frequently used measure for play preferences that shows large sex differences (Hönekopp & Thierfelder, 2009; Mitsui et al., 2016; Wong & Hines, 2016) which are typically larger than in direct observations of toy preferences in a single, unnatural laboratory situation (Hines, 2010; Wong & Hines, 2016). It also has the advantage that it not only asks for toy preferences, but also for activity preferences and temperamental characteristics during the last month, giving a more comprehensive picture of sex-typed play behavior in comparison with observational studies of toy preferences.

Additionally, the number of older brothers and sisters (living in the same household) was assessed as an indicator for socialization effects on sex-typed play behavior (Mitsui et al., 2016; Rust et al., 2000). This allows us, for the first time, to assess the combination of potential biological (2D:4D) and socialization influences (older siblings) on sex-typed play behavior (Berenbaum et al., 2011).

Based on previous research, we predicted more male-typical play behavior in boys than in girls (operationalized by the PSAI score) and lower digit ratios in boys than in girls, independent of age and hand (right/left). For the temporal stability of 2D:4D, we expected a slight increase with age independent of the sex difference. Moreover, we predicted that boys and girls with lower digit ratios (supposedly higher prenatal testosterone exposure) should display more male-typical and less female-typical play behavior (higher PSAI scores). With respect to older siblings, more older brothers should lead to more masculine and less feminine play behavior (higher PSAI scores) and more older sisters to less masculine and more feminine behavior (lower PSAI scores) in both boys and girls.

Method

Participants

Between 2010 and 2012, *Praenatal.de* (Düsseldorf, Germany) recruited mothers who underwent prenatal screening. A total of 388 mothers agreed to be contacted by the Department of Experimental Psychology of the Heinrich Heine University Düsseldorf after birth of their children. The prerequisites for participation in the study were no identified prenatal abnormalities of the child and an uncomplicated course of pregnancy. We did not contact mothers that had significant problems later during pregnancy or childbirth. All remaining 274 mothers whose children had an APGAR score (Apgar, 1953) of at least nine were contacted for the study. All children were Caucasian and were born between January 2011 and February 2013 in and around Düsseldorf.

Children were tested on four occasions, at the age of 5 months (T1), 9 months (T2), 20 months (T3), and 40 months (T4; for the number of participants and mean ages at all four time points, see Table 1) in the Department of Experimental Psychology at the University of Düsseldorf. For 78 boys and 75 girls, parents' answers on the Preschool Activities Inventory (PSAI; Golombok & Rust, 1993) were assessed at T4.

At each measurement, parents gave informed consent for participation, recording, and storage of data and received a refund of their travel expenses. The study was approved by the local Ethics Committee of the Science Faculty of the University of Düsseldorf, Germany.

Materials and Procedure

Preschool Activities Inventory

The German version of the standardized questionnaire Preschool Activities Inventory (PSAI; Golombok & Rust, 1993)

by Hönekopp and Thierfelder (2009) was used to measure the children's sex-typed play behavior. The questionnaire was presented on a laptop screen with the software *Presentation* (Neurobehavioral Systems, USA) and completed by the accompanying parent (116 mothers, 9 fathers, 28 both parents) at T4 (age: 40 months). The PSAI consists of 24 items assessing children's toy preferences (7 items), activity preferences (11 items), and temperamental characteristics (6 items). Twelve items assess typical masculine and 12 typical feminine behavior on a 5-point scale. The PSAI is scored by subtracting the sum of the "female items" from the sum of "the male items," and transforming the score into a pseudo-*T* scale (Golombok & Rust, 1993). Thus, a higher score indicates more masculine behavior and a lower score more feminine behavior.

2D:4D

As hand scans (or photocopies) produce both larger sex differences (Hönekopp & Watson, 2010) and a higher measurement precision (Kemper & Schwerdtfeger, 2009; Mikac, Buško, Sommer, & Hildebrandt, 2016) than direct measurements (with calipers), both hands of the children were scanned at T1, T2, T3, and T4, respectively. The freeware program *Autometric* (DeBruine, 2004) was used to determine the ratio between the second and the fourth digit (2D:4D, the midpoint of the ventral proximal crease of the second digit to the tip of the second digit, and the ventral proximal crease of the fourth digit to the tip of the fourth digit). *Autometric* has been specially designed to measure finger length ratios and has been shown to be the superior program for indirect 2D:4D measurement compared to other computer-based measurement methods because of its high reliability (Kemper & Schwerdtfeger, 2009; Mikac et al., 2016). Scans in which the tips or the ventral creases could not be identified were excluded. One rater measured all hand scans (max. 8 per child) twice, and a second independent rater measured all hand scans once. The raters were blind to the sex of the children. The intra- and inter-rater reliabilities were determined

