

Second to Fourth Digit Length Ratio (2D:4D) and Adult Sex Hormone Levels: New Data and a
Meta-Analytic Review

Running title: 2D:4D and adult sex hormone levels

[Accepted for Publication in Psychoneuroendocrinology. Final Unedited Draft]

Johannes Hönekopp*, Luise Bartholdt

Technische Universität Chemnitz, Institut für Psychologie, Wilhelm-Raabe Str. 43, D-09120
Chemnitz

Lothar Beier, Andreas Liebert

Klinikum Chemnitz, Institut für Laboratoriumsmedizin, Flemmingstr. 2, D-09009 Chemnitz

* Corresponding author. Address: Technische Universität Chemnitz, Institut für Psychologie,
Wilhelm-Raabe Str. 43, D-09120 Chemnitz, Germany; e-mail: johannes.hoenekopp@phil.tu-
chemnitz.de; phone: ++49 (371) 531 63 99; fax: ++49 (371) 531 63 33.

The relative length of the second (index) to the fourth (ring) finger (2D:4D) is a putative negative correlate of prenatal testosterone (T) exposure. Therefore, 2D:4D (and to a lesser extent D_{r-l} , the difference between 2D:4D in the right hand and in the left hand) has often been used to study effects of prenatal androgenization on human behavior and cognition. However, evidence suggests that 2D:4D may also be related to levels of circulating sex hormones in adults. This would question the validity of 2D:4D as a means of studying the effects of prenatal sex hormones. Here we present new data from two non-clinical samples (64 women and 102 men) regarding the relationships of 2D:4D and D_{r-l} with circulating sex hormone levels. We then present a meta-analytic review of all the present evidence regarding this issue. The results suggest that, in the normal population, 2D:4D and D_{r-l} are not associated with adult sex hormone levels. The findings from this current study add to the growing body of evidence demonstrating that 2D:4D is a suitable tool to study the effects of prenatal androgenization on human behavior and cognition.

Keywords: 2D:4D; finger length; prenatal androgenization; sex hormones; testosterone

1. Introduction

The ratio of the length of the second (index) to the fourth (ring) finger, 2D:4D, is sexually dimorphic in humans. Averaged across samples from various populations, female values were found to be about .25 standard deviations higher than male values (Manning et al., 2000).

Several lines of evidence suggest that foetal sex hormone (especially androgen) levels bring about the sex difference in 2D:4D and also affect intra-sexual variability in this measure. First, the sex difference in 2D:4D is already observable at the end of the first trimester of foetal development (Malas et al., 2006). Second, the development of both, genitals and digits, is controlled by the same genes HoxA and HoxD (Kondo et al., 1997). Third, the sex difference in 2D:4D appears unaffected by puberty as is evidenced by cross sectional (Manning et al., 1998) and longitudinal (McIntyre et al., 2005; Trivers et al., 2006) data. Fourth, right hand 2D:4D at the age of two years was found to be negatively correlated with the T/estrogen ratio as measured by amniocentesis in the second trimester (Lutchmaya et al., 2004). Fifth, individual 2D:4D values were shown to have high longitudinal stability from age 10 to age 14 (Trivers et al., 2006) and some stability from infancy throughout to age 17 (McIntyre et al., 2005). Sixth, children with congenital adrenal hyperplasia, a condition that results in abnormally high T levels during gestation, were found to have lower 2D:4D values than normal controls in two studies (Brown et al., 2002; Ökten et al., 2002). A third study did not find such a difference (Buck et al., 2003), but this study shows some methodological shortcomings (for a thorough critique, see McIntyre, 2006). Seventh, females with male co-twins appear to be exposed to higher T levels in utero and they have lower 2D:4D values than females with female co-twins (van Anders et al., 2006). Eighth, 2D:4D negatively correlates with T sensitivity in androgen receptors as measured by the number of CAG repeats in the androgen receptor gene (Manning et al., 2003). Ninth, autism is believed to result from prenatal hypermasculinization of the brain, and autistic subjects were found to have lower 2D:4D than normal controls (Manning et al., 2001).

For these reasons, 2D:4D may be a valid marker of prenatal T exposure. As prenatal T affects human behavior and cognition and since other ways of studying these effects in humans are laborious and pose various difficulties (Cohen-Bendahan et al., 2005; Collaer and Hines, 1995) 2D:4D has become popular as a means to study the effects of prenatal androgenization in humans, especially regarding sex-linked behaviors and traits (Manning, 2002a). In some areas, findings regarding the relationship between 2D:4D and research variables appear quite diverse. For example, a negative relationship between mental rotation ability and 2D:4D has been found for females but not for males (Burton et al., 2005), for males but not for females (Sanders et al., 2005), or for neither sex (Alexander, 2006; Coolican and Peters, 2003; Falter et al., 2006). In other areas, research results have been remarkably consistent. For example, all studies focusing at 2D:4D and athletic ability have found negative relationships (Hönekopp et al., 2006b; Manning, 2002a, 2002b; Pokrywka et al., 2005; Paul et al., 2006). Interestingly, evidence for homologous effects of prenatal steroids as indicated by 2D:4D and of circulating steroids has been found in some fields. For example, not only prenatal T but also circulating T appears to boost sport performance in men (Neave and Wolfson, 2003); and both, low prenatal androgenization as indicated by high 2D:4D values and high levels of circulating estrogen, appear to increase the risk of eating disorders in women (Klump et al., 2006).

