Evolutionary Psychological Science
https://doi.org/10.1007/s40806-020-00247-9

RESEARCH ARTICLE

Digit Ratio (2D:4D), Transgendered Belief, and Transsexual Drug Therapy in the BBC Internet Study

John T. Manning[1] · Robert Trivers[2] · Bernhard Fink[3]

Received: 10 April 2020 / Revised: 19 May 2020 / Accepted: 21 May 2020
© Springer Nature Switzerland AG 2020

Abstract

Transgendered belief—the conviction that one is the opposite gender to one's natal gender—may be influenced by prenatal sex steroids. We consider this possibility by examining the relationship between digit ratio (2D:4D—a suggested correlate of fetal testosterone and estrogen), natal gender, felt gender, and transsexual drug therapy in a large online survey (the BBC Internet Study). There were 209,317 participants who reported their gender, their felt gender, and whether they were taking/had taken transsexual drug therapy (male-to-female (*MtF*) or female-to-male (*FtM*)). Participants included natal males who felt male ($M \rightarrow M$, $n = 104{,}939$) and those who felt female ($M \rightarrow F$, $n = 4760$) and natal females who felt female ($F \rightarrow F$, $n = 84{,}904$) and those who felt male ($F \rightarrow M$, $n = 4705$). Transsexual drug therapy (*MtF* and *FtM*) was reported by 830 and 223 participants, respectively. Digit length was determined by direct self-measurement. Mean 2D:4D of $M \rightarrow F$ and *MtF* individuals was higher (more "feminized") than for $M \rightarrow M$ and natal males, respectively. These effects were found in the total sample, the most numerous ethnic group (Whites) and the two largest national White samples (the UK and the USA). The mean 2D:4D of $F \rightarrow M$ and *FtM* participants did not differ from that of $F \rightarrow F$ and natal females, respectively. We conclude that $M \rightarrow F$ and *MtF* individuals may have experienced lower prenatal testosterone and higher estrogen than $M \rightarrow M$ and natal males, respectively. There was no evidence for an effect of prenatal sex steroids on transgendered belief or transsexualism in $F \rightarrow M$ and *FtM* individuals.

Keywords Digit ratio · Transsexual · Transgendered belief · Prenatal · Sex steroids

Introduction

Transgendered belief is the conviction that there is a mismatch between one's biological sex and gender. The experience of discomfort or distress because of such belief is termed gender dysphoria. Gender identity is a multi-factorial complex trait with significant heritable polygenic components (Polderman et al. 2018). With regard to masculinity/femininity, there is evidence for a general factor, which is dependent on personality and sociosexuality (Pozzebon et al. 2015). Findings from digit ratio (2D:4D) studies suggest that prenatal sex steroids, and in particular a balance of fetal testosterone and estrogen, influence masculinity/femininity (Manning et al. 2017; Salmon and Hehman 2018). Such reports give reason to suggest that 2D:4D may be related to transgendered belief and gender dysphoria. However, comparisons of mean 2D:4D in transgendered and control individuals have suffered from small sample sizes in the former and an overreliance on one measurement method (indirect experimenter-measurement of digits from photocopies; see Manning 2017).

Digit ratio, particularly right 2D:4D, has been suggested to be a negative correlate of prenatal testosterone and a positive correlate of prenatal estrogen that is stable during ontogeny (Manning et al. 1998; Manning 2002). It is sexually dimorphic (males < females), the dimorphism appears early in fetal development (Malas et al. 2006; Galis et al. 2010), and it is independent of subsequent age-related changes in digit length (Manning and Fink 2018; Králík et al. 2017). The purpose of this study was to investigate patterns of self-measured direct 2D:4D from a large ($n > 255{,}000$) multi-ethnic and multi-national Internet survey (the BBC Internet Study) in participants with transgendered belief ($n > 9000$) and those who received transsexual drug therapy ($n > 1000$).

✉ Bernhard Fink
bfink@gwdg.de

1 Applied Sports, Technology, Exercise, and Medicine (A-STEM), Swansea University, Swansea, UK

2 Biosocial Research Foundation, Southfield, St Elizabeth, Jamaica

3 Biosocial Science Information, Ortsstraße 63, 2362 Biedermannsdorf, Austria

For individuals taking transsexual drug therapy, male-to-female (*MtF*) transsexuals might have higher 2D:4D (lower prenatal testosterone and higher prenatal estrogen) than natal males (*MtF* > *Controls*). In contrast, female-to-male (*FtM*) transsexuals might have lower 2D:4D (higher prenatal testosterone and lower prenatal estrogen) than natal females (*FtM* < *Controls*). A meta-analysis of seven studies (Sadr et al. 2020) supported the hypothesis of *MtF* > *Controls* for the right ($d = 0.29$, $p = 0.0009$) and left ($d = 0.19$, $p = 0.01$) hand 2D:4D. The direction of effects for *FtM* was consistent with the hypothesized *FtM* < *Controls* relationship. However, there was substantial heterogeneity in the direction of differences between group means and neither the right ($d = 0.36$, $p = 0.12$) nor the left ($d = 0.20$, $p = 0.195$) hand 2D:4D showed significant effects. We note that in all seven studies, digit length was experimenter-measured, and in six of the studies, measurement was indirect from photocopies. The single exception to this was that of Leinung and Wu (2017) in which digits were measured directly. The authors of this study reported a significant *FtM* < *Controls* effect but a non-significant difference between *MtF* and natal male 2D:4D, i.e., an opposite trend to that of the other six studies using indirect measurements. The predominance of a single measurement type is not the only problem in this group of studies. In general, numbers of *MtF* and *FtM* participants were low. Sadr et al. (2020) reported the largest samples of transsexuals (*MtF*, $n = 88$; *FtM*, $n = 104$). In the remaining studies, numbers of *MtF* transsexuals varied from 39 to 70 and for *FtM* transsexuals from 17 to 50.

Large online studies that use direct self-measurement of digit length is one way of correcting the problems that may arise in consequence of relying on small samples employing a single indirect measurement method. Veale (2011) reported means for transsexuals from an online study that employed two forms of digit measurement, i.e., self-measurement with either an online ruler or a conventional ruler. There were no significant group differences in 2D:4D (Veale et al. 2010; Veale 2011). The numbers of participants were higher than those reported from experimenter-measured studies in this literature (*MtF* physical ruler $n = 191$ and online ruler $n = 117$; *FtM* physical ruler $n = 41$ and online ruler $n = 38$). However, compared with indirect 2D:4D data from photocopies, self-measured 2D:4Ds have high SDs, presumably resulting from high rates of random errors (Manning et al. 2007), and effect sizes with target traits are generally low. Thus, 2D:4D online studies must probably yield larger numbers of participants than those reported by Veale if they are to contribute useful data to this subject area.

