Arch Sex Behav
DOI 10.1007/s10508-007-9202-3

ORIGINAL PAPER

A Comparison of Finger 2D:4D by Self-Report Direct Measurement and Experimenter Measurement from Photocopy: Methodological Issues

Noreen Caswell · John T. Manning

Received: 20 June 2006 / Revised: 12 December 2006 / Accepted: 4 February 2007

Abstract The ratio of 2nd and 4th digit length (2D:4D) is sexually dimorphic and may be a correlate of prenatal sex steroids. 2D:4D is often calculated from measurements of photocopies of fingers. However, 2D:4D from photocopies is lower than 2D:4D from direct measurements of the fingers. A new and promising source of 2D:4D measures is self-reports from Internet studies. This necessitates self-report of direct finger measurements and such measurements may be unreliable. In the present study, we compared 2D:4D from self-reported finger lengths measured directly from the fingers (S-R 2D:4D) and experimenter-measured finger lengths from photocopies of the fingers (photo 2D:4D). There were 329 participants (77 men, 252 women) recruited from a first-year undergraduate psychology pool. Compared to photo 2D:4D, (1) S-R 2D:4D tended to include some extreme values; (2) S-R 2D:4D was higher; (3) S-R 2D:4D showed weak similarities which increased when extreme values of S-R 2D:4D were removed; (4) photo 2D:4D and S-R 2D:4D showed lower values for males compared to females but the dimorphism was significant for the former but not for the latter. We conclude that, insofar as S-R 2D:4D has similarities to 2D:4D from Internet studies, the 2D:4D from Internet studies will show extreme values which should be removed, mean 2D:4D will be higher than from photocopy studies, and the sexual dimorphism will be weaker than in photo 2D:4D. We suggest that large samples are necessary in Internet studies of 2D:4D because measurement error will reduce effect sizes.

N. Caswell · J. T. Manning (✉)
Department of Psychology, University of Central Lancashire, Preston PR1 2HE, England
e-mail: jtmanning@uclan.ac.uk

Keywords Digit ratio · 2D:4D · Measurement

Introduction

It has been suggested that the ratio of the lengths of the 2nd (“index”) finger and 4th (“ring”) finger (2D:4D) is a negative correlate of prenatal testosterone and a positive correlate of prenatal estrogen (Manning, 2002; Manning, Scutt, Wilson, & Lewis-Jones, 1998). This possibility has led to much interest in 2D:4D as a correlate of several sex-linked traits thought to be influenced by levels of prenatal sex steroids such as sexual orientation (e.g., McFadden et al., 2005) and transsexualism (Schneider, Pickel, & Stalla, 2006).

2D:4D is sexually dimorphic such that males have a lower mean 2D:4D than females (Manning, 2002). There is evidence from studies of congenital adrenal hyperplasia and fetal sex hormones from amniocentesis that the sex difference in 2D:4D is determined by prenatal sex steroids. However, the sexual dimorphism may be influenced by genes, including Homeobox genes, or may be a product of the interaction of fetal sex hormones and genetic factors (for discussion, see Manning, 2002; Manning, Churchill, & Peters, 2007). The dimorphism appears early and is robust to different measurement protocols: (1) directly from the fingers of fetuses as young as 9 weeks (Malas, Dogan, Hilal, Evcil, & Desdicioglu, 2006) and in children as young as 2 years (Manning et al., 1998); (2) from photocopies of the fingers in children aged from 5 to 17 years (Manning, Stewart, Bundred, & Trivers, 2004; Trivers, Manning, & Jacobson, 2006) and (3) from X-rays of the fingers of children from 1 month to 18 years (McIntyre, Ellison, Liebermann, Demerath, & Towne, 2005).

There are, however, measurement concerns with 2D:4D which center on differences in 2D:4D obtained from direct finger measurements versus measurements from photocopied fingers. Following Manning et al.'s (1998) suggestion of a link between 2D:4D and prenatal sex steroids, most of the early studies employed direct measurements of the fingers (e.g., Manning et al., 1998; Manning & Taylor, 2001). Such direct measurements are difficult to make because of participant movement and this is particularly so in studies of children. Measurements from photocopies of fingers are much easier to make and a permanent record may be retained. Consequently, many recent studies have reported 2D:4D calculated from measurements of photocopied fingers (e.g., McFadden et al., 2005; Schneider et al., 2006). However, mean 2D:4D from direct finger measurement tends to be higher than mean 2D:4D measured from photocopies, and the magnitude of this difference is influenced by sex, sexual orientation, and hand (Manning, Fink, Neave, & Caswell, 2005).