In sum, 2D:4D is promising as a marker for studying effects of prenatal androgenization in humans (for a more thorough discussion, see McIntyre, 2006). However, questions remain open. For example, it is unclear why the sex difference in 2D:4D is so low (about .25 standard deviations) whereas the sex difference in prenatal T levels appears to be much higher ($d \approx 1.9$; cf. Knickmeyer et al., 2005; van de Beek et al., 2004). This difference in effect sizes suggests that other, yet to be discovered, factors than prenatal T affect 2D:4D and that 2D:4D may not be a very accurate indicator of prenatal T. To clarify the latter point, more longitudinal studies comparing prenatal T with 2D:4D values later in life would be of great value.

Here, we focus on another possible problem of 2D:4D: Potential associations between 2D:4D and levels of circulating sex hormones threaten its validity as a measure of prenatal androgenization as we argue below. Our aim is therefore to clarify the relationship between 2D:4D and circulating sex hormone levels. In order to do so, we present new data regarding associations between 2D:4D and current levels of various sex hormones in women and men. We then review the literature on the relationship between sex hormone levels and 2D:4D, using meta-analytic procedures when appropriate. However, before we turn to the new data and the literature review, we show why potential associations between adult sex hormone levels and 2D:4D would threaten the validity of the latter as a means of studying prenatal androgenization, and we argue why such associations might be expected.

As 2D:4D appears to reflect prenatal T levels, a correlation between 2D:4D and X (the study variable of interest) is usually interpreted as indicating an effect of prenatal T on X (see Figure 1a). Although this interpretation is plausible, it is not necessarily correct. 2D:4D may not exclusively reflect prenatal androgenization but also adult levels of sex hormones (see below). In this case, an observed relationship between 2D:4D and X may not arise from prenatal T but from activational effects of circulating sex hormones on X, from effects of X on circulating sex hormones, or from bidirectional causation (Figure 1b). Of course, an association between circulating adult hormones and 2D:4D does not undermine possible prenatal hormonal influences on digit ratios. But in this case, it may be wrong to conclude, on grounds of an observed relationship between the study variable and 2D:4D, that prenatal androgenization affects this variable. It is thus important to clarify the relationship between 2D:4D and adult circulating sex hormone levels, which is our objective.

Is it plausible to assume a correlation between 2D:4D and adult sex hormone titres? A causal effect of adult sex hormones on 2D:4D is unlikely because the onset of puberty does not change the sex difference in 2D:4D (Manning et al., 1998; McIntyre et al., 2005; Trivers et al., 2006). Nonetheless, theoretical considerations make an association plausible. This is, because

prenatal exposure to sex hormones, affecting 2D:4D, may be associated with adult circulating sex hormone levels (Figure 1c). Prenatal T appears to stem from the fetuses and not from their mothers. This is indicated by the fact that amniotic T levels are not correlated with maternal serum T levels (van de Beek et al, 2004). But if prenatal and adult T stem from the same organs their titres may show a positive correlation. This should effect an association between 2D:4D and adult levels of circulating sex hormones (Figure 1c).

So far, we have focused on a potential association between adult T levels and 2D:4D. For three reasons, however, it is necessary to broaden the scope to other sex hormones. First, 2D:4D may not exclusively reflect prenatal T levels but also prenatal estradiol (Lutchmaya et al, 2004). Second, complex relationships between sex hormones are common. And third, empirical findings suggest an association of 2D:4D with adult FSH and LH levels (see below).

We now turn to the empirical evidence regarding potential associations between 2D:4D and adult sex hormones. Indirect evidence supports the idea of an association between 2D:4D and adult T. Manning et al. (2002) found a negative relationship between parents' 2D:4D and offspring sex ratio (i.e., proportion of sons). Because parents' current T levels at the time of conception are positively related to sex ratio, this result suggests a negative association between 2D:4D and adult circulating T levels. Previous direct findings regarding associations between 2D:4D and adult sex hormone levels are listed in detail in Tables 3 and 4. In brief, these findings are incomplete and inconsistent. Relationships between 2D:4D and adult T have been studied best. Conflicting results have been obtained for women (e.g., van Anders and Hampson, 2005, vs. Benderlioglu and Nelson, 2004) and men (e.g., Manning et al., 1998, vs. Bang et al., 2005). Whereas studies on women have focused exclusively on T, studies on men have obtained (partly conflicting) results, which suggest an association of 2D:4D with FSH and LH (Manning et al., 2004, vs. Bang et al. 2005).

We present new data investigating, for the first time, the relationship of 2D:4D with 17- β -estradiol (E2), FSH, LH, and progesterone in women and with E2 in men. Studying E2 is

especially important because 2D:4D may be related to prenatal estradiol (Lutchmaya et al, 2004); studying FSH and LH in women is important because of the already mentioned findings in men (Bang et al. 2005; Manning et al., 2004). We also investigated the association between 2D:4D and T, because conflicting findings have been obtained before. We then review all studies that have investigated associations between adult sex hormone levels and 2D:4D, using meta-analytic techniques to summarize multiple results.