We considered self-measured direct 2D:4D of individuals who reported transgendered belief and those who elect for transsexual drug therapy in the BBC Internet Study. It may be that the distress associated with transgender issues is less in the former than the latter, but we predicted that the pattern of relationship with 2D:4D would be essentially the same. In particular, we hypothesized that natal males who feel female (*M→F*) would have higher 2D:4D than male controls who feel male (*M→M*), i.e., *M→F* > *Controls*. In contrast, we predicted that natal females who feel male (*F→M*) would have lower 2D:4D than female controls (*F→F*), i.e. *F→M* < *Controls*. We expected the pattern of these relationships to map on to the associations between 2D:4D and transsexual drug therapy, such that natal males taking transsexual drug therapy (*MtF*) would have higher 2D:4D than male controls (*MtF* > *Controls*) and natal females taking such therapy (*FtM*) would have lower 2D:4D than control females, i.e., *FtM* < *Controls*.

Methods

The BBC Internet Study included around 200 questions concerning sex-dependent aspects of demographics and behavior along with self-measurement of morphological traits including the lengths of 2D and 4D. Reimers (2007) set out the history of the study together with details of recruitment and ethical issues in detail. It was made up of six sections, and all sections were completed by 255,116 participants. It was multi-ethnic and multi-national. The predominant ethnicity was "White," which was reported by 84.1% of participants. The most commonly represented nationalities were the UK (46.9%), the US (27.7%), Canada (5.2%), and Australia (3.6%) with 11 other nations represented by > 1000 participants.

The first items in the Study were *Gender* (male or female), *Age* (integer between 0 and 99), *Ethnicity* (Asian/Asian British, Black/Black British, Black other, Chinese, Middle/Near Eastern, Mixed Ethnic, White), and *Where I live* (United Kingdom, then 240 other countries). In addition, of relevance to this report, there were the following items: (i) *Regardless of your biological sex, what sex do you feel?* (male or female) and (ii) *Are you taking any of the following? This includes any you may have taken within the last 3 months. Tick only those that apply, otherwise leave blank:* male-to-female transsexual drug therapy or female-to-male transsexual drug therapy.

Participants self-measured 2D and 4D of right and left hands using the methodology of Manning et al. (1998). A diagram of the hand was provided, and they were instructed to measure their fingers on the ventral side of the digit with a conventional ruler. Measurements were reported to the nearest millimeter using dropdown menus, with values between 10 and 100 mm in 1 mm increments. The 2D:4D was calculated by dividing the 2D by 4D digit lengths. Manning et al. (2007) have examined the effects of sex and ethnicity on 2D:4D in the study. Means of 2D:4D from the study showed the expected sex difference (males < females), and this extends across ethnic groups. Moreover, there were also expected ethnic differences in 2D:4D in the direction of Whites > Asians > Chinese = Blacks. However, the standard deviations (SDs) of mean (*M*) 2D:4D were greater than expected from a consideration of experimenter-measured studies, presumably reflecting high rates of random error in self-measurement. Restricting the analyses to participants ≥ 18 years of age and a 2D:4D range of 0.80 to 1.20 had the effect of reducing SDs to about 0.05. We have adopted this range restriction for 2D:4D. However, it should be noted that SDs of about 0.05 are in excess of experimenter-measured reports of 2D:4D, which average about 0.03 (Manning et al. 2007).

Results

Transgendered Belief

With regard to males, 194,939 reported they "felt male" (*M*→*M*) and 4769 "felt female" (*M*→*F*). For females, 84,904 "felt female" (*F*→*F*) and 4705 "felt male" (*F*→*M*). Table 1 provides means and standard deviations (M ± SD) of right and left 2D:4D for natal gender and felt gender for males (*M*→*M* and *M*→*F*) and females (*F*→*F* and *F*→*M*).

Digit Ratio, Gender, and "Felt Gender" in the Total Sample

A two-way ANOVA with factors *Gender* and *Felt Gender* for right hand 2D:4D as dependent variable revealed effects for *Gender* ($F_{1,199313} = 253.12$, $p < 0.0001$) and *Felt Gender* ($F_{1,199313} = 9.55$, $p < 0.01$) and an interaction of *Gender* * *Felt Gender* ($F_{1,199313} = 22.71$, $p < 0.0001$). Post-hoc tests (Fisher's PLSD) showed females to have a higher mean difference in right 2D:4D compared with males ($p < 0.0001$), and felt females had higher right 2D:4D than felt males ($p < 0.0001$). *F*→*F* and *F*→*M* had similar mean right 2D:4D (Cohen's $d = 0.02$), and *M*→*M* had lower right 2D:4D than *M*→*F* ($d = 0.08$) (Fig. 1).

A similar picture pertained with left 2D:4D although effect sizes were smaller for *Gender* and the interaction of *Gender* * *Felt Gender*. A two-way ANOVA with factors *Gender* and *Felt Gender* gave effects for *Gender* ($F_{1,199313} = 128.57$, $p < 0.0001$) and *Felt Gender* ($F_{1,199313} = 10.93$, $p < 0.001$) and an interaction of *Gender* * *Felt Gender* ($F_{1,199313} = 13.82$, $p < 0.001$). Post-hoc tests showed females to have a higher mean difference in left 2D:4D compared with males ($p < 0.0001$), and felt females had higher right 2D:4D than felt males ($p < 0.0001$). *F*→*F* and *F*→*M* had identical mean left 2D:4D, and *M*→*M* had lower right 2D:4D than *M*→*F* ($d = 0.08$) (Fig. 1).

Ethnicity Effects

We performed a three-way ANOVA with factors *Gender*, *Felt Gender*, and *Ethnicity* for right or left 2D:4D as dependent variables. For right 2D:4D, there were significant effects of *Gender* ($F_{1,198158} = 40.69$, $p < 0.0001$) and *Ethnicity* ($F_{6,198185} = 10.66$, $p < 0.0001$) but no effect of *Felt Gender* ($F_{1,198158} = 3.23$, $p = 0.07$). However, there was an interaction of *Gender* * *Felt Gender* ($F_{1,198158} = 10.61$, $p < 0.001$). None of the remaining interactions was significant.