Another source of 2D:4D data is from Internet surveys, such as the recent British Broadcasting Corporation (BBC) Sex Differences Survey (Reimers, 2007). Internet surveys have potential for gathering data from very large numbers of participants, but there are concerns regarding non-serious report of data and multiple responses from individuals. A study by Gosling, Vasire, Srivastava, and John (2004) compared the findings from an Internet sample of 361,703 participants with data from 510 papers from conventional samples. Gosling et al. concluded that the findings from Internet studies were not substantially affected by repeat responses or non-serious participation, and were consistent with reports from experimenter-measured samples. Accurate recall and reliable reporting of information by participants may also be a problem in large surveys. However, there is evidence that recall of physical variables, such as birth weight of children reported by their mothers, is highly reliable in large samples with correlations between actual and recalled birth weight as high as $r = .90$ (Blanchard & Ellis, 2001; Blanchard et al., 2002). With regard to self-report of participant-measured traits, the recent BBC Internet survey recruited > 250,000 participants who reported their finger length (Manning et al., 2007). Nothing is known about the characteristics of self-reported directly measured finger lengths and the 2D:4Ds calculated from them. It seems likely that self-measurement will introduce a considerable element of random error into the measurement of finger lengths. However, direct measurement of one's own fingers may not be subject to directional distortion of 2D:4D as has been found in 2D:4D from photocopies.

Therefore, the purpose of this study was to compare self-reported finger lengths and the corresponding 2D:4D obtained from direct measurements of the fingers with finger lengths and the corresponding 2D:4D obtained from photocopies of the fingers.

Method

Participants

We recruited from undergraduate psychology students in the University of Central Lancashire. We do not think that this is a potential bias of our study because it is likely that Internet Surveys will recruit many students as most have skills in IT. For example, in the BBC Survey, there were 250,923 participants who gave their occupation and this included 81,985 (33%) students. There were a number of 2D:4D projects in the department and some second and third year students were familiar with aspects of 2D:4D. Therefore, we chose to recruit from the first year intake of session 2005/2006. Measurements were made early in the academic year in order that participants were relatively naive to the measurement characteristics of 2D:4D. The sample consisted of 329 participants (77 men and 252 women).

Procedure

All participants were provided with a photocopy of the ventral surface of a hand and a 15 cm clear plastic ruler. The 2nd and 4th fingers were indicated on the photocopy and landmarks were marked at the crease at the base of the finger proximal to the palm and the tip of the finger. The length of the 2nd and 4th digit of the left and right hand was reported by the participants and all indicated their sex. The participants were instructed to "Hold your left hand in front of you. Look at where your ring finger joins the palm of your hand. Find the bottom crease. Go to the middle of the crease. Put the 0 of your ruler on the middle of the bottom crease. Make sure the ruler runs straight up the middle of your finger. Measure to the tip of your finger (not your nail) in millimetres."

Photocopies of the ventral surface of the left and right hands were made. The participants were asked to straighten their fingers and very gently press their hand on the glass plate of the photocopier. The quality of the photocopies was checked and second copies were made if the proximal creases of the fingers or the tips of the fingers were indistinct.

Measures

The participants reported their finger lengths (2nd and 4th fingers for the right and left hands) measured directly from the fingers.

The length of the 2nd and 4th fingers on the left and right hands was measured from the photocopies using as landmarks the finger crease proximal to the palm and the tip of the finger. Each finger was measured independently by two observers blind to each others measurements. A clear plastic ruler was used and finger lengths were scored to the nearest millimetre.

Results

2D, 4D, and 2D:4D from photocopies

We calculated intra-class correlation coefficients (r_1) when comparing two sets of finger measurements or two sets of 2D:4D ratios. Values of r_1 were calculated from Model II single factor ANOVA tests:

$$r_1 = (\text{Groups MS - Error MS})/(\text{Groups MS + Error MS})$$

where MS = mean squares.