Synthesizing multiple results by meta-analyses has several advantages. First, meta-analyses have a higher statistical power than the primary studies they are based upon. Thus, they may reveal effects that tend not to be detected in single studies yield. Moreover, meta-analyses yield especially precise estimates of the size of such effects. In the present context, this is highly desirable because primary studies have rarely studied large samples. The statistical power of primary studies in this field also suffers from the limited reliability of single hormone measurements. Sex hormone levels tend to vary from day to day, across the day, and they respond to external events (van Anders and Watson, 2006). Thus, multiple hormone measures would be desirable to arrive at precise measurements. Probably for economic reasons, this has not been done in most studies reviewed here (including our own). Second, meta-analyses can resolve conflicts arising from conflicting primary results. Meta-analyses help to understand whether conflicting evidence stems from chance, from differences in statistical power, or from the fact that the studied effect varies across populations (e.g., clinical vs. normal samples) or across particulars of methods (e.g., time of hormone measurement). Conflicting evidence is likely due to particulars of samples or methods if effect sizes vary more strongly across studies than would be expected by chance (Shadish and Haddock, 1994).

Similar to 2D:4D, left hand 2D:4D minus right hand 2D:4D (D_{r-l}) has also been suggested to be a negative correlate of prenatal androgenization in humans (Manning, 2002a; Manning et al., 2003). As D_{r-l} is sometimes used as a putative negative measure of prenatal androgenization

(e.g., Manning, 2002a) we extend our present analyses to relationships between D_{r-1} and adult circulating sex hormone levels.

2. New data

2.1. Methods

The original female sample (Sample 1) consisted of 66 women who were recruited via newspaper advertisements, leaflets, and newsgroup postings in Chemnitz, Dresden, and Leipzig (Saxony, Germany). All women gave their informed consent to participating in a larger study on sex hormones and attractiveness. Only women who did not use hormonal contraceptives were eligible to participate. All women were nulliparous and happened to be Caucasian. One woman was dropped from the original sample because her finger lengths were not measured; another participant was dropped because her sex hormones could not be measured within the intended time frame (day four to ten after menstruation onset). The male sample (Sample 2) consisted of 102 men from Chemnitz and vicinity who participated in a study on physical fitness and attractiveness (Hönekopp et al., in press). For a large subgroup of the male sample, Hönekopp et al. (2006a) previously reported on the association between 2D:4D and sexual behavior, and Hönekopp et al. (2006b) previously reported on the association between 2D:4D and physical fitness.

Between 0800h and 1000h, 5 ml venous blood was taken from all participants. For women, this took place four to ten days after menstruation onset. We chose this time frame because sex hormone levels remain relatively stable across this part of the menstrual cycle. Serum was obtained, aliquoted, and stored at -20°C . Economic constraints caused the use of different methods of analysis for both samples. Electrochemiluminescence immunoassays Elecsys 17- β -estradiol, progesterone, LH, FSH (Roche Diagnostics GmbH, Mannheim, Germany) and LIA ACCESS Testosterone (Beckman Coulter Inc., Fullerton CA, USA) were used to analyze the women's samples. The kits show inter-assay variabilities $< 13\%$ (E2 and T), $< 12\%$

(progesterone), < 10% (SHBG), and < 5% (LH and FSH). Radioimmunoassay was used to analyze the men's samples for E2, T, and SHBG. The kits RIA ACTIVE™ TESTOSTERONE, RIA SHBG, and RIA ESTRADIOL ULTRA-SENSITIVE (Diagnostic System Laboratories, Webster TX) were used. These kits show inter-assay variabilities <10%, <5%, and <12%, respectively. All samples from women were analyzed together; the same was true for all samples from men. All analyzes were performed at the Klinikum Chemnitz, following the manufacturer's protocol. All control values were within the range specified by the manufacturer.

For the men, we computed a free androgen index (FAI) as total T/SHBG, which is highly correlated with bioavailable T (Nanjee and Wheeler, 1985).

Participants' finger lengths were measured from digital photographs of the palmar surfaces of both hands. Participants lightly pressed their palms against a glass plate perpendicularly mounted on a table. Pictures were then taken through the glass plate. Three people blind to all other data independently measured the finger lengths using morph man 2.01 software, which allows storing the coordinates of specified pixels. Men's 2D:4D were taken from Hönekopp et al. (2006b), who successfully used this method for measuring 2D:4D.

2.2. Results

2D:4D measurements proved to be highly reliable for women's right hands (Cronbach's $\alpha = .93$) and women's left hands ($\alpha = .94$). Accordingly, we averaged, for each hand, all three 2D:4D measurements. All analyses are based on these averages.

Descriptive statistics for the sample of 64 women are listed in Table 1. As expected, none of the hormone measurements were correlated with the time span between the onset of menstruation and serum extraction (all unsigned $r < .08$, all $P > .54$). Descriptive statistics for the sample of 102 men are also listed in Table 1.

In women, 2D:4D in the right and in the left hand were correlated $r = .70$ ($P < .001$); in men, this correlation was $r = .68$ ($P < .001$). As expected, women's as compared to men's

2D:4Ds were larger in the right hand (independent-samples t test: $t_{164} = 3.6, P < .001$) and in the left hand ($t_{164} = 2.9, P = .005$). In line with Putz et al. (2004), D_{r-l} did not differ significantly between the sexes ($t_{164} = 0.5, P = .48$).

The correlations between digit ratios and sex hormones are listed in Table 2 for both samples. None of them proved statistically significant. This includes LH and FSH in women, which have been found to be positively associated with 2D:4D in men before (e.g., Manning et al., 2004).