With regard to the *Gender* * *Felt Gender* interaction, for males, there were seven felt males/felt females ethnic pairs. Considering right 2D:4D, for six of these pairs, mean felt males 2D:4D was lower than mean felt females (Fig. 2). White was the largest ethnic group, and this showed a significant difference between *M*→*M* ($n = 89{,}428$, 2D:4D $M = 0.984 \pm 0.049$, 95% CI (0.984, 0.984)) and *M*→*F* ($n = 3616$, 2D:4D $M = 0.989 \pm 0.050$, 95% CI (0.984, 0.984), $d = .10$).

Table 1 Samples sizes and descriptive statistics (*M* and SDs) of right (R) and left (L) 2D:4D for natal gender and felt gender for males (*M*→*M* and *M*→*F*) and females (*F*→*F* and *F*→*M*) for the total sample and for White participants in the UK and the USA

	Natal gender and felt gender											
	M→*M*			*M*→*F*			*F*→*F*			*F*→*M*		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
All participants												
R2D:4D	104,939	0.984	0.049	4769	0.988	0.053	84,904	0.994	0.051	4705	0.995	0.055
L2D:4D	104,939	0.984	0.047	4769	0.988	0.053	84,904	0.992	0.049	4705	0.992	0.053
White participants R2D:4D by nation												
UK	45,431	0.985	0.047	1571	0.989	0.051	35,082	0.993	0.049	1470	0.995	0.055
USA	19,348	0.985	0.052	961	0.991	0.054	16,470	0.997	0.054	851	0.998	0.057
White participants L2D:4D by nation												
UK	45,431	0.986	0.045	1571	0.989	0.049	35,082	0.992	0.047	1470	0.993	0.051
USA	19,348	0.985	0.051	961	0.992	0.054	16,470	0.994	0.052	851	0.993	0.055

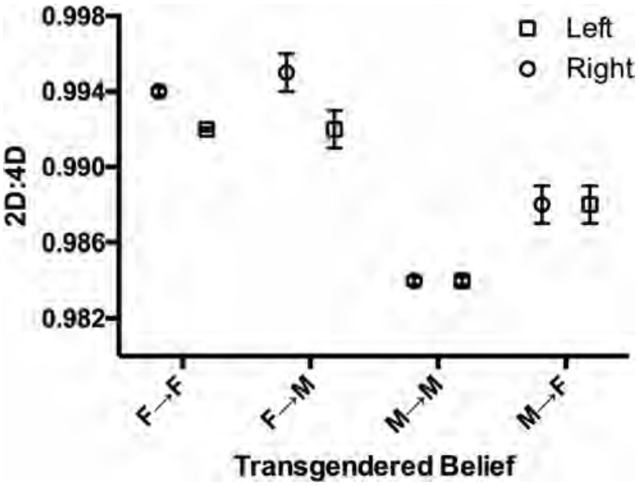

Fig. 1 Right and left mean 2D:4D (± SEM) for females who feel female (*F→F*), females who feel male (*F→M*), males who feel male (*M→M*), and males who feel female (*M→F*)

Thus, restricting ethnicity to White increased the effect size of *M→F* > *M→M* from $d = 0.08$ to $d = 0.10$. For females, there were five ethnic groups where mean 2D:4D felt females < felt males. White was the largest ethnic group with a marginally significant difference between *F→F* ($n = 73{,}106$, right 2D:4D $M = 0.994 \pm 0.051$), 95% CI (0.994, 0.994)) and *F→M* ($n = 3480$, right 2D:4D $M = 0.996 \pm 0.055$), 95% CI (0.994, 0.998), $d = 0.04$).

For left 2D:4D, there were effects of *Gender* ($F_{1,198158} = 27.94$, $p < 0.0001$) and *Ethnicity* ($F_{6,198185} = 20.90$, $p < 0.0001$) but no effect of *Felt Gender* ($F_{1,198158} = 0.28$, $p = 0.60$). However, there was an interaction of *Gender * Felt Gender* ($F_{1,198158} = 13.03$, $p < 0.001$). None of the remaining interactions were significant. Considering the *Gender * Felt Gender* interaction, for males, of the seven felt males/felt females ethnic pairs in five of these, the mean 2D:4D of felt males was lower than mean felt females. White was the largest ethnic group and *M→M* ($n = 89{,}428$, 2D:4D $M = 0.985 \pm 0.047$) was less than *M→F* ($n = 3616$, 2D:4D $M = 0.990 \pm 0.050$), $d = .10$). Thus, as with right 2D:4D, restricting ethnicity to White increased the effect size of *M→F* > *M→M* from $d = 0.08$ to $d = 0.10$. For females, there were six ethnic groups where mean 2D:4D felt female < felt male. However, in a consideration of the largest ethnic group (i.e., White), we found identical means for *F→F* ($n = 73{,}106$, 2D:4D $M = 0.993 \pm 0.049$) and *F→M* ($n = 3480$, 2D:4D $M = 0.993 \pm 0.052$).

National Effects among White Participants

We have shown that *F→M* individuals who have similar mean 2D:4D in comparison with *F→F* and *M→F* have significantly higher mean 2D:4D than *M→M*. The most numerous group in the study, i.e., White participants, drive the latter effect. In addition to sex and ethnic differences in 2D:4D, there are also national differences in mean 2D:4D among predominantly White nations (Manning et al. 2007). Thus, "felt gender" effects may also vary based on nationality. The problem with testing for such effects arises because *M→F* and *F→M* are present in a small proportion of the population (in the study about 4%). The four nations with the largest representation in the study are the UK (46.8%), the USA (27.7%), Canada (5.2%), and Australia (3.6%). For Canada and Australia, there were fewer than 200 White individuals in each of the groups of *M→F* and *F→M*. Therefore, we considered White participants from the UK and the USA in the analysis of national effects (see Table 1 for sample sizes).

