Early Human Development 127 (2018) 28–32

Contents lists available at ScienceDirect

Early Human Development

journal homepage: www.elsevier.com/locate/earlhumdev

Direct versus indirect measurement of digit ratio: New data from Austria and a critical consideration of clarity of report in 2D:4D studies

Bernhard Fink[a,b,*], John T. Manning[c]

[a] Department of Behavioral Ecology, University of Goettingen, Germany
[b] Hanse-Wissenschaftskolleg, Institute for Advanced Study, Delmenhorst, Germany
[c] Applied Sports, Technology, Exercise, and Medicine (A-STEM), Swansea University, Swansea, United Kingdom

ARTICLE INFO

Keywords:
Digit ratio
2D:4D
Direct
Indirect
Measurement
Austria

ABSTRACT

Background: Digit ratio (2D:4D), has been the subject of many studies. However, the best method of measuring digit length, either directly (d2D:4D) or indirectly (i2D:4D), is controversial. In many reports i2D:4D has been found to show a directional effect such that d2D:4D > i2D:4D. The exception is three studies from one group of researchers in Austria. Thus, it is unclear whether this effect is nation- or lab-specific.
Aims: To examine evidence for effects of direct versus indirect measurements of mean 2D:4D in Austrians.
Study design: We compared 2D:4D based on direct and indirect measurements of digit lengths of Austrians.
Subjects: There were 80 participants, 21 adults and 59 children.
Outcome measures: 2D:4D of right and left hands, measured directly (from the palm) and indirectly (from hand scans).
Results: Repeatability was high for both d2D:4D and i2D:4D, with the latter slightly higher than the former. d2D:4D and i2D:4D correlated strongly and the sex difference in 2D:4D (males < females) was greater for i2D:4D. With regard to directional differences, we found d2D:4D > i2D:4D for both right ($d = -0.53$) and left hands ($d = -0.80$).
Conclusion: There was no evidence of an "Austrian" effect on direct versus indirect measurements of 2D:4D, i.e. mean d2D:4D was greater than i2D:4D. We discuss our findings in the light of issues regarding "clarity of report" from earlier Austrian studies.

1. Introduction

Digit ratio (2D:4D) - the relative lengths of the index (2D) and ring (4D) finger - is thought to be a negative correlate of prenatal testosterone and a positive correlate of prenatal oestrogen [1,2]. Studies have shown that 2D:4D is a correlate of many sex-dependent traits with effect sizes varying from strong (performance in sport) to weak (personality traits; see for review [4]). However, determining accurate effect sizes in 2D:4D has proved difficult because there is no consensus as to how to measure digit length.

In order to measure digit length, it is desirable to measure the overall length of the bones or phalanges. However, in practice soft tissue measurements are usually made. Manning et al. [3] reported digit lengths measured directly (with a Vernier calliper) from the ventral surface of the fingers. The measurement points for this method are the fleshy tip of the digit to the crease proximal to the palm of the hand. The latter measurement point lies approximately at the mid-point of the proximal phalanx. This direct method of measuring 2D:4D (d2D:4D) has its difficulties. It requires that the participant's digits are measured (and re-measured for the assessment of measurement accuracy) during data collection with the fingers kept steady and straight in all experimental conditions. Therefore, images of digit lengths from photocopies of the hands were measured and indirect 2D:4D's (i2D:4D) reported [5]. Later, Robinson and Manning [6] compared d2D:4D and i2D:4D and reported very strong intra-class correlations (ICC's) for d2D:4D versus i2D:4D. However, ICC's measure correlations between d2D:4D and i2D:4D. They do not indicate differences in scale (i.e. whether d2D:4D and i2D:4D differ in magnitude).

Subsequent to these studies, researchers collected and reported i2D:4D (from photocopies or hand scans, many of which were measured with Vernier callipers) in many studies. However, concerns regarding issues of scale in image-based 2D:4D's emerged when Manning et al. [7] reported that d2D:4D > i2D:4D in samples of heterosexual men and women and homosexual men. This was a matter of substantial concern.