3. Meta-analytic review

We next review all findings on relationships between 2D:4D and adult levels of circulating sex hormones that are available at present. If more than one study addressed a particular relationship, we integrated the multiple study results using meta-analytic techniques.

3.1. Method

We used the data bases PsycINFO and Entrez PubMed to find relevant studies, using the search terms *2D:4D*, *finger length**, and *digit ratio*. We used the reference section in the retrieved documents to look for further studies pertaining to our issue. All studies that reported relationships between 2D:4D and adult levels of circulating sex hormones were included in our review. Our analyses were based on 10 documents reporting the results of 15 samples, plus Samples 1 and 2 reported above. Altogether, these 17 samples comprised 332 female and 850 male participants.

Relationships between 2D:4D and circulating levels of sex hormones may differ between women and men. Also, potential associations may depend on the hand from which 2D:4D is being measured (Manning, 2002a). We therefore analyzed all data separately for women and men, and for right hand 2D:4D, left hand 2D:4D, and D_{r-l} . Studies that analyzed T from saliva and FAI from serum were analyzed together because both methods measure bioavailable T.

Most studies provided the relevant effect sizes (i.e., correlations) directly. In some cases, we had to compute them from test statistics, which we did as suggested by Rosenthal (1994). If studies reported non-significant correlations but did not give the precise values, we wrote to the authors, who kindly provided the exact results. In all cases, the analyzed effect sizes were zero-order correlations. If more than one study reported a specific relationship, we computed the mean effect size from r -transformed correlations, weighting each effect size according to its sample size (i.e., with $n-3$). For each mean weighted effect size, we computed the lower and upper bound of the .95-confidence interval. We tested if the primary results varied more than would be expected by chance (Shadish & Haddock, 1994). Significantly heterogeneous results indicate that the studied effect depends on the population studied, on the method employed, or on some other systematic factor. If between-study variance was greater than expected by chance ($P \leq .05$), we used a random effects model to compute the confidence interval, following Equation 18-23 from Shadish and Haddock (1994). In this case, we also computed SD_{md} , which indicates how strongly the studied effect varies across populations, methods of study, or other systematic factors. If the results of the primary studies were homogeneous (i.e., varied within the limits expected by chance) the computation of the confidence interval followed a fixed effects model.

3.2. Results and discussion

We first turn to those associations between 2D:4D and adult sex hormones that have been studied in more than one sample. The primary findings are listed in the left part of Table 3. Its right part shows the aggregated results. In general, associations between sex hormones and digit ratios proved to be non-significant and the variance in results across studies was not greater than expected by chance. However, three statistically significant exceptions are noteworthy, all of which pertain to men. D_{r1} showed a low negative correlation with total T; right hand 2D:4D showed a small positive correlation with FSH; and results regarding right hand 2D:4D and LH showed substantial between-studies variance.

We next turn to those associations that have been studied in only one sample (Table 4). As can be seen, non-significant results (most of them stemming from our Samples 1 and 2) prevail. Two noteworthy exceptions stem from Manning et al. (2004), who found substantial associations of left hand 2D:4D with FSH and LH in men.

In sum, Tables 3 and 4 contain five unusual (i.e., statistically significant) results. It is noteworthy that all five exceptions are partly or entirely based on a highly atypical sample of infertile men, many of which probably had severely compromised testicular function (Manning et al., 2004). Three of the five exceptional findings involved at least one other sample; in all three cases, the sample of infertile men produced the most extreme result. When we omitted the results from this atypical sample from our analyses, the relationship between D_{r-l} and total T from serum in men was not statistically significant any more (mean weighted $r = -.13$, $r_{lower} = -.28$, $r_{upper} = .03$). Thus, only one out of five unusual results remained significant (Table 3, figure in bold face): Bang et al. (2005) found a small correlation between FSH and right hand 2D:4D in normal men. Given the high number (29) of remaining associations between 2D:4D or D_{r-l} and levels of circulating sex hormones that have been studied a single significant result is about what one would expect by chance.

Manning et al. (2004) conclude from their findings that the variability in circulating T levels among male patients of infertility clinics may be greater than in normal controls. An unusually high variance in circulating T may explain why there appears to be a relationship with 2D:4D in this sub-population.

4. General Discussion

The length ratio of the second (index) to the fourth (ring) finger (2D:4D), a putative negative correlate of prenatal T, has become a popular variable for studying effects of prenatal androgenization in humans. However, 2D:4D may not exclusively reflect prenatal sex hormones but also adult circulating sex hormone levels. Therefore, observed relations between 2D:4D and

variables of interest in adults may not reflect effects of prenatal androgenization, as is commonly assumed, but the effects of circulating sex hormones. Here, we set out to clarify the relationship between 2D:4D and adult sex hormone levels. We did so by providing new data from two non-clinical samples (comprising of 64 women and 102 men) and by reviewing and meta-analyzing the available evidence.