We performed a three-way ANOVA with factors *Gender*, *Felt Gender*, and *Country* and right or left 2D:4D as dependent variable. For right 2D:4D, there were effects of *Gender* ($F_{1,121176} = 122.86$, $p < 0.0001$), *Felt Gender* ($F_{1,121176}) = 7.08$, $p < 0.01$), and *Country* ($F_{1,121176} = 9.42$, $p < 0.01$). There was also an interaction of *Gender * Felt Gender* (Fig. 3) ($F_{1,121176} = 20.26$, $p < 0.0001$). Post-hoc tests for *Gender* showed female > male ($p < 0.0001$), for *Felt Gender* felt females > felt males ($p < 0.0001$), and *Country* UK < USA ($p < 0.0001$). For left 2D:4D, there were effects of *Gender*

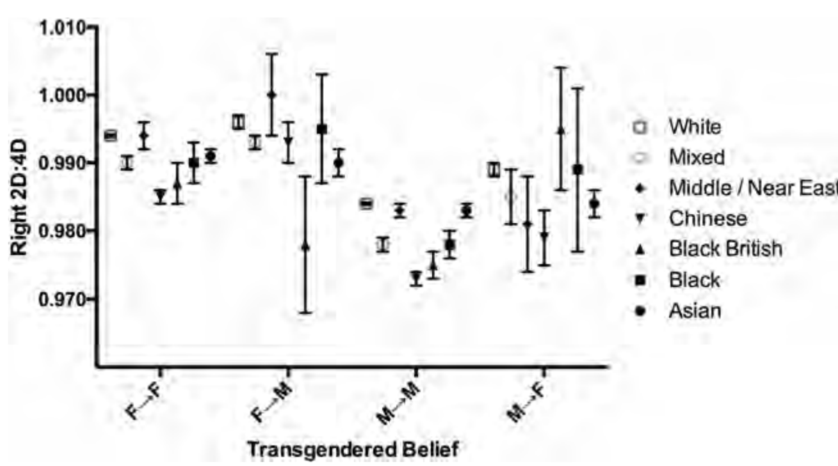

Fig. 2 Right mean 2D:4D (± SEM) by ethnicity for females who feel female (*F→F*); females who feel male (*F→M*), males who feel male (*M→M*), and males who feel female (*M→F*)

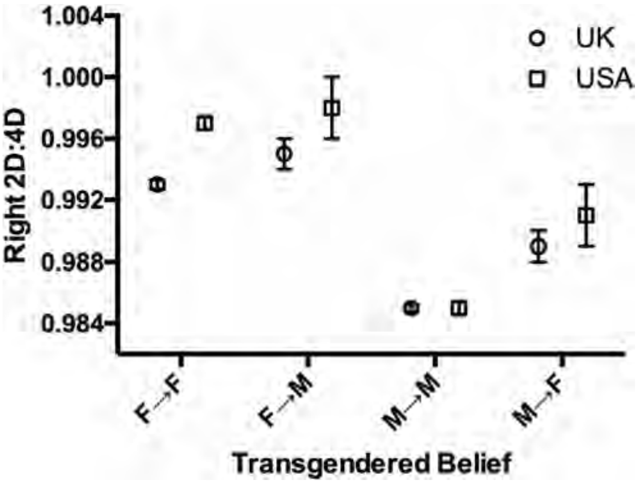

Fig. 3 White female and male right mean 2D:4D (± SEM) by nation (UK, USA) for females who feel female (*F→F*); females who feel male (*F→M*), males who feel male (*M→M*), and males who feel female (*M→F*)

($F_{1,121176} = 45.12$, $p < 0.0001$) and *Felt Gender* ($F_{1,121176}$) = 15.22, $p < 0.01$) but not for *Country* ($F_{1,121176} = 2.62$, $p = 0.11$). There was an interaction of *Gender* * *Felt Gender* ($F_{1,121176} = 9.99$, $p < 0.002$) and *Felt Gender* * *Country* ($F_{1,121176} = 4.18$, $p < 0.04$). Post-hoc tests for *Gender* showed female > male ($p < 0.0001$) for *Felt Gender* felt females > felt males ($p < 0.0001$) and *Country* UK < USA ($p < 0.0001$) (Fig. 3). Table 1 provides means and SDs of right and left 2D:4D for natal gender and felt gender for males (*M→M* and *M→F*) and females (*F→F* and *F→M*) for White participants in the UK and the USA.

Male-to-Female and Female-to-Male Transsexual Drug Therapy

There were 1053 participants who reported taking transsexual drug therapy. Of these, 830 were male-to-female (*MtF*) and 223 female-to-male (*FtM*) transsexuals. Means and SDs of right and left 2D:4D for natal male and female gender controls and transsexuals (*MtF* and *FtM*) for the total sample and for White participants only are provided in Table 2.

MtF Transsexuals: Total Sample

With regard to right 2D:4D, the mean value for *MtF* individuals ($n = 830$, $M = 0.989 \pm 0.050$, 95% CI (0.986, 0.992)) was significantly higher than the mean for males ($n = 110{,}518$, $M = 0.984 \pm 0.049$, 95% CI (0.984, 0.984), $d = .10$) but still significantly lower than the female mean ($n = 90{,}517$, $M = 0.994 \pm 0.051$, 95% CI (0.994, 0.994), $d = .10$). There was a similar pattern for left 2D:4D albeit with lower effect size (*MtF*, $n = 830$, $M = .988 \pm 0.049$, 95% CI (0.985, 0.991); males, $n = 110{,}518$, $M = 0.984 \pm 0.048$, 95% CI (0.984, 0.984); females, $n = 90{,}517$, $M = 0.992 \pm 0.049$, 95% CI (0.992, 0.992), $d = .08$).

Of the 830 *MtF* individuals, 393 reported they were female, and 437 reported they were male. This raised the possibility that within the sample, there may be a difference in 2D:4D between those that reported they were male but are now female and those that were male and at present remain male. We performed a two-way ANOVA with factors *Gender* (male or female*)* and *Gender after Therapy* (male or female) for right 2D:4D or left 2D:4D as dependent variables. For right 2D:4D, there was an effect of *Gender* ($F_{1,201861} = 8.81$, $p < 0.01$) but no effect of *Gender after Therapy* ($F_{1,201861} = 0.01$, $p = 0.94$). However, there was an interaction of *Gender* * *Gender after Therapy* ($F_{1,201861} = 7.45$, $p < 0.01$). The interaction arose because the *MtF* transsexuals who described themselves as female had lower 2D:4D than control females ($M = 0.989 \pm 0.048$ and $M = 0.994 \pm 0.051$, respectively), while *MtF* transsexuals who described themselves as male had higher 2D:4D than control males ($M = 0.988 \pm 0.052$ and $M = 0.984 \pm 0.049$, respectively). For left 2D:4D, there was no effect of Gender ($F_{1,201861} = 1.28$, $p = .26$) and no effect of Gender after Therapy ($F_{1,201861} = 0.03$, $p = .87$). However, there was an interaction of Gender * Gender after Therapy ($F_{1,201861} = 10.35$, $p < .001$). As with right 2D:4D the interaction arose because the MtF transsexuals who described themselves as female had lower 2D:4D than control females (.987 ± .044) and .992 ± .049 respectively) while MtF transsexuals who described themselves as male had higher 2D:4D than control males (.990 ± .053 and .984 ± .048, respectively) (Fig. 4).