Table 1 Number and age (in months) of participants at T1–T4

| Time point | | Boys | Girls | Total |
|---------------|----------|-------|-------|-------|
| T1: 5 months | <i>N</i> | 114 | 111 | 225 |
| | <i>M</i> | 5.40 | 5.45 | 5.43 |
| | SD | 0.29 | 0.31 | 0.30 |
| T2: 9 months | <i>N</i> | 101 | 91 | 192 |
| | <i>M</i> | 9.38 | 9.36 | 9.37 |
| | SD | 0.35 | 0.39 | 0.37 |
| T3: 20 months | <i>N</i> | 86 | 80 | 166 |
| | <i>M</i> | 20.54 | 20.52 | 20.53 |
| | SD | 0.34 | 0.44 | 0.39 |
| T4: 40 months | <i>N</i> | 80 | 78 | 158 |
| | <i>M</i> | 40.49 | 40.55 | 40.52 |
| | SD | 0.51 | 0.65 | 0.59 |

with intra-class correlations (all ICC > .90; see Table 2). The three measurements were averaged to increase reliability. Due to dropout of participants in the course of the study and exclusion of non-measurable hand scans, the number of available hand scans per child varied between one and eight scans. Since it is more difficult to scan the hands of babies and infants as they cannot be instructed to hold still and place the hand flat on the scanner, there are less measurable hand scans at T1 to T3 compared to T4 (measurable hand scans: T1: 64.9%, T2: 60.7%, T3: 60.8%, T4: 78.5%).

Statistical Analyses

To test whether age, sex, and hand affected 2D:4D, we calculated a multilevel linear regression model (cf. Kenny, Korchmaros, & Bolger, 2003; Krull & MacKinnon, 2001) using the R packages lme4 and lmerTest (Bates, Mächler, Bolker, & Walker, 2014; Kuznetsova, Brockhoff, & Christensen, 2014; R Core Team, 2019), with participants as random effects. As such, the analysis accounts for participant heterogeneity as well as missing data. We entered age (months since birth, z -standardized), sex (dummy-coded as boys = 0, girls = 1), and hand (dummy-coded as left = 0, right = 1) as predictors of 2D:4D (z -standardized). Thus, the intercept is the average z -standardized 2D:4D for left-handed boys (with z -standardized age = 0). To follow-up on the effect of age on 2D:4D, three additional multilevel linear regression models comparing the effect of age between T1 and T2, T2 and T3, and T3 and T4 were calculated. Mixed models were also used in the longitudinal study of 2D:4D by Knickmeyer et al. (2011).

The sex difference in sex-typed play behavior (PSAI) was evaluated using an independent sample t test. Pearson correlations were calculated to examine the relationships between PSAI and 2D:4D in both sexes separately for every time point and both hands. Correlations were tested for significant differences with Fisher z -transformations. Finally, multiple regression analyses (forced entry) were conducted to predict sex-specific play preference from the predictors 2D:4D (average of left and right hand) at T4 (age: 40 months), older brothers and older sisters, separately for girls and boys.

Significance levels were set to .05 for all comparisons. Effect sizes (Cohen's d and Pearson r) were interpreted according to Cohen (1988)—small effect $d \geq 0.20$ or $r \geq .10$,

medium effect $d \geq 0.50$ or $r \geq .30$, and large effect $d \geq 0.80$ or $r \geq .50$.

Results

The multilevel linear regression model indicated a significant main effect of sex. As expected, girls had larger 2D:4D ratios than boys (see Fig. 1). The regression weight indicated that this effect was in the order of 0.49 standard deviations. Further, there was a significant main effect of age with a standardized regression weight of 0.11. Older participants had larger 2D:4D. There was no main effect of hand (see Table 3). Three additional multilevel linear regression models compared the effect of age between T1 and T2, T2 and T3, and T3 and T4 (for the standardized regression weights, see Table 4). These models showed that age did not affect 2D:4D between T1 and T2, and between T2 and T3. However, there was a significant age effect between T3 and T4, indicating that older participants had larger 2D:4D (for descriptive statistics, see Table 5).

As expected, PSAI scores for boys ($M = 64.99$, $SD = 9.22$) were significantly higher than for girls (girls: $M = 36.03$, $SD = 10.07$), indicating more masculine and less feminine behavior, $t(151) = 18.52$, $p < .001$, $d = 3.00$. Since the parents completed the PSAI at T4, we do not have measurements of the sex-typed play preference for children that did not take part at T4. Consequently, the sample sizes for the correlational analyses of the relationship between 2D:4D and PSAI scores differ depending on the number of measureable hand scans and participation in T4 (for the results of the correlational analyses of every single measurement and Fisher's z -test for differences in correlations between girls and boys, see Table 5). In accordance with the hypothesis, girls with lower 2D:4D showed more masculine and less feminine behavior (higher PSAI scores). This correlation was shown for every time point and both hands except from right 2D:4D at T2 and T3. There were no significant correlations between 2D:4D and PSAI scores in boys.