Our review pertains to 32 relationships between putative indicators of prenatal androgenization (left hand 2D:4D, right hand 2D:4D, D_{r-l}) and measures of circulating levels of sex hormones (total T, bioavailable T, E2, LH, FSH, and progesterone). Several significant relationships, involving T, FSH, and LH, were obtained for men. However, these relationships depended largely or entirely on one highly atypical sample of infertile men (Manning et al., 2004). Without the results from this sample, we obtained data on 29 relationships. Only one of them, pertaining to right hand 2D:4D and FSH in men, was statistically significant. For a number of reasons, this finding does not challenge the usefulness of 2D:4D as an indicator of prenatal androgenization. First, given the large number of associations examined, this finding may well have arisen from chance. Second, although statistically significant, the relationship was very small ($r = .12$). Third, and most importantly, considerations regarding the effects of sex hormones on behavior and cognition focus on T in males and on T, estrogen, and progesterone in females (e.g., Fitch and Denenberg, 1998; Keefe, 2002). We are not aware of results or considerations that suggest an effect of current FSH levels on men's behavior or cognition. For the theoretically more important hormones, T, estrogen, and progesterone, the present evidence does not suggest an association with 2D:4D or D_{r-l} in the normal population.

A word of caution is in order regarding the relationship between 2D:4D and circulating T levels in men. Recently, van den Bergh and Dewitte (2006) found that 2D:4D predicts T spikes triggered by sexual cues in men. Thus, researchers should ensure that obtained relationships between 2D:4D and actual behavior do not reflect effects of a transient T spike (e.g. caused by the presence of an attractive experimenter) on that behavior.

Our review was entirely based on published studies. The results of published studies may not be representative for the results of all studies that have investigated the particular question of interest. This is because statistically significant studies as compared to statistically non significant studies have a higher chance to get published. The resulting underrepresentation of non significant results is a potential threat to the validity of meta-analyses (e.g., Begg, 1994). This is not true in our case because unpublished, non significant results would strengthen but not weaken our conclusions.

However, one should keep in mind that the absence of evidence is not the same as the evidence of absence. Those of our results that did not pertain to T were often based on only one sample, and the respective confidence intervals were often large. Thus, we cannot rule out that future studies may reveal moderate associations between 2D:4D or D_{r-1} and levels of circulating sex hormones in the normal population. However, the present evidence suggests that neither 2D:4D nor D_{r-1} are associated with adults' sex hormone levels in healthy adults. This supports the validity of 2D:4D as a means to study the effects of prenatal androgenization on human behavior and cognition.

Acknowledgements

We would like to thank all participants. We also thank Nina Asperger, Astrid Lang, Ute Lausmann, and Anja Miethe for help with the collection of data. We are grateful to Georg Jahn and Udo Rudolph for helpful comments on an earlier draft. This research was supported by Deutsche Forschungsgemeinschaft grants HO 2506/1-1 and HO 2506/3-1 to J.H.

References

- Alexander, G. M., 2006. Associations among gender-linked toy preferences, digit ratio, and spatial ability: evidence from eye-tracking analysis. *Arch. Sex. Behav.* 35, 699-709.
- Bang, A.K., Carlsen, E., Holm, M., Petersen, J.H., Skakkebaek, N.E., Jørgensen, N., 2005. A study of finger lengths, semen quality and sex hormones in 360 young men from the general Danish population. *Hum. Reprod.* 20, 3109-3113.
- Begg, C.B., 1994. Publication bias. In: H. Cooper and L.V. Hedges (Eds.), *The Handbook of Research Synthesis*. Russell Sage, New York, pp. 399-410.
- Benderlioglu, Z., Nelson, R.J., 2004. Digit length ratios predict reactive aggression in women, but not in men. *Horm. Behav.* 46, 558-564.
- 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. *Horm. Behav.* 42, 380-386.
- Buck, J.J., Williams, R.M., Hughes, I.A., Acerini, C.L., 2003. In utero exposure and 2nd to 4th digit length ratio. Comparisons between healthy controls and females with classical congenital adrenal hyperplasia. *Hum. Reprod.* 18, 976-979.
- Burton, L.A., Henninger, D., Hafetz, J., 2005. Gender differences in relations of mental rotation, verbal fluency, and SAT scores to finger length ratios as hormonal indexes. *Dev. Neuropsychol.* 28, 493-505.
- 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. *Neurosci. Biobehav. Rev.* 29, 353-384.
- Collaer, M.L., Hines, M., 1995. Human behavioral sex differences: a role for gonadal hormones during early development? *Psych. Bull.* 118, 55-107.
- Coolican, J., Peters, M., 2003. Sexual dimorphism in the 2D/4D ratio and its relation to mental rotation performance. *Evol. Hum. Behav.* 24, 179-183.

- Falter, C.M., Arroyo, M., Davis, G.J., 2006. Testosterone: activation or organization of spatial cognition? *Biol. Psychol.* 73, 132-140.
- Fitch, R.H., Denenberg, V.H., 1998. A role of ovarian hormones in sexual differentiation of the brain. *Behav. Brain Sci.* 21, 311-352.
- Hönekopp, J., Manning, J.T., Müller, C., 2006b. Digit ratio (2D:4D) and physical fitness in males and females: evidence for effects of prenatal androgens on sexually selected traits. *Horm. Behav.* 49, 545-549.
- Hönekopp, J., Rudolph, U., Beier, L., Liebert, A., Müller, C., in press. Physical attractiveness of face and body as indicators of physical fitness. *Evol. Hum. Behav.*
- Hönekopp, J., Voracek, M., Manning, J.T., 2006a. 2nd to 4th digit ratio (2D:4D) and number of sex partners: evidence for effects of prenatal testosterone in men. *PNEC* 31, 30-37.
- Keefe, D.L., 2002. Sex hormones and neural mechanisms. *Arch. Sex. Behav.* 31, 401-403.
- Kempel, P., Gohlke, B., Klempau, J., Zinsberger, P., Reuter, M., Hennig, J., 2005. Second-to-fourth digit length, testosterone and spatial ability. *Intelligence* 33, 215-230.
- Klump, K.L., Gobrogge, K.L., Perkins, P.S., Thorne, D., Sisk, C.L., Breedlove, S.M., 2006. Preliminary evidence that gonadal hormones organize and activate disordered eating. *Psychol. Med.* 36, 539-546.
- Knickmeyer, R.C., Wheelwright, S., Taylor, K., Raggatt, P., Hackett, G., Baron-Cohen, S. 2005. Gender-typed play and amniotic testosterone. *Dev. Psychol.* 41, 517-528.
- Kondo, T., Zákány, J., Innis, J. W., Duboule, D., 1997. Of fingers, toes, and penises. *Nature* 390, 29.
- Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., Manning, J.T., 2004. 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Hum. Dev.* 77, 23-28.
- Malas, M.A., Dogan, S., Evcil, E.H., Desdicioglu, K., 2006. Fetal development of the hand, digits and digit ratio (2D:4D). *Early Hum. Dev.* 82, 469-475.