Considering "felt gender," there were 681 *MtF* individuals who "felt female" and 141 who "felt male." The mean right 2D:4Ds were 0.987 ± 0.048 and 0.994 ± 0.058, respectively ($F_{1,820} = 2.18$, $p = 0.14$).

MtF Transsexuals: White Participants

Restricting the sample to the largest ethnic group (i.e., Whites), there were 702 *MtF* transsexuals (Table 2).

Considering right 2D:4D, the mean value for *MtF* individuals ($n = 702$, $M = 0.988 \pm 0.049$), 95% CI (0.984, 0.992)) was marginally significantly higher than the mean for males ($n = 93{,}476$, $M = 0.984 \pm 0.049$), 95% CI (0.984, 0.984), $d = .08$) but still significantly lower than the female mean ($n = 77{,}120$, $M = 0.994 \pm 0.051$), 95% CI (0.994, 0.994), $d = 0.12$). There was a similar pattern for left 2D:4D with marginally significant differences between *MtF* and male and female controls (*MtF*, $n = 702$, $M = 0.989 \pm 0.048$, 95% CI (0.985, 0.993); males, $n = 93{,}476$, $M = .985 \pm 0.047$, 95% CI (0.985, 0.985), $d = 0.08$; females, $n = 77{,}120$, $M = 0.993 \pm 0.049$, 95% CI (0.993, 0.993), $d = 0.08$).

Of the 702 *MtF* individuals, 342 reported they were female, and 360 reported they were male. We performed a two-way ANOVA with factors *Gender* and *Gender after Therapy* for right 2D:4D or left 2D:4D as dependent variables. For right 2D:4D, there was an effect of *Gender*

Table 2 Samples sizes and descriptive statistics (M, SDs, and SEMs) of right and left 2D:4D for natal male and female gender controls and transsexuals (*MtF* and *FtM*) for the total sample and for White participants

	n	*M*	*SD*	*SEM*	*n*	*M*	*SD*	*SEM*	*n*	*M*	*SD*	*SEM*
					Male-to-female							
	Transsexuals				Males				Females			
All participants												
R2D:4D	830	0.989	0.050	0.002	110,518	0.984	0.049	0.0002	90,517	0.994	0.051	0.0002
L2D:4D	830	0.988	0.049	0.002	110,518	0.984	0.048	0.0001	90,517	0.992	0.049	0.0002
White participants												
R2D:4D	702	0.988	0.049	0.002	93,476	0.984	0.049	0.0002	77,120	0.994	0.051	0.0002
L2D:4D	702	0.989	0.048	0.002	93,476	0.985	0.047	0.0002	77,120	0.993	0.049	0.0002
					Female-to-male							
	Transsexuals				Males				Females			
All participants												
R2D:4D	223	0.992	0.057	0.004	110,789	0.984	0.049	0.0002	90,853	0.994	0.051	0.0002
L2D:4D	223	0.989	0.055	0.004	110,789	0.984	0.048	0.0001	90,853	0.992	0.049	0.0002
White participants												
R2D:4D	144	0.993	0.059	0.005	93,722	0.984	0.049	0.0002	77,432	0.994	0.051	0.0002
L2D:4D	144	0.989	0.053	0.004	93,722	0.985	0.047	0.0002	77,432	0.993	0.049	0.0002

($F_{1,171294} = 7.04$, $p < 0.01$) but no effect of *Gender after Therapy* ($F_{1,171294} = 0.49$, $p = 0.48$). However, there was an interaction of *Gender * Gender after Therapy* ($F_{1,171294} = 7.04$, $p < 0.01$) such that the means for "female" ($M = 0.988 \pm 0.046$) and "male" ($M = 0.988 \pm 0.052$) *MtF* individuals were lower than that for females who were not transsexuals ($M = 0.994 \pm 0.051$) and higher than that of males who were not transsexuals ($M = .984 \pm 0.049$) (Fig. 5). For left 2D:4D, effect sizes were smaller with no effect of *Gender* ($F_{1,171294} = 0.17$, $p = 0.68$) or *Gender after Therapy* ($F_{1,171294} = 0.04$, $p = 0.84$), but as with right 2D:4D, there was an interaction of *Gender * Gender after Therapy* ($F_{1, 171294} = 13.73$, $p < 0.001$) (Fig. 5).

FtM Transsexuals: Total Sample

There were 223 *FtM* participants. The mean right 2D:4D for *FtM* individuals ($n = 223$, $M = 0.992 \pm 0.057$, 95% CI (0.985,

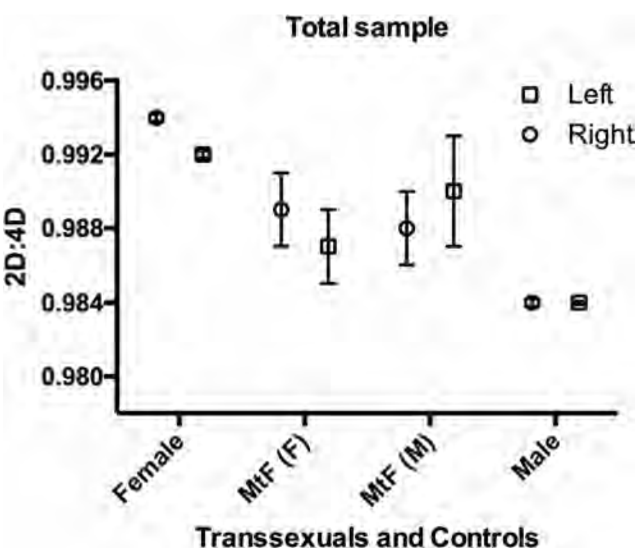

Fig. 4 Right and left mean 2D:4D (± SEM) for male-to-female transsexuals (*MtF*) reporting their gender as female (F) or male (M), natal female controls (female) and natal male controls (male)