With regard to measurements of finger lengths from photocopies, we found high r_1 values when comparing measurements made by the first and second raters. These were $r_1 = .978$ for left 2D, $r_1 = .986$ for left 4D, $r_1 = .981$ for right 2D, and $r_1 = .985$ for right 4D. Repeated measures ANOVA tests showed that the ratio of differences in finger lengths between individuals to measurement error (F) was highly significant for all fingers (2D left, F [1, 328] = 89.60, $p < .0001$; 4D left F [1, 328] = 141.81, $p < .0001$; 2D right, F [1, 328] = 104.35, $p < .0001$; 4D right F [1, 328] = 131.92, $p < .0001$). There was no tendency for consistent differences in the length of finger measurements between rater 1 and 2. Thus, in comparison to measurements of finger lengths from rater 1, the lengths from rater 2 were very slightly longer for 2D left and 4D right and very slightly shorter for 4D left and 2D right (Table 1).

As expected (Manning, 2002) r_1 values for 2D:4D were lower than for finger lengths (left 2D:4D, $r_1 = .884$; right 2D:4D, $r_1 = .858$) but F values were highly significant (left 2D:4D, F [1, 328] = 16.29, $p < .0001$; right 2D:4D, F [1, 328] = 13.12, $p < .0001$).

Therefore, we concluded that measurements of finger length and 2D:4D were very similar for rater 1 and 2 and there were no directional biases. Thus, our measurements from photocopies reflected real differences in 2D:4D between individuals. Mean finger lengths and 2D:4D values from the two raters were calculated and used in all further analyses.

Table 1 Mean finger lengths (in mm) measured by raters 1 and 2

	2D Left		4D Left		2D Right		4D Right	
	M	SD	M	SD	M	SD	M	SD
Rater 1	69.16	4.79	72.23	5.06	70.18	4.72	71.87	5.01
Rater 2	69.33	4.79	72.09	5.09	70.10	4.77	71.88	5.00

2D:4D from direct measurement and photocopies

With regard to 2D:4D, we performed a 2 (Mode: Self-Report versus Photocopy) × 2 (Hand: Right vs. Left) × 2 (Sex: Men versus Women) mixed factorial ANOVA, with Mode and Hand as the within subjects factors, and Sex as the between subjects factor. This revealed a significant main effect of Mode (F [1, 326] = 36.29, $p < .0001$, Eta^2 = .10), Hand (F [1, 326] = 22.15, $p < .0001$, Eta^2 = .06), and Sex (F [1, 326] = 5.68, $p = .02$, Eta^2 = .02). There was also a significant Mode × Hand interaction effect (F [1, 326] = 9.63, $p < .002$, Eta^2 = .03). We were concerned primarily with differences in 2D:4D between Self-report (S-R 2D:4D) and Photocopies (Photo 2D:4D). For Mode, mean 2D:4D was greater for S-R 2D:4D than for Photo 2D:4D (S-R 2D:4D = .978, Photo 2D:4D = .967). For Hand, the right hand 2D:4D was greater than left hand 2D:4D (right hand .980, left hand .965), and for Sex male 2D:4D was lower than female (male .967, female .978). The significant Mode × Hand interaction arose because the Self-Report and Photocopies difference was greater for the left hand (S-R 2D:4D = .973, Photo 2D:4D = .957) than it was for the right hand (S-R 2D:4D = .983, Photo 2D:4D = .977).

The distribution of right hand ratios from S-R 2D:4D and from Photo 2D:4D is shown in Fig. 1A and B, respectively. An inspection of the distributions of left and right S-R 2D:4D showed outliers which were unlikely to be accurate, and these were higher SDs for S-R 2D:4D (.978 [.054]) than for Photo 2D:4D (.967 [.030]) which were significantly different for both hands (right hand $t = 3.34$, $p = .0009$; left hand $t = 4.86$, $p < .0001$). A comparison of S-R 2D:4D with Photo 2D:4D showed marked dissimilarities, with low but significant r_1 values for both left ($r_1 = .29$, F [1, 328] = 1.83, $p < .0001$) and right ($r_1 = .39$, F [1, 328] = 2.26, $p < .0001$) hands. The outliers were from four participants, and they all reported very different S-R 2D:4D compared to their Photo 2D:4D. We excluded the four outliers in our sample (1% of the total sample) by restricting left and right 2D:4D to a range of .80–1.20 inclusive. As a result of the exclusion, the means of S-R 2D:4D were little affected but as expected the SDs were reduced (left .979 [.046]; right .987 [.043]). With this restricted range of S-R 2D:4D, comparisons with Photo 2D:4D showed increased but still rather weak r_1 values (left $r_1 = .409$, F [1, 324] = 2.38, $p = .0001$; right $r_1 = .518$, F [1, 324] = 3.15, $p = .0001$).