Boys and girls did not differ in the number of older brothers, $t(150) = 0.26$, $p = .792$, or sister, $t(126.85) = 1.56$, $p = .122$, unequal variances. Correlations between the PSAI score and the number of older brothers and the number of older sisters revealed that, in girls, having more older brothers resulted in higher PSAI scores (more masculine and less

Table 2 Intra- and inter-rater reliabilities (ICCs) for 2D:4D

| | 5 mos. Right hand | 5 mos. Left hand | 9 mos. Right hand | 9 mos. Left hand | 20 mos. Right hand | 20 mos. Left hand | 40 mos. Right hand | 40 mos. Left hand |
|-------------|-------------------------|------------------------|-------------------------|------------------------|--------------------------|-------------------------|--------------------------|-------------------------|
| Intra-rater | .97 | .96 | .96 | .95 | .92 | .96 | .96 | .96 |
| Inter-rater | .94 | .94 | .96 | .95 | .94 | .92 | .95 | .96 |

mo. months, *Intra-rater* intra-class correlations for two measurements of first rater; *Inter-rater* intra-class correlations for mean of two measurements of first rater and measurement of second rater

Fig. 1 2D:4D at all ages (T1 = 5 months, T2 = 9 months, T3 = 20 months, T4 = 40 months) for both hands, separately for boys and girls. The error bars depict standard deviations. The number in the bars indicates the sample sizes



Table 3 Multilevel linear model results regarding effects of sex, age, and hand on 2D:4D

| Effect | Estimate | SE | df | t | p |
|-----------|----------|------|-------|------|--------|
| Intercept | -0.22 | 0.07 | 283.2 | 3.21 | .001 |
| Sex | 0.49 | 0.09 | 211.1 | 5.42 | < .001 |
| Age | 0.11 | 0.03 | 848.0 | 4.17 | < .001 |
| Hand | -0.04 | 0.05 | 782.8 | 0.74 | .458 |

SE standard error

Estimates represent standardized (age) and semi-standardized (sex, hand) regression weights. Analyses were performed with the R procedures lme4 and lmerTest with restricted maximum likelihood estimation. For sex, 0 = boys, 1 = girls. For hand, 0 = left, 1 = right

Table 4 Multilevel linear model results regarding the age effect on 2D:4D

| Measurement | Estimate | SE | df | t | p |
|-------------|----------|------|-------|------|--------|
| T1 vs. T2 | -0.02 | 0.04 | 437.6 | 0.56 | .579 |
| T2 vs. T3 | 0.03 | 0.04 | 379.7 | 0.56 | .574 |
| T3 vs. T4 | 0.16 | 0.04 | 375.8 | 3.88 | < .001 |

SE standard error

Estimates represent standardized regression weights. Analyses were performed with the R procedures lme4 and lmerTest with restricted maximum likelihood estimation. For sex, 0 = boys, 1 = girls. For hand, 0 = left, 1 = right. For age, T1 = 5 months (4.6–6.9 months), T2 = 9 months (7.6–10.3 months), T3 = 20 months (18.6–21.7 months), T4 = 40 months (38.3–42.6 months)

feminine behavior; $r(75) = .42, p < .001$) and having more older sisters resulted in lower PSAI scores (less masculine and more feminine behavior; $r(75) = -.28, p = .014$). There

were no significant correlations for boys (older brothers: $r(77) = .17, p = .137$; older sisters: $r(77) = -.07, p = .528$).

The regression analysis for girls identified a lower 2D:4D at T4 and a higher number of older brothers as significant predictors of more masculine-typed and less feminine-typed play behavior (higher PSAI scores; see Table 6). Partial correlations revealed that both predictors explained independent proportions of variance of the sex-specific play preference in girls (see Table 6). In boys, neither 2D:4D nor the number of older brothers or the number of older sisters significantly explained the PSAI scores (see Table 6).