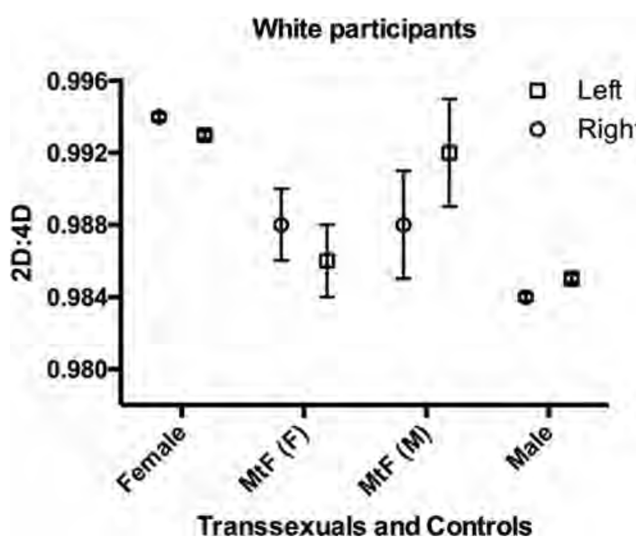

Fig. 5 White participants only: Right and left mean 2D:4D (± SEM) for male-to-female transsexuals (*MtF*) reporting their gender as female (F) or male (M), natal female controls (female), and natal male controls (male)

0.999)) did not differ significantly from the mean of the female controls ($n = 90{,}853$, $M = 0.994 \pm 0.051$), 95% CI (0.994, 0.994)) and was significantly higher than the mean for males ($n = 110{,}789$, $M = 0.984 \pm 0.049$, 95% CI (0.984, 0.984)). The mean left 2D:4D for *FtM* ($n = 223$, $M = 0.989 \pm 0.055$, 95% CI (0.982, 0.996)) did not significantly differ from that of control females ($n = 90{,}853$, $M = 0.992 \pm 0.049$, 95% CI (0.992, 0.992)) or control males ($n = 110{,}789$, $M = 0.984 \pm 0.048$, 95% CI (0.984, 0.984)).

Splitting the *FtM* sample into participants who reported they were male ($n = 166$) and those who regarded themselves as female ($n = 57$), we performed a two-way ANOVA with factors *Gender* and *Gender after Therapy* for right 2D:4D or left 2D:4D as dependent variables. For right 2D:4D, there was an effect of *Gender* ($F_{1,201861} = 5.33$, $p < 0.05$) but not for *Gender after Therapy* ($F_{1,201861} = 1.55$, $p = 0.21$), and there was no interaction of *Gender * Gender after Therapy* ($F_{1,201861} = 0.07, p = 0.80$). For left 2D:4D, there was no effect of *Gender* ($F_{1,201861} = 0.99, p = 0.32$) or *Gender after Therapy* ($F_{1,201861} = 0.03$, $p = 0.86$) and no interaction of *Gender * Gender after Therapy* ($F_{1,201861} = 0.93$, $p = 0.33$) (Fig. 5).

FtM Transsexuals: White Participants

Restricting the sample to the largest ethnic group (i.e., Whites), there were 144 *FtM* transsexuals. The mean right 2D:4D for White *FtM* individuals ($n = 114$, $M = 0.993 \pm 0.059$, 95% CI (0.983, 1.00)) did not differ significantly from the mean of the female controls ($n = 77{,}432$, $M = 0.994 \pm 0.051$, 95% CI (0.994, 0.994)) or from the mean for males ($n = 93{,}722$, $M = 0.984 \pm 0.049$, 95% CI (0.984, 0.984)). The mean left 2D:4D for *FtM* ($n = 144$, $M = 0.989 \pm 0.053$, 95% CI (0.980, 0.998)) did not significantly differ from that of control females ($n = 77{,}432$, $M = 0.993 \pm 0.049$, 95% CI (0.993, 0.993)) or control males ($n = 93{,}722$, $M = 0.985 \pm 0.047$, 95% CI (0.985, 0.985)).

We split the White *FtM* sample into participants who reported they were male ($n = 114$) and those who regarded themselves as female ($n = 30$). We performed a two-way ANOVA with factors *Gender* and *Gender after Therapy* for right 2D:4D or left 2D:4D as dependent variables. For right 2D:4D, there was no effect of *Gender* ($F_{1,171294} = 1.78$, $p = 0.18$) or *Gender after Therapy* ($F_{1,171294} = 0.73, p = 0.39$) and no interaction of *Gender * Gender after Therapy* ($F_{1,171294} = 0.38, p = 0.54$). For left 2D:4D, there was no effect of *Gender* ($F_{1,171294} = 0.62$, $p = 0.43$) or *Gender after Therapy* ($F_{1,171294} = 0.33$, $p = 0.94$) but an interaction of *Gender * Gender after Therapy* ($F_{1,171294} = 5.27, p = 0.02$) (Fig. 6).

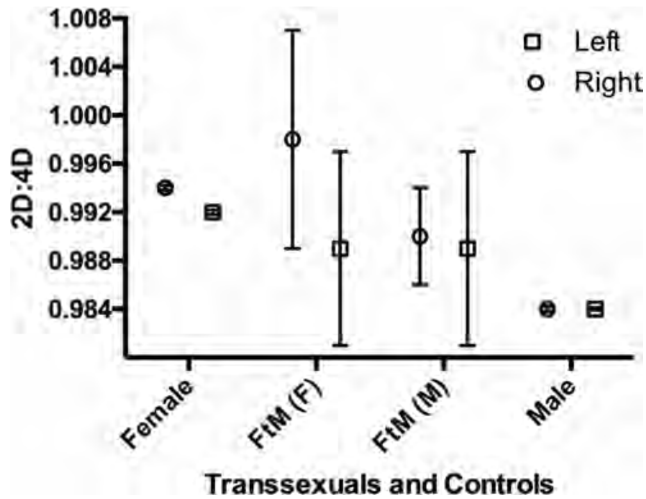

Fig. 6 Right and left mean 2D:4D (± SEM) for female-to-male transsexuals (*FtM*) reporting their gender as female (F) or male (M), (ii) natal female controls (female), and natal male controls (male)

Discussion

We investigated 2D:4D (a proxy for fetal testosterone and estrogen) in relation to natal gender, felt gender, and transsexual drug therapy in a large multi-ethnic and multinational online survey (the BBC Internet Study). Earlier reports on the BBC Study have found sex differences in 2D:4D (males < females) and ethnicity differences (Whites > East Asians = Blacks). This pattern of sex and ethnic variation has been interpreted as indicating levels of prenatal testosterone relative to estrogen. That is, males have experienced higher testosterone relative to estrogen than females, and East Asians and Blacks have higher testosterone relative to estrogen than Whites (Manning et al. 2007).