- Manning, J.T., 2002a. Digit ratio: a pointer to fertility, behavior, and health. Rutgers University Press, New Brunswick.
- Manning, J.T., 2002b. The ratio of 2nd to 4th digit length and performance in skiing. *J. Sports Med. Phys. Fitness* 42, 446-450.
- Manning, J.T., Barley, L., Walton, J., Lewis-Jones, D.I., Trivers, R.L., Singh, D., Thornhill, R., Rhode, P., Bereczkei, T., Henzi, P., Soler, M., Szwed, A., 2000. The 2nd:4th digit ratio, sexual dimorphism, population differences, and reproductive success: Evidence for sexually antagonistic genes? *Evol. Hum. Behav.* 21, 163-183.
- Manning, J.T., Baron-Cohen, S., Wheelwright, S., Sanders, G., 2001. The 2nd to 4th digit ratio and autism. *Dev. Med. Child Neurol.* 43, 160-164.
- Manning, J.T., Bundred, P.E., Newton, D.J., Flanagan, B.F., 2003. The second to fourth digit ratio and variation in the androgen receptor gene. *Evol. Hum. Behav.* 24, 399-405.
- Manning, J.T., Martin, S., Trivers, R.L., Soler, M., 2002. 2nd to 4th digit ratio and offspring sex ratio. *J. Theor. Biol.* 217, 93-95.
- 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 levels of testosterone, LH and oestrogen. *Hum. Reprod.* 13, 3000-3004.
- Manning, J.T., Wood, S., Vang, E., Walton, J., Bundred, P.E., van Heyningen, C., Lewis-Jones D.I., 2004. Second to fourth digit ratio (2D:4D) and testosterone in men. *Asian J Androl.* 6, 211-5.
- McIntyre, M.H., 2006. The use of digit ratios as markers for prenatal androgen action. *Reprod. Biol. Endocrin.*, 4:10.
- 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. *Proc. R. Soc. Lond. B* 272, 1473-1479.

- Nanjee, N.M., Wheeler, M.J., 1985. Plasma free testosterone – is an index sufficient? *Annals of Clin. Biochem.* 22, 387-390.
- Neave, N., Laing, S., Fink, B., Manning, J.T., 2003. Second to fourth digit ratio, testosterone and perceived male dominance. *Proc. R. Soc. Lond. B* 270, 2167-2172..
- Neave, N., Wolfson, S., 2003. Testosterone, territoriality, and the 'home advantage'. *Physiol. Behav.* 78, 269-275.
- 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 Hum. Dev.* 70, 47-54.
- Paul, S.N., Kato, B.S., Hunkin, J.L., Vivekanandan, S., Spector, T.D., 2006. The Big Finger: the second to fourth digit ratio is a predictor of sporting ability in women. *Brit. J. Sports Med.* 40, 981-983.
- Pokrywka, L., Rachon, D., Suchecka-Rachon, K., Bitel, L., 2005. The second to fourth digit ratio in elite and non-elite female athletes. *Am. J. Hum. Biol.* 17, 796-800.
- Putz, D.A., Gaulin, S.J.C., Sporter, R.J., McBurney, D.H., 2004. Sex hormones and finger length. What does 2D:4D indicate? *Evol. Hum. Behav.* 25, 182-199.
- Roney, J.R., Maestripieri, D., 2004. Relative digit lengths predict men's behaviour and attractiveness during social interactions with women. *Hum. Nature* 15, 271-282.
- Rosenthal, R., 1994. Parametric measures of effect size. In: H. Cooper and L.V. Hedges (Eds.), *The Handbook of Research Synthesis*. Russell Sage, New York, pp. 231-244.
- Sanders, G., Bereczkei, T., Csatho, A., Manning, J., 2005. The ratio of the 2nd to 4th finger length predicts spatial ability in men but not women. *Cortex* 41, 789-95.
- Shadish, W.R., Haddock, C.K., 1994. Combining estimates of effect sizes. In: H. Cooper and L.V. Hedges (Eds.), *The Handbook of Research Synthesis*. Russell Sage, New York, pp. 261-285.