Discussion

Our study aimed at clarifying the relationship between 2D:4D digit ratio as a putative marker for prenatal testosterone exposure and sex-typed play behavior. Additionally, we explored the effect of the number of older brothers and sisters as markers for the impact of socialization on sex-typed play behavior. As expected, a smaller 2D:4D was associated with more male-typical and less female-typical play behavior (higher PSAI scores), replicating previous studies (Hönekopp & Thierfelder, 2009; Mitsui et al., 2016; Wong & Hines, 2016). However, these expected correlations were only found for girls for every measurement and both hands except for right 2D:4D at T2 (age: 9 months) and T3 (age: 20 months). In addition to the relationship between digit ratio and play behavior, in the present sample, girls were reported to be more masculine and less feminine the more older brothers and less masculine and more feminine the more older sisters they had. These results match those observed in earlier studies (Mitsui et al., 2016; Rust et al., 2000), albeit in contrast to these studies, we found no

Table 5 Pearson correlations between PSAI and right and left 2D:4D at T1–T4 and Fisher's z -tests

| | Boys | | | Girls | | | Fisher's z -test | |
|----------------|------|------|-----|-------|--------|-----|--------------------|--------|
| | r | p | N | r | p | N | z | p |
| T1: Right hand | .08 | .308 | 46 | -.39 | .005** | 43 | 2.24 | .013* |
| T1: Left hand | -.01 | .481 | 47 | -.34 | .005** | 56 | 1.69 | .046* |
| T2: Right hand | .13 | .225 | 37 | .02 | .448 | 37 | 0.46 | .324 |
| T2: Left hand | .14 | .175 | 45 | -.29 | .029* | 45 | 2.01 | .022* |
| T3: Right hand | -.05 | .382 | 36 | -.21 | .093 | 43 | 0.69 | .244 |
| T3: Left hand | .15 | .190 | 38 | -.43 | .002** | 43 | 2.64 | .004** |
| T4: Right hand | -.20 | .069 | 55 | -.34 | .003** | 63 | 0.80 | .212 |
| T4: Left hand | -.20 | .063 | 58 | -.27 | .017* | 64 | 0.40 | .345 |

T1: age 5 months, T2: age 9 months, T3: age 20 months, T4: age 40 months

* $p < .05$, one-tailed; ** $p < .01$, one-tailed

Table 6 Linear model of predictors (2D:4D at T4, older brothers and older sisters) of sex-typed play preference (PSAI scores)

| Sex | R^2 | Predictors | b | SE b | β | p | Partial correlation |
|-------|-------|------------|---------|--------|---------|------|---------------------|
| Boys | .12 | 2D:4D | -66.60 | 36.02 | -.22 | .069 | -.23 |
| | | Brothers | 3.14 | 1.67 | .22 | .064 | .23 |
| | | Sisters | -2.36 | 2.17 | -.13 | .282 | -.13 |
| Girls | .30 | 2D:4D | -117.09 | 42.75 | -.29 | .008 | -.32* |
| | | Brothers | 5.02 | 1.53 | .36 | .002 | -.38* |
| | | Sisters | -2.83 | 1.67 | -.18 | .095 | -.21 |

2D:4D = averaged right and left digit ratios at T4 (age: 40 months); Brothers = number of older brothers, Sisters = number of older sisters

* $p < .05$

significant correlations for boys. Additionally, in girls, 2D:4D at T4 and older brothers explained independent proportions of variance of sex-typed play behavior.

The now repeatedly shown relationship between sex-typed play behavior and digit ratios further encourages the assumptions that (1) 2D:4D could be a marker for prenatal sex hormone exposure and (2) sex-typed play preference could be partly affected by these hormones that have organizational effects on the brain. However, we cannot infer from the available data whether 2D:4D is indeed influenced by prenatal sex hormones. Interestingly, in the female sample, six of eight correlations between 2D:4D (four time points, right and left hand) and sex-typed play preference (PSAI at T4) were significant, while Wong and Hines (2016) demonstrated significant correlations only for their older sample (mean age of the girls: $M = 36.10$ months, $SD = 5.78$ months). This difference might be due to the fact that Wong and Hines (2016) administered the PSAI, a questionnaire designed to assess the behavior of preschoolers, at both time points even though the children were strictly speaking not preschoolers at the time of the first measurement (mean age of the girls: $M = 29.22$ months, $SD = 5.51$ months).

In contrast to previous research (Hönekopp & Thierfelder, 2009; Mitsui et al., 2016; Wong & Hines, 2016), the expected correlations were only found for girls. One possible explanation could be a threshold effect with prenatal sex hormones affecting 2D:4D and/or play behavior to a specific point with higher levels having no further effect (Breedlove, 2010; Cohen-Bendahan et al., 2005). The fact that boys are prenatally exposed to much higher levels of testosterone than girls might explain the nonsignificant correlations for boys (Hines et al., 2002a). This hypothesis fits with studies on children affected by CAH. Many studies have shown behavioral differences between CAH girls and unaffected female controls but not between CAH boys and unaffected male controls (Berenbaum et al., 2011). Likewise, the only study that has found a significant negative relationship between amniotic testosterone levels and 2D:4D only found it for girls. It was argued that male fetuses are exposed to such high levels of testosterone that levels above average do not impact the 2D:4D ratio any further (Ventura et al., 2013). The only amniocentesis study that has found the expected positive correlation between male-typed play behavior (PSAI) and prenatal amniotic fluid testosterone has also shown a significantly stronger correlation for girls ($r = .42$) than for boys ($r = .20$;