With regard to transgendered belief, we found for 2D:4D *M→M < M→F*, with *d* varying from 0.08 to 0.10. Where there were differences in effect sizes between hands, right 2D:4D tended to be greater than left 2D:4D. There was no significant difference between *F→F* and *F→M*. Among the most numerous ethnic group (Whites, $n = 169{,}630$ (93,044 males)), there was a similar pattern to the total sample, i.e., a significant difference was detected with *M→M < M→F*, $d = 0.10$, but there was no difference between *F→F* and *F→M*. Restricting the sample to White participants from the UK and the USA (the two largest national samples), the mean 2D:4D again showed a pattern of *M→M < M→F* but *F→F = F→M*.

The pattern present in transgendered belief also applied to those participants who reported taking transsexual drug therapy. Comparing *MtF* ($n = 839$) and *FtM* ($n = 223$) participants in the total sample with controls, respectively, gave a pattern of mean 2D:4Ds of *MtF* > male controls and *FtM* = female controls. The effect sizes for *MtF* > male controls were $d = 0.10$ (right 2D:4D) and $d = 0.08$ (left 2D:4D). Including (self-reported) gender or felt gender for *MtF* and *FtM* individuals in the analysis made little difference to these effect sizes. Considering White participants only reduced sample sizes (*MtF* $n = 702$ and *FtM* $n = 144$). However, the relative magnitude of 2D:4D remained similar to the findings for the total multi-ethnic sample, that is, *MtF* > male controls with $d = .08$

for both right and left hands. There was no significant effect for *FtM* (i.e., *FtM* = female controls). Presumably, the conviction that one is the opposite gender to one's natal gender is strong for individuals who elect to take transsexual drug therapy. For those that express gender dissatisfaction, it is—on average—likely to be less strong. Nevertheless, our results suggest that the patterns of 2D:4D in the latter map on to those of the former.

To date, there have been two meta-analyses of transgender studies (Sadr et al. 2020; Voracek et al. 2018). The former concluded that *MtF* transsexuals showed feminized 2D:4D ($d \approx 0.24$ for right and left 2D:4D). Conversely, *FtM* transsexuals showed masculinized 2D:4D ($d \approx 0.28$), but there was much variation around this mean, which makes the effect less meaningful. Voracek et al. (2018) reported a similar effect size metric (Hedges' g) to that of Cohen's d. In contrast to Sadr et al. (2020), these authors found only one significant effect: *MtF* transsexuals showed feminized right 2D:4D compared with controls ($g = 0.19$). Left 2D:4D was feminized over controls, but the effect was statistically not significant ($g = 0.13$). Moreover, *FtM* showed very small ("masculinized") effects for right ($g = 0.09$) and left ($g = 0.06$) 2D:4D. The authors concluded, "such effects are small at best".

Sadr et al. (2020) considered six studies (Hisasue et al. 2012; Kraemer et al. 2009; Leinung and Wu 2017; Mas et al. 2009; Schneider et al. 2006; Wallien et al. 2008) plus their own. In all seven studies, 2D:4D was experimenter-measured, and SDs are reported in five. There was evidence of low mean *SD*s of 2D:4D of about 0.03 in the Sadr et al. (2020) sample; these could reflect the high reliability of experimenter-measurement (Hisasue et al., right 0.031, left 0.033; Kraemer et al., right 0.030, left 0.029; Leinung & Wu right 0.030; Wallien et al., right 0.039, left 0.035; Sadr et al., right 0.031, left 0.032). Voracek et al. (2018) considered three additional samples: Veale (2011, p. 187–188, two samples) and Vujović et al. (2014). There were two reported self-measured *SD*s which were high, one with a physical ruler $SD = 0.048$ and one with an online ruler $SD = 0.056$ (Veale 2011). The third study (Vujović et al. 2014) measured 2D:4D from radiographs, but the protocol was unclear. Voracek et al. (2018) in their meta-analysis assessed means and *SD*s from the published error bar chart of Vujović et al. study but did not provide these values. The text and figure in Vujović et al. (2014) were contradictory, and upon inquiry by Voracek et al. (2018), no clarification was received. Nevertheless, the study was included in the meta-analysis. We reassessed the means and *SD*s from the error bars of the figure in Vujović et al.'s report and found large values for the *SD*s, right hand 0.060 and left hand 0.083; the latter was considerably influenced by the read off of the very high mean and SD of the *FtM* group ($M = 1.022 \pm 0.181$). Thus, there is evidence that the Voracek et al. (2018) meta-analysis includes three samples and six effect sizes that may show high rates of random measurement error of 2D:4D. Such a pattern of inclusion could result in a reduction of mean effect sizes and/or misleading influences on the direction of effect sizes. There was no effort to point this out in the analysis nor any attempt to adjust this effect. We suggest that experimenter-measured and self-measured reports should not be mixed in the same meta-analysis. Such mixing is likely to reduce and/or directionally distort overall effect sizes. Moreover, studies that are contradictory in their report should not be included without satisfactory clarification. In this instance, without such clarification, Voracek et al. (2018) included two very large and theory-inconsistent *FtM* effect sizes (*FtM* > *Controls*: right $g = .664$; left $g = .810$) in their analysis with a questionable influence on the overall mean direction of the effect size. Their findings that "associations of digit ratio and transgender identity are small at best" should be treated with caution.

The findings of this present study suggest that mean 2D:4D of *M*→*F* and *MtF* is higher than 2D:4D of *M*→*M* and natal males, respectively. Thus, *M*→*F* and *MtF* individuals may have experienced lower prenatal testosterone and higher estrogen than *M*→*M* and natal males. We found this pattern in a large multi-national sample, in Whites only and White participants in the UK and the USA. Digit length was self-measured and 2D:4D means showed high SDs of about 0.05. Therefore, imprecise digit measurement probably attenuated the effect sizes (Sadr et al. 2020 in their meta-analysis reported effect sizes of $d \approx 0.28$ for right and left 2D:4D). However, the sample sizes are such that we feel this pattern is representative of real differences in 2D:4D between *M*→*F* and *MtF* compared with controls (*M*→*M* and natal males, respectively). For *F*→*M* and *FtM*, we found little support that their mean 2D:4Ds differed significantly from those of *F*→*F* or *FtM*. This null finding was also in accord with the report of Sadr et al. (2020). Future work should focus on experimenter-measured 2D:4D as this will establish accurate effect sizes related to transgendered belief and transsexual drug therapy. All but one of the studies considered by Sadr et al. (2020) measured digit length indirectly from photocopies. In comparison with direct digit measurement, this method results in low mean 2D:4D and an increased sex difference (Fink and Manning 2018). A mix of effect sizes from direct and indirect experimenter-measured studies would give a further assurance of the real magnitude of 2D:4D differences in transgender studies (Manning 2017).