- Trivers, R., Manning, J, Jacobson, A., 2006. A longitudinal study of digit ratio (2D:4D) and other finger ratios in Jamaican children. *Horm. Behav.* 49, 150-156.
- van Anders, S.M., Hampson, E., 2005. Testing the prenatal androgen hypothesis: measuring digit ratios, sexual orientation, and spatial abilities in adults. *Horm. Behav.* 47, 92-98.
- van Anders, S.M., Watson, N.V., 2006. Social neuroendocrinology. Effects of social contexts and behaviors on sex steroids in humans. *Hum. Nature* 17, 212-237.
- van Anders, S.M., Vernon, P.A., Wilbur, C.J., 2006. Finger-length ratios show evidence of prenatal hormone-transfer between opposite-sex twins. *Horm. Behav.* 49, 315-319.
- van de Beek, C., Thijssen, J.H.H., Cohen-Kettenis, P.T., van Goozen, S.H.M., Buitelaar, J.K., 2004. Relationships between sex hormones assessed in amniotic fluid, and maternal and umbilical cord serum: what is the best source of information to investigate the effects of fetal hormonal exposure? *Horm. and Behav.* 46, 663-669.
- van den Bergh, B., Dewitte, S., 2006. Digit ratio (2D:4D) moderates the impact of sexual cues on men's decisions in ultimatum games. *Proc. R. Soc. Lond. B* 273, 2091–2095.

Figure captions

Figure 1: How should a correlation between 2D:4D and the study variable (X) be interpreted?

Usually it is seen as indicating an effect of prenatal testosterone on X (a). However, given a correlation between 2D:4D and adult steroid levels, the same correlation may instead indicate unidirectional or bidirectional causation between adult hormone levels and X (b). The critical correlation between 2D:4D and adult steroids may arise from an association between foetal T and adult steroids (c). (Dotted lines indicate correlations, arrows indicate causation).

Table 1: Descriptive statistics for Sample 1 and 2 (M \pm SD).

	Sample 1	Sample 2
Age (years)	22.4 \pm 2.5	22.4 \pm 1.3
Progesteron (nmol/l)	2.2 \pm 0.8	
LH (U/l)	7.0 \pm 4.0	
FSH (U/l)	6.6 \pm 1.5	
E2 (pmol/l)	172 \pm 72	90 \pm 23
total T (nmol/l)	1.6 \pm 0.9	18.1 \pm 5.1
FAI		81.5 \pm 38.1
right hand 2D:4D	0.967 \pm 0.027	0.958 \pm 0.034
left hand 2D:4D	0.981 \pm 0.030	0.966 \pm 0.035
D _{r-l}	-0.005 \pm .022	-0.008 \pm .027

Table 2: Correlations between digit ratios and sex hormones.

	Sample 1			Sample 2		
	r2D:4D	l2D:4D	D _{r-l}	r2D:4D	l2D:4D	D _{r-l}
Progesteron	.16	.20	-.08			
LH	.12	-.03	.19			
FSH	.11	-.01	.14			
E2	-.01	.07	-.10	.04	.12	-.10
total T	.11	.10	-.01	.04	.10	-.08
FAI				.08	.11	-.04

Note: r2D:4D = right hand 2D:4D; l2D:4D = left hand 2D:4D. All $P \geq .13$.

Table 3. Associations between digit ratios and levels of circulating sex hormones in adults (multiple samples).

Single studies							Aggregated findings		
study	sex	<i>n</i>	age	sample	digits	<i>r</i>	<i>SD</i> _{rnd}	mean <i>r</i>	<i>r</i> _{low}
bioavailable testosterone									
Bang et al. (2005)	m	360	19.5±1.0	normal	r2D:4D	-.03			
Sample 2	m	102	22.4±1.3	normal	r2D:4D	.08			
Benderlioglu and Nelson (2004)	m	97	20.5±1.9	normal	r2D:4D	-.10 ^a			
Neave et al. (2003)	m	48	21.3±3.4	normal	r2D:4D	-.05	n.s.	-.02	-.1
Roney and Maestripieri (2004)	m	39	21.4±3.7	normal	r2D:4D	.20 ^a			
Falter et al. (2006)	m	35	24.1±3.0	normal	r2D:4D	-.21 ^a			
Kempel et al. (2005)	m	17	24.2±4.2	normal	r2D:4D	.03			
Sample 2	m	102	22.4±1.3	normal	l2D:4D	.11			
Benderlioglu and Nelson (2004)	m	97	20.5±1.9	normal	l2D:4D	.08 ^a			
Neave et al. (2003)	m	48	21.3±3.4	normal	l2D:4D	.03	n.s.	.09	-.0
Kempel et al. (2005)	m	17	24.2±4.2	normal	l2D:4D	.11			
Sample 2	m	102	22.4±1.3	normal	D _{r-1}	-.04			
Benderlioglu and Nelson (2004)	m	97	20.5±1.9	normal	D _{r-1}	-.20 ^a	n.s.	-.14	-.2
Benderlioglu and Nelson (2004)	f	77	20.5±1.9	normal	r2D:4D	.03 ^a			
van Anders et al. (2005)	f	75	23.3±??	normal	r2D:4D	-.10			
Kallai et al. (2005)	f	40	21.3±1.6	normal	r2D:4D	-.16	n.s.	-.05	-.1
Falter et al. (2006)	f	34	24.1±4.6	normal	r2D:4D	-.20 ^a			
Kempel et al. (2005)	f	23	23.5±4.3	normal	r2D:4D	-.22			
van Anders et al. (2005)	f	18	25.9±??	normal	r2D:4D	.56			
Benderlioglu and Nelson (2004)	f	77	20.5±1.9	normal	l2D:4D	.01 ^a			
van Anders et al. (2005)	f	75	23.3±??	normal	l2D:4D	-.04			
van Anders et al. (2005)	f	19	25.9±??	normal	l2D:4D	.11	n.s.	-.02	-.1
Kallai et al. (2005)	f	40	21.3±1.6	normal	l2D:4D	-.02			
Kempel et al. (2005)	f	23	23.5±4.3	normal	l2D:4D	-.02			