Fisher's $z = 1.76$, $p = .04$; Auyeung et al., 2009). Another explanation could be the different socialization of boys and girls. Boys are generally more often reinforced for sex-typical behavior and more likely to be discouraged from sex-untypical behavior than girls (Fagot, 1978; Langlois & Downs, 1980; Pasternski et al., 2005). Accordingly, previous studies showed that, although the sex difference in play behavior persists throughout childhood, the preference of boys for male-typed play behavior increases with age, while that of girls for female-typed play behavior decreases in favor of an increasing preference for typically masculine toys (Servin et al., 1999; Todd et al., 2017). For this reason, a potential effect of prenatal sex hormones on play behavior in boys could have been masked by a stronger social impact on boys play behavior as compared to girls (Hines et al., 2002a, 2002b).

In addition to the presumed biological effect on sex-typed play preference, reflected in the shown correlations between 2D:4D and play preference (at least for girls), socialization seems to affect sex differences in behavior to an even greater extent (Berenbaum & Beltz, 2016; Berenbaum et al., 2011). In line with this, girls were more masculine and less feminine the more older brothers and less masculine and more feminine the more older sisters they had. It is worth highlighting that, at least for girls, our data provide some evidence that biology (digit ratio) and socialization (impact of older brothers) serve as independent predictors for the amount of sex-specific play behavior (Berenbaum et al., 2011). Nevertheless, it cannot be ruled out that the effect of older siblings on their younger siblings' play behavior is not only due to socializing but also biological. In this context, it is known that the probability for younger brothers to be gay as well as to show more gender variance in childhood increases with the number of older brothers (for a meta-analysis, see Blanchard, 2018; Coome et al., 2018). However, in the present study, the effect of older siblings on sex-typed play behavior could only be shown for girls for whom no such birth-order effect is known. Conversely, it is obvious that by growing up in the same household with older siblings, observational learning takes place which most likely affects the behavior of the younger siblings (Mitsui et al., 2016; Rust et al., 2000).

Besides the behavioral correlations, the present study, being the first longitudinal study with four measurements over infancy and early childhood, sheds light on the temporal stability of 2D:4D. The digit ratio has to be sexual dimorphic early in life to present a marker for prenatal testosterone exposure (Manning et al., 1998). The present study revealed a medium-sized sex difference in 2D:4D with higher values for girls than boys comparable to previous studies (for a meta-analysis, see Hönekopp & Watson, 2010). The sex difference in 2D:4D was present at all measurements, confirming previous research suggesting 2D:4D to be a reliable measure (McIntyre et al., 2005, 2006; Trivers et al., 2006; Wong & Hines, 2016). Nevertheless, in line with previous research,

2D:4D increased significantly in size (Knickmeyer et al., 2011; McIntyre et al., 2005; Trivers et al., 2006; Williams et al., 2003; Wong & Hines, 2016) from T3 (age: 20 months) to T4 (age: 40 months).

Conclusion

In our study, we set out to evaluate the relationship between sex-typed play behavior and digit ratios as a proposed marker for prenatal sex hormone exposure. In line with previous research (Mitsui et al., 2016; Wong & Hines, 2016), we found that girls with lower 2D:4D were described as behaving more masculine and less feminine, suggesting a biological component in the development of sex-related play preference. This biological effect acts independent of socialization from older brothers on the girls' play behavior. A key strength of the present study was its longitudinal design with four measurements throughout infancy and early childhood which demonstrated the stability of the sex difference in 2D:4D.

Acknowledgements Open Access funding provided by Projekt DEAL.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alexander, G. M., & Hines, M. (2002). Sex differences in response to children's toys in nonhuman primates (*Cercopithecus aethiops sabaeus*). *Evolution and Human Behavior*, 23(6), 467–479.
- Apgar, V. (1953). A proposal for a new method of evaluation of the newborn infant. *Current Researches in Anesthesia & Analgesia*, 32(4), 260–267.

- Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R. C., Taylor, K., Hackett, G., & Hines, M. (2009). Fetal testosterone predicts sexually differentiated childhood behavior in girls and in boys. *Psychological Science*, *20*(2), 144–148.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *Fitting linear mixed-effects models using lme4*. arXiv preprint [arXiv:1406.5823](https://arxiv.org/abs/1406.5823).
- Berenbaum, S. A., & Beltz, A. M. (2016). How early hormones shape gender development. *Current Opinion in Behavioral Sciences*, *7*, 53–60.
- Berenbaum, S. A., Blakemore, J. E. O., & Beltz, A. M. (2011). A role for biology in gender-related behavior. *Sex Roles*, *64*(11–12), 804–825.
- Berenbaum, S. A., & Hines, M. (1992). Early androgens are related to childhood sex-typed toy preferences. *Psychological Science*, *3*(3), 203–206.
- Blanchard, R. (2018). Fraternal birth order, family size, and male homosexuality: Meta-analysis of studies spanning 25 years. *Archives of Sexual Behavior*, *47*(1), 1–15.
- Bogaert, A. F., Skorska, M. N., Wang, C., Gabrie, J., MacNeil, A. J., Hoffarth, M. R., ... Blanchard, R. (2018). Male homosexuality and maternal immune responsivity to the Y-linked protein NLGN4Y. *Proceedings of the National Academy of Sciences*, *115*(2), 302–306.
- Breedlove, S. M. (2010). Minireview: Organizational hypothesis: Instances of the fingerpost. *Endocrinology*, *151*(9), 4116–4122.
- Brown, W. M., Hines, M., Fane, B. A., & Breedlove, S. M. (2002). Masculinized finger length patterns in human males and females with congenital adrenal hyperplasia. *Hormones and Behavior*, *42*(4), 380–386.
- Çetin, R., Can, M., & Özcan, E. (2016). The relationship between testosterone and oestrogen level of the cord blood and length of fingers of newborns 2D:4D. *Balikesir Health Sciences Journal*, *5*, 75–82.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
- Cohen-Bendahan, C. C. C., van de Beek, C., & Berenbaum, S. A. (2005). Prenatal sex hormone effects on child and adult sex-typed behavior: Methods and findings. *Neuroscience and Biobehavioral Reviews*, *29*(2), 353–384.
- Cooze, L. A., Skorska, M. N., van der Miesen, A. I., Peragine, D. E., & VanderLaan, D. P. (2018). An examination of the biodevelopment of gender expression in children ages 6-to 12-years. *Canadian Journal of Human Sexuality*, *27*(2), 133–143.
- Davis, J. T. M., & Hines, M. (2020). How large are gender differences in toy preferences? A systematic review and meta-analysis of toy preference research. *Archives of Sexual Behavior*, *49*, 373–394. <https://doi.org/10.1007/s10508-10019-01624-10507>.
- DeBruine, L. M. (2004). *AutoMetric software for measurement of 2D:4D ratios*. Retrieved March 25, 2020 from www.facelab.org/debruine/Programs/autometric.php.
- Fagot, B. I. (1978). The influence of sex of child on parental reactions to toddler children. *Child Development*, *49*(2), 459–465.
- Galis, F., Ten Broek, C. M. A., Van Dongen, S., & Wijnaendts, L. C. D. (2010). Sexual dimorphism in the prenatal digit ratio (2D:4D). *Archives of Sexual Behavior*, *39*(1), 57–62.
- Golombok, S., & Rust, J. (1993). The pre-school activities inventory: A standardized assessment of gender role in children. *Psychological Assessment*, *5*(2), 131–136.
- Golombok, S., Rust, J., Zervoulis, K., Croudace, T., Golding, J., & Hines, M. (2008). Developmental trajectories of sex-typed behavior in boys and girls: A longitudinal general population study of children aged 2.5–8 years. *Child Development*, *79*(5), 1583–1593.
- Grimbos, T., Dawood, K., Burriss, R. P., Zucker, K. J., & Puts, D. A. (2010). Sexual orientation and the second to fourth finger length ratio: A meta-analysis in men and women. *Behavioral Neuroscience*, *124*(2), 278–287.