* Corresponding author at: Department of Behavioral Ecology, University of Goettingen, Kellnerweg 6, 37077 Goettingen, Germany.
E-mail address: bfink@gwdg.de (B. Fink).

https://doi.org/10.1016/j.earlhumdev.2018.09.007
Received 10 July 2018; Received in revised form 31 August 2018; Accepted 11 September 2018

The measurement points for d2D:4D and i2D:4D are identical, but the mean 2D:4D's are significantly different. Thus, many studies conducted between 1998 and 2005 reported values of i2D:4D, measured with a protocol, which was poorly understood. We emphasize that the comparison of mean 2D:4D's is very important in digit ratio literature. Many papers test the difference in mean 2D:4D between the sexes, ethnic groups, nations, athletes and non-athletes, patients and controls etc. If mean d2D:4D and i2D:4D show directional differences, it is essential that in such comparisons the protocol of digit measurement (direct or indirect) is the same. If it is not and the protocol influences the group mean such comparisons are meaningless. This also applies to pooling data. A mix of participants measured directly and indirectly will give misleading mean values of 2D:4D.

Concern regarding the measurement of finger length increased further when one group of Austrian researchers [8–10] published three papers that reported non-replication of the Manning et al. [7] directional effect of i2D:4D. A surprising pattern of replication studies then emerged. The directional effect of i2D:4D was reported in eight studies from six nations (England, Canada, Saudi Arabia, Spain, Korea, USA), either with experimenter-reported measurements [11–16] or self-reported measurements of digit lengths [17,18]. Furthermore, a meta-analysis of Chinese 2D:4D studies showed a significant directional effect (d2D:4D > i2D:4D; [19]). Thus, we have here an example of the "replication crisis", which has troubled science over the last few years [20]. Ribeiro et al. [16] have considered this singular pattern of "Austrian" non-replication and have cautioned against assuming that it is necessarily lab-specific. There are significant ethnic and national effects on 2D:4D. It may be that there is some aspect of Austrian 2D:4D, which lends itself to distortional effects of i2D:4D such that (in comparison to d2D:4D) it shows a mix of positive and negative effects. However, in contrast to the nation-specific explanation, there is some evidence of a lack of "clarity in report" in the studies by Voracek and colleagues [16]. That is, the findings of the authors do not map on to their reports such that the conclusions of the reports do not appear to accurately represent patterns of data within the report. One study [8] reported d2D:4D < i2D:4D, but this effect ($d = -0.31$, a negative value of d indicates d2D:4D < d2D:4D) may have arisen from power issues (a sample size of $N = 30$). The other studies found significant effects that supported a relationship of d2D:4D > i2D:4D in six out of the possible eight comparisons (d varied from 0.12 to 0.62), and non-significance ($d = -0.03$ and -0.10) for the remaining two [9,10]. This support for the relationship of d2D:4D > i2D:4D was not clearly reported (see for discussion [16]).

Therefore, the purpose of the present study was an attempt to search for nation-specific effects in a comparison of direct versus indirect 2D:4D among Austrians. If the reports by Voracek and colleagues [8–10] are correct we should find that Austrian 2D:4D shows a mix of effects such that d2D:4D < i2D:4D, d2D:4D = i2D:4D, d2D:4D > i2D:4D. In contrast, if reports from all other nations are correct we expect that in an Austrian sample we should find d2D:4D > i2D:4D across all comparisons. We emphasize that we do not consider whether d2D:4D is "better" than i2D:4D. Nor do we focus on evidence for why there are differences in mean d2D:4D and i2D:4D. Rather, we are concerned with examining whether differences in mean 2D:4D such that d2D:4D > i2D:4D are found in Austrian samples in addition to all other national reports.

2. Methods

In order to exclude the possibility of age effects in direct versus indirect comparisons of 2D:4D, our sample included children and adults. We recruited participants from the population of a rural village in Lower Austria (Austria). All participants were Caucasian. Adult participants ($N = 21$) were members of the local administration and schoolteachers. Children ($N = 59$) were recruited at the elementary school, following verbal consent and informed written consent by their parents. To preserve anonymity, each participant received a consecutive number at arrival and we used only this number to assign his or her measurements.

We measured the index and the ringer finger of both right and left hands, twice from each participant. One of the investigators (B.F.) took the measurements, directly from the palm, from the proximal crease to the tip of the finger, with a digital Vernier calliper (Preisser products, Germany) measuring down to 0.01 mm. First and second measurements of right and left hands were alternated. The investigator called out each measurement to an assistant sitting next to the investigator, who wrote down numbers to ensure that the investigator could not memorize them.

Following direct measurements, hand scans of the right and left hand of each participant were collected with a flatbed scanner (Canon LiDE 210, Canon Inc., Japan), in colour, and at a resolution of 300 dpi. For identification purposes, we attached a sticky note to the palm with the participants' number and information on right vs. left hand on it. These image files were transferred electronically to the second investigator (J.T.M), who was then blind to the