(Table 3 continues)

study	Single studies						Aggregated findings		
	sex	<i>n</i>	Age	sample	digits	<i>r</i>	<i>SD</i> _{rnd}	mean <i>r</i>	<i>r</i> _{low}
total testosterone									
Bang et al. (2005)	m	360	19.5±1.0	normal	r2D:4D	.03			
Sample 2	m	102	22.4±1.3	normal	r2D:4D	.04			
Manning et al. (1998)	m	58	34.1±0.7	clinical	r2D:4D	-.29	n.s.	-.02	-.5
Manning et al. (2004)	m	51	56.5±4.8	normal	r2D:4D	-.08			
Manning et al. (2004)	m	43	39.0±7.6	clinical	r2D:4D	-.09			
Sample 2	m	102	22.4±1.3	normal	l2D:4D	.10			
Manning et al. (1998)	m	58	34.1±0.7	clinical	l2D:4D	-.23	n.s.	.05	-.0
Manning et al. (2004)	m	51	56.5±4.8	normal	l2D:4D	.11			
Manning et al. (2004)	m	43	39.0±7.6	clinical	l2D:4D	.23			
Sample 2	m	102	22.4±1.3	normal	D _{rl}	-.08			
Manning et al. (2004)	m	51	56.5±4.8	normal	D _{rl}	-.23	n.s.	<u>-.20</u>	-.3
Manning et al. (2004)	m	43	39.0±7.6	clinical	D _{rl}	-.44			
FSH from serum									
Bang et al. (2005)	m	360	19.5±1.0	normal	r2D:4D	.12			
Manning et al. (2004)	m	43	39.0±7.6	clinical	r2D:4D	.39	n.s.	.15	.0
LH from serum									
Bang et al. (2005)	m	360	19.5±1.0	normal	r2D:4D	.04			
Manning et al. (2004)	m	43	39.0±7.6	clinical	r2D:4D	.50	<u>.34</u>	.09	-

Note: *SD*_{rnd} = unaccounted variability in effect sizes between studies; mean *r* = weighted mean

effect size; *r*_{lower}, *r*_{upper} = lower/upper bound of .95 confidence interval; r2D:4D = right hand

2D:4D; l2D:4D = left hand 2D:4D; ^a = data obtained through personal communication.

Exceptional (i.e., statistically significant) results are underlined. All involve one highly atypical

sample from Manning et al. (2004). The only result that remained statistically

significant after removing the highly atypical sample is printed in bold face.

Table 4. Associations between digit ratios and levels of circulating sex hormones in adults (single samples).

study	sex	<i>n</i>	age	sample	digits	<i>r</i>
FSH from serum						
Manning et al. (2004)	m	43	39.0±7.6	clinical	l2D:4D	<u>.33</u>
Manning et al. (2004)	m	43	39.0±7.6	clinical	D _{r-l}	.05
Sample 1	f	64	22.4±2.5	normal	r2D:4D	.11
Sample 1	f	64	22.4±2.5	normal	l2D:4D	-.01
Sample 1	f	64	22.4±2.5	normal	D _{r-l}	.14
LH from serum						
Manning et al. (2004)	m	43	39.0±7.6	clinical	l2D:4D	<u>.59</u>
Manning et al. (2004)	m	43	39.0±7.6	clinical	D _{r-l}	-.15
Sample 1	f	64	22.4±2.5	normal	r2D:4D	.12
Sample 1	f	64	22.4±2.5	normal	l2D:4D	-.03
Sample 1	f	64	22.4±2.5	normal	D _{r-l}	.19
total testosterone						
Sample 1	f	64	22.4±2.5	normal	r2D:4D	.11
Sample 1	f	64	22.4±2.5	normal	l2D:4D	.10
Sample 1	f	64	22.4±2.5	normal	D _{r-l}	-.01
bioavailable testosterone						
Benderlioglu et al. (2004)	f	77	20.5±1.9	normal	D _{r-l}	-.03 ^a
17-β-estradiol from serum						
Sample 2	m	102	22.4±1.3	normal	r2D:4D	.04
Sample 2	m	102	22.4±1.3	normal	l2D:4D	.12
Sample 2	m	102	22.4±1.3	normal	D _{r-l}	-.10
Sample 1	f	64	22.4±2.5	normal	r2D:4D	-.01
Sample 1	f	64	22.4±2.5	normal	l2D:4D	.07
Sample 1	f	64	22.4±2.5	normal	D _{r-l}	-.10
progesterone from serum						
Sample 1	f	64	22.4±2.5	normal	r2D:4D	.16
Sample 1	f	64	22.4±2.5	normal	l2D:4D	.20
Sample 1	f	64	22.4±2.5	normal	D _{r-l}	-.08

Note: r2D:4D = right hand 2D:4D; l2D:4D = left hand 2D:4D; ^a = data obtained through personal communication. Both statistically significant results (underlined) stem from the same atypical sample.