- Hassett, J. M., Siebert, E. R., & Wallen, K. (2008). Sex differences in rhesus monkey toy preferences parallel those of children. *Hormones and Behavior*, *54*(3), 359–364.
- Hickey, M., Doherty, D., Hart, R., Norman, R., Mattes, E., Atkinson, H., & Sloboda, D. (2010). Maternal and umbilical cord androgen concentrations do not predict digit ratio (2D: 4D) in girls: A prospective cohort study. *Psychoneuroendocrinology*, *35*(8), 1235–1244.
- Hines, M. (2003). Sex steroids and human behavior: Prenatal androgen exposure and sex-typical play behavior in children. *Annals of the New York Academy of Sciences*, *1007*(1), 272–282.
- Hines, M. (2010). Sex-related variation in human behavior and the brain. *Trends in Cognitive Science*, *14*(10), 448–456.
- Hines, M., Brook, C., & Conway, G. S. (2004). Androgen and psychosexual development: Core gender identity, sexual orientation, and recalled childhood gender role behavior in women and men with congenital adrenal hyperplasia (CAH). *Journal of Sex Research*, *41*(1), 75–81.
- Hines, M., Golombok, S., Rust, J., Johnston, K. J., Golding, J., & Avon Longitudinal Study of Parents and Children Study Team. (2002a). Testosterone during pregnancy and gender role behavior of preschool children: A longitudinal, population study. *Child Development*, *73*(6), 1678–1687.
- Hines, M., Johnston, K. J., Golombok, S., Rust, J., Stevens, M., Golding, J., & ALSPAC Study Team. (2002b). Prenatal stress and gender role behavior in girls and boys: a longitudinal, population study. *Hormones and Behavior*, *42*(2), 126–134.
- Hollier, L. P., Keelan, J. A., Jamnadass, E. S., Maybery, M. T., Hickey, M., & Whitehouse, A. J. (2015). Adult digit ratio (2D: 4D) is not related to umbilical cord androgen or estrogen concentrations, their ratios or net bioactivity. *Early Human Development*, *91*(2), 111–117.
- Hönekopp, J. (2013). No evidence that 2D:4D is related to the number of CAG repeats in the androgen receptor gene. *Frontiers in Endocrinology*, *4*. <https://doi.org/10.3389/fendo.2013.00185>.
- Hönekopp, J., & Thierfelder, C. (2009). Relationships between digit ratio (2D:4D) and sex-typed play behavior in pre-school children. *Personality and Individual Differences*, *47*(7), 706–710.
- Hönekopp, J., & Watson, S. (2010). Meta-analysis of digit ratio 2D:4D shows greater sex difference in the right hand. *American Journal of Human Biology*, *22*(5), 619–630.
- Judd, H. L., Robinson, J. D., Young, P. E., & Jones, O. W. (1976). Amniotic-fluid testosterone levels in mid-pregnancy. *Obstetrics and Gynecology*, *48*(6), 690–692.
- Kemper, C. J., & Schwerdtfeger, A. (2009). Comparing indirect methods of digit ratio (2D:4D) measurement. *American Journal of Human Biology*, *21*(2), 188–191.
- Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods*, *8*(2), 115–128.
- Knickmeyer, R. C., Wheelwright, S., Taylor, K., Raggatt, P., Hackett, G., & Baron-Cohen, S. (2005). Gender-typed play and amniotic testosterone. *Developmental Psychology*, *41*(3), 517–528.
- Knickmeyer, R. C., Woolson, S., Hamer, R. M., Konneker, T., & Gilmore, J. H. (2011). 2D:4D ratios in the first 2 years of life: Stability and relation to testosterone exposure and sensitivity. *Hormones and Behavior*, *60*(3), 256–263.
- Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*, *36*(2), 249–277.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). *lmerTest: Tests for random and fixed effects for linear mixed effect models* [Software]. Retrieved March 25, 2020 from <http://cran.r-project.org/package=lmerTest>.