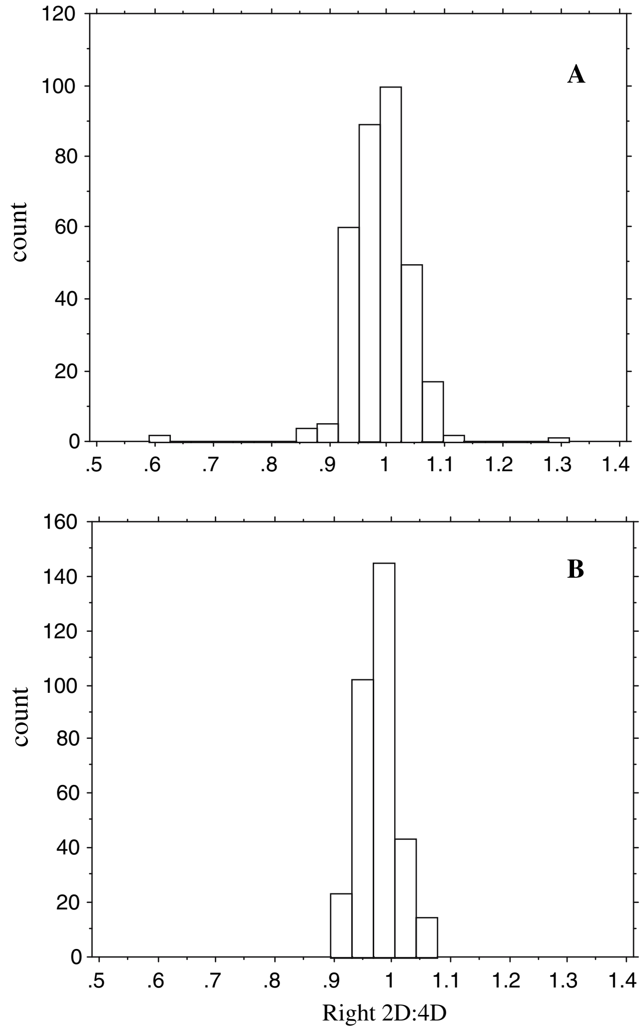

Fig. 1 Distributions of 2D:4D from self-measured finger lengths (**A**) and from rater-measured photocopied fingers (**B**)

Using the restricted range of 2D:4D, we performed a 2 (Mode: Self-Report versus Photocopy) × 2 (Hand: Right versus Left) × 2 (Sex: Men versus Women) mixed factorial ANOVA, with Mode and Hand as the within subjects factors, and Sex as the between subjects factor. As with 2D:4D that included outliers, we found a significant main effect for Mode (F [1, 322] = 66.15, $p < .0001$. Eta^2 = .17), Hand (F [1, 322] = 45.77, $p < .0001$, Eta^2 = .12) and Sex (F [1, 322] = 4.22, $p = .04$, Eta^2 = .01). There was also a significant Mode × Hand interaction (F [1, 322] = 10.03, $p < .002$, Eta^2 = .03). Mean differences between self-reported measurements and photocopy measurements of finger length and 2D:4D are shown in Table 2. It can be seen that the mean of S-R 2D:4D with outliers removed (.982) remained higher than the mean for Photo 2D:4D (.965). There was a tendency for self-reported direct finger measurements to be higher for the 2nd digit and lower for the 4th digit compared to measurements from photocopies but this was only significant for the left 2nd digit.