References

Fink, B., & Manning, J. T. (2018). Direct versus indirect measurement of digit ratio: new data from Austria and a critical consideration of clarity of report in 2D:4D studies. *Early Human Development, 127*, 28–32. https://doi.org/10.1016/j.earlhumdev.2018.09.007.

Galis, F., Ten Broek, C. M., Van Dongen, S., & Wijnaendts, L. C. (2010). Sexual dimorphism in the prenatal digit ratio (2D:4D). *Archives of*

Sexual Behavior, 39(1), 57–62. https://doi.org/10.1007/s10508-009-9485-7.

Hisasue, S. I., Sasaki, S., Tsukamoto, T., & Horie, S. (2012). The relationship between second-to-fourth digit ratio and female gender identity. *Journal of Sexual Medicine, 9*(11), 2903–2910. https://doi.org/10.1111/j.1743-6109.2012.02815.x.

Kraemer, B., Noll, T., Delsignore, A., Milos, G., Schnyder, U., & Hepp, U. (2009). Finger length ratio (2D:4D) in adults with gender identity disorder. *Archives of Sexual Behavior, 38*(3), 359–363. https://doi.org/10.1007/s10508-007-9262-4.

Králík, M., Ingrová, P., Kozieł, S., Hupková, A., & Klíma, O. (2017). Overall trends vs. individual trajectories in the second-to-fourth digit (2D:4D) and metacarpal (2M:4M) ratios during puberty and adolescence. *American Journal of Physical Anthropology, 162*(4), 641–656. https://doi.org/10.1002/ajpa.23153.

Leinung, M., & Wu, C. (2017). The biological basis of transgender identity: 2D:4D finger length ratios implicate a role for prenatal androgen activity. *Endocrine Practice, 23*(6), 669–671. https://doi.org/10.4158/EP161528.

Malas, M. A., Dogan, S., Evcil, E. H., & Desdicioglu, K. (2006). Fetal development of the hand, digits and digit ratio (2D:4D). *Early Human Development, 82*(7), 469–475. https://doi.org/10.1016/j.earlhumdev.2005.12.002.

Manning, J. T. (2002). *Digit ratio: A pointer to fertility, behavior and health*. New Brunswick, NJ: Rutgers University Press.

Manning, J. T. (2017). Prenatal sex steroids and transgender identity: Is there a link with digit ratio? *Endocrine Practice, 23*(6), 738–740. https://doi.org/10.4158/EP171843.CO.

Manning, J. T., & Fink, B. (2018). Digit ratio (2D:4D) and personality and individual differences. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *The SAGE handbook of personality and individual differences* (pp. 40–50). Thousand Oaks: SAGE Publications.

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. https://doi.org/10.1093/humrep/13.11.3000.

Manning, J. T., Churchill, A. J., & Peters, M. (2007). The effects of sex, ethnicity, and sexual orientation on self-measured digit ratio (2D:4D). *Archives of Sexual Behavior, 36*(2), 223–233. https://doi.org/10.1007/s10508-007-9171-6.

Manning, J. T., Trivers, R., & Fink, B. (2017). Is digit ratio (2D:4D) related to masculinity and femininity? Evidence from the BBC internet study. *Evolutionary Psychological Science, 3*, 316–324. https://doi.org/10.1007/s40806-017-0098-4.

Mas, M., Alonso, C., Hernandez, P., Fernandez, M., Gutierrez, P., Salido, E., & Baez, D. (2009). Androgen receptor CAG and GGN polymorphisms and 2D:4D finger ratio in male to female transsexuals. *Journal of Sexual Medicine, 6*(Suppl 5), 419–420. https://doi.org/10.1111/j.1743-6109.2009.016081.x.

Polderman, T. J., Kreukels, B. P. C., Irwig, M. S., Beach, L., Chan, Y.-M., Derks, E. M., et al. (2018). The biological contributions to gender identity and gender diversity: Bringing data to the table. *Behavior Genetics, 48*(2), 95–108. https://doi.org/10.1007/s10519-018-9889-z.

Pozzebon, J. A., Visser, B. A., & Bogaert, A. F. (2015). Vocational interests, personality, and sexual fantasies as indicators of a general masculinity/femininity factor. *Personality and Individual Differences, 86*, 291–296.

Reimers, S. (2007). The BBC internet study: General methodology. *Archives of Sexual Behavior, 36*, 147–161. https://doi.org/10.1007/s10508-006-9143-2.

Sadr, M., Khorashad, B. S., Talaei, A., Fazeli, N., & Honekopp, J. (2020). 2D:4D suggests a role of prenatal testosterone in gender dysphoria. *Archives of Sexual Behavior, 49*, 421–432. https://doi.org/10.1007/s10508-020-01630-0.

Salmon, C. A., & Hehman, C. A. (2018). Second to fourth digit ratio (2D:4D), tomboyism, and temperament. *Personality and Individual Differences, 123*, 131–134.

Schneider, H. J., Pickel, J., & Stalla, G. K. (2006). Typical female 2nd–4th finger length (2D:4D) ratios in male-to-female transsexuals. Possible implications for prenatal androgen exposure. *Psychoneuroendocrinology, 31*(2), 265–269. https://doi.org/10.1016/j.psyneuen.2005.07.005.

Veale, J. F. (2011). *Biological and psychosocial correlates of gender-variant and gender-typical identities* [dissertation]. Albany (New Zealand): Massey University.

Veale, J. F., Clarke, D. E., & Lomax, T. C. (2010). Biological and psychosocial correlates of adult gender-variant identities: new findings. *Personality and Individual Differences, 49*(3), 252–257. https://doi.org/10.1016/j.paid.2010.03.045.

Voracek, M., Kaden, A., Kossmeier, M., Pietschnig, J., & Tran, U. S. (2018). Meta-analysis shows associations of digit ratio (2D:4D) and transgender identity are small at best. *Endocrine Practice, 24*(4), 386–390. https://doi.org/10.4158/EP-2017-0024.

Vujović, S., Popović, S., Mrvošević Marojević, L., Ivović, M., Tančić-Gajić, M., Stojanović, M., et al. (2014). Finger length ratios in Serbian transsexuals. *The Scientific World Journal, 763563*, 1–4. https://doi.org/10.1155/2014/763563.

Wallien, M. S., Zucker, K. J., Steensma, T. D., & Cohen-Kettenis, P. T. (2008). 2D:4D finger-length ratios in children and adults with gender identity disorder. *Hormones and Behavior, 54*(3), 450–454. https://doi.org/10.1016/j.yhbeh.2008.05.002.

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