- Langlois, J. H., & Downs, A. C. (1980). Mothers, fathers, and peers as socialization agents of sex-typed play behaviors in young children. *Child Development, 51*(4), 1237–1247.
- Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., & Manning, J. T. (2004). 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development, 77*(1–2), 23–28.
- Malas, M. A., Dogan, S., Hilal Evcil, E., & Desdicioglu, K. (2006). Fetal development of the hand, digits and digit ratio (2D:4D). *Early Human Development, 82*(7), 469–475.
- Manning, J. T., Scutt, D., Wilson, J., & Lewis-Jones, D. I. (1998). The ratio of 2nd to 4th digit length: A predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. *Human Reproduction, 13*(11), 3000–3004.
- McIntyre, M. H., Cohn, B. A., & Ellison, P. T. (2006). Sex dimorphism in digital formulae of children. *American Journal of Physical Anthropology, 129*(1), 143–150.
- McIntyre, M. H., Ellison, P. T., Lieberman, D. E., Demerath, E., & Towne, B. (2005). The development of sex differences in digital formula from infancy in the Fels Longitudinal Study. *Proceedings of the Royal Society B: Biological Sciences, 272*(1571), 1473–1479.
- Mikac, U., Buško, V., Sommer, W., & Hildebrandt, A. (2016). Analysis of different sources of measurement error in determining second-to-fourth digit ratio, a potential indicator of perinatal sex hormones exposure. *Review of Psychology, 23*(1–2), 39–49.
- Mitsui, T., Araki, A., Miyashita, C., Ito, S., Ikeno, T., Sasaki, S., & Nonomura, K. (2016). The relationship between the second-to-fourth digit ratio and behavioral sexual dimorphism in school-aged children. *PLoS One, 11*. <https://doi.org/10.1371/journal.pone.0146849>.
- Nordenström, A., Servin, A., Bohlin, G., Larsson, A., & Wedell, A. (2002). Sex-typed toy play behavior correlates with the degree of prenatal androgen exposure assessed by CYP21 genotype in girls with congenital adrenal hyperplasia. *Journal of Clinical Endocrinology and Metabolism, 87*(11), 5119–5124.
- Ökten, A., Kalyoncu, M., & Yaris, N. (2002). The ratio of second- and fourth-digit lengths and congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Early Human Development, 70*(1–2), 47–54.
- Pasterski, V. L., Geffner, M. E., Brain, C., Hindmarsh, P., Brook, C., & Hines, M. (2005). Prenatal hormones and postnatal socialization by parents as determinants of male-typical toy play in girls with congenital adrenal hyperplasia. *Child Development, 76*(1), 264–278.
- Richards, G. (2017). What is the evidence for a link between digit ratio (2D:4D) and direct measures of prenatal sex hormones? *Early Human Development, 113*, 71–72.
- Rust, J., Golombok, S., Hines, M., Johnston, K., Golding, J., & Team, A. S. (2000). The role of brothers and sisters in the gender development of preschool children. *Journal of Experimental Child Psychology, 77*(4), 292–303.
- Servin, A., Bohlin, G., & Berlin, L. (1999). Sex differences in 1-, 3-, and 5-year-olds' toy-choice in a structured play-session. *Scandinavian Journal of Psychology, 40*(1), 43–48.
- R Core Team (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>.
- Thornton, J., Zehr, J. L., & Loose, M. D. (2009). Effects of prenatal androgens on rhesus monkeys: A model system to explore the organizational hypothesis in primates. *Hormones and Behavior, 55*(5), 633–644.
- Todd, B. K., Barry, J. A., & Thommessen, S. A. (2017). Preferences for 'gender-typed' toys in boys and girls aged 9 to 32 months. *Infant and Child Development, 26*(3). <https://doi.org/10.1002/icd.1986>.
- Todd, B. K., Fischer, R. A., Di Costa, S., Roestorf, A., Harbour, K., Hardiman, P., & Barry, J. A. (2018). Sex differences in children's toy preferences: A systematic review, meta-regression, and meta-analysis. *Infant and Child Development, 27*(2), 15–20. <https://doi.org/10.1002/icd.2064>.
- Trivers, R., Manning, J. T., & Jacobson, A. (2006). A longitudinal study of digit ratio (2D:4D) and other finger ratios in Jamaican children. *Hormones and Behavior, 49*(2), 150–156.
- van de Beek, C., van Goozen, S. H., Buitelaar, J. K., & Cohen-Kettenis, P. T. (2009). Prenatal sex hormones (maternal and amniotic fluid) and gender-related play behavior in 13-month-old infants. *Archives of Sexual Behavior, 38*(1), 6–15.
- Ventura, T., Gomes, M. C., Pita, A., Neto, M. T., & Taylor, A. (2013). Digit ratio (2D:4D) in newborns: Influences of prenatal testosterone and maternal environment. *Early Human Development, 89*(2), 107–112.
- Voracek, M. (2014). No effects of androgen receptor gene CAG and GGC repeat polymorphisms on digit ratio (2D:4D): A comprehensive meta-analysis and critical evaluation of research. *Evolution and Human Behavior, 35*(5), 430–437.
- Whitehouse, A. J., Gilani, S. Z., Shafait, F., Mian, A., Tan, D. W., Maybery, M. T., ... Goonawardene, M. (2015). Prenatal testosterone exposure is related to sexually dimorphic facial morphology in adulthood. *Proceedings of the Royal Society B, 282*(1816). <https://doi.org/10.1098/rspb.2015.1351>.
- Williams, J. H. G., Greenhalgh, K. D., & Manning, J. T. (2003). Second to fourth finger ratio and possible precursors of developmental psychopathology in preschool children. *Early Human Development, 72*(1), 57–65.
- Wong, W. I., & Hines, M. (2016). Interpreting digit ratio (2D:4D)-behavior correlations: 2D:4D sex difference, stability, and behavioral correlates and their replicability in young children. *Hormones and Behavior, 78*, 86–94.
- Zheng, Z. G., & Cohn, M. J. (2011). Developmental basis of sexually dimorphic digit ratios. *Proceedings of the National Academy of Sciences of the United States of America, 108*(39), 16289–16294.
- Zosuls, K. M., & Ruble, D. N. (2018). Gender-typed toy preferences among infants and toddlers. In E. S. Weisgram & L. M. Dinella (Eds.), *Gender typing of children's toys: How early play experiences impact development* (pp. 49–72). Washington, DC: American Psychological Association.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.