Table 2 Differences in mean finger lengths and mean 2d:4d between fingers self-measured directly on the hand (x) and fingers measured from photocopies (y)

Trait	x–y	SD	t	p
2nd digit left	1.138	6.396	3.21	.0001
4th digit left	–.141	6.538	.39	ns
2nd digit right	.247	6.293	.71	ns
4th digit right	–.506	6.225	1.47	ns
2D:4D left	.019	.041	8.11	.0001
2D:4D right	.011	.037	5.28	.0001

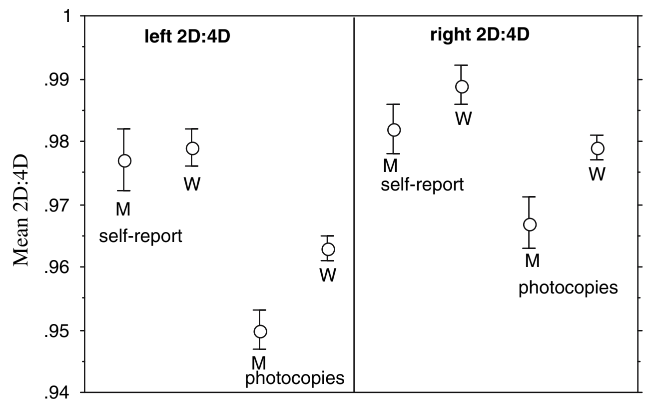

Fig. 2 Mean 2D:4D in men (M) and women (W) from self-reported finger lengths and from rater-measured finger lengths from photocopies

Mean 2D:4D by measurement procedure and by sex is shown in Fig. 2. S-R 2D:4D with outliers removed showed lower 2D:4D in males than females but this had low effect size and was not significant for the left hand (males .977 (.047), females .979 (.046), $t < 1$, *ns*, Cohen's $d = .04$) or the right hand (males .982 (.037), females .989 (.045), $t = 1.26$, *ns*, Cohen's $d = .17$). As with S-R 2D:4D, the Photo 2D:4D showed the expected sex difference but here the effect size was higher for the left hand and right hand and both were significant (left hand; males .950 (.030), females .963 (.035), $t = 2.97$, $p = .003$, Cohen's $d = .40$: right hand; males .967 (.032), females .979 (.033), $t = 2.84$, $p = .005$, Cohen's $d = .40$.

Discussion

We have found the following results: (1) for photocopies, the rater 1 and 2 measurements of finger length and photo 2D:4D were very similar and there was no evidence of directional differences in lengths of 2D and 4D; (2) S-R 2D:4D included some participants with outliers for both left and right hands which were very different from their photo 2D:4D; (3) left and right S-R 2D:4D, which included the outliers, had high SD and low r_1 values in comparisons with photo 2D:4D; (4) removal of the outliers from right and left S-R 2D:4D reduced sample size by a very small amount (1%) and the resulting samples had lower SD and higher r_1 values in comparison with photo 2D:4D; (5) S-R 2D:4D, with or without outliers, had higher means for left and right hands than photo 2D:4D; and (6) there were lower mean values of 2D:4D for males compared to females in both S-R 2D:4D and photo 2D:4D but these differences showed small effect sizes and were not significant in the former but had medium effect sizes and were significant in the latter.

Photo 2D:4D is often calculated from measurements taken by two raters. Therefore, it may be argued that consistent idiosyncratic differences between the raters may lead to directional differences in finger lengths. This possible weakness in measurement rigor would not be found in self-report measurements in which there are many raters. We found no evidence for this in the present data. Raters trained in measurement protocol will use the same measurement landmarks. Thus, we found very small differences in mean finger length for rater 1 and 2 and these were not specific for 2D or 4D. For self-report the training of raters is necessarily less detailed and is likely to be less effective than training given to raters measuring photocopies. Therefore, idiosyncratic differences in measurements of finger lengths are likely to be more common in the former compared to the latter. For these reasons, measurement error in S-R 2D:4D may be predicted to be higher than in Photo 2D:4D.

With regard to S-R 2D:4D measurements, our findings suggest that they are prone to major errors and/or deliberately misleading responses which lead to extreme values of 2D:4D. However, it is encouraging that these "errors" were rare (1% of the total sample in this study), and removal of extremely low or high values of 2D:4D is probably an effective means of management. Of more concern are the high SD and low values of r_1 of the S-R 2D:4D when compared to photo 2D:4D. However, the former was reduced and the latter increased after removal of extreme values of S-R 2D:4D.

We found that mean S-R 2D:4D was higher than photo 2D:4D for both left and right hands. We made this comparison because most studies use measurements of 2D:4D obtained from photocopies. Manning et al. (2005) reported a similar difference in comparisons of 2D:4D calculated from rater-measured direct finger measurements and photocopies. Thus, this directional difference may be a general property of 2D:4D calculated from direct digit measurements compared to 2D:4D from photocopies of the fingers. However, we can not be completely sure of this because we have no direct measurements of finger length by trained raters. Further work is required to confirm that both rater- and self-report direct finger measurements yield high mean 2D:4D. The discrepancy we have found between mean S-R 2D:4D and Photo 2D:4D may arise as a result of patterns of distortions of the tips of the 2nd and 4th digits which result from transforming three-dimensional fingers into two-dimensional photocopy images (for discussion, see Manning et al., 2005).

The effect size of the sexual dimorphism in 2D:4D was weak for S-R 2D:4D and this may well have been the result of many random errors in the self-measurement of finger lengths. A consideration of sexual dimorphism of 2D:4D in the BBC internet study of 255,116 participants showed that self-reported finger lengths gave significant dimorphism for the right and left hand with very low effect sizes of $d = .09$ for the right hand and $d = .06$ for the left hand. After excluding outliers of 2D:4D, the effect sizes increased somewhat to $d = .20$ and $d = .15$ for right and left, respectively. As expected from earlier work, the effect size of the dimorphism in photo 2D:4D was approximately of medium strength.

We have found that self-measurements directly on the fingers are prone to a number of drawbacks. It is likely that they will include a few major "errors" which result in outlier values of 2D:4D. When these are excluded, there are weak but significant associations between 2D:4D calculated from self-measurements and 2D:4D from rater-reported measurements. In addition, many minor "errors" in self-measurements probably result in a reduction in the effect size of the sexual dimorphism of 2D:4D, and in effect sizes for correlations for 2D:4D with target variables. These drawbacks can be countered if large numbers of participants are recruited. We conclude that self-reported finger lengths can be valuable in the study of 2D:4D if large samples are available and outliers of 2D:4D are removed. Internet studies, such as the BBC Survey, can provide large samples and may be a valuable tool in the study of sex dependent patterns of behavior in 2D:4D.

References

Blanchard, R., & Ellis, L. (2001). Birth weight, sexual orientation and the sex of preceding siblings. *Journal of Biosocial Science, 33*, 451–467.

Blanchard, R., Zucker, K. J., Cavacas, A., Allin, S., Bradley, S. J., & Schachter, D. C. (2002). Fraternal birth order and birth weight in probably prehomosexual feminine boys. *Hormones and Behavior, 41*, 321–327.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist, 59*, 93–104.

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

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

Manning, J. T., Churchill, A. J. G., & Peters, M. (2007). The effects of sex, ethnicity, and sexual orientation on self-measured digit ratio (2D:4D). *Archives of Sexual Behavior, 36*, 223–233.

Manning, J. T., Fink, B., Neave, N., & Caswell, N. (2005). Photocopies yield lower digit ratios (2D:4D) than direct finger measurements. *Archives of Sexual Behavior, 34*, 329–333.

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*, 3000–3003.

Manning, J. T., Stewart, A., Bundred, P. E., & Trivers, R. L. (2004). Sex and ethnic differences in 2nd to 4th digit ratio of children. *Early Human Development, 80*, 161–168.

Manning, J. T., & Taylor, R. P. (2001). Second to fourth digit ratio and male ability in sport: Implications for sexual selection in humans. *Evolution and Human Behavior, 22*, 61–69.

McFadden, D., Loehlin, J. C., Breedlove, S. M., Lippa, R. A., Manning, J. T., & Rahman, Q. (2005). A reanalysis of five studies on sexual orientation and the relative length of the 2nd and 4th fingers (the 2D:4D ratio). *Archives of Sexual Behavior, 34*, 341–356.

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 Royal Society London B, 272*, 1473–1479.

Reimers, S. (2007). The BBC Internet Study: General methodology. *Archives of Sexual Behavior, 36*, 147–161.

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*, 265–269.

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*, 150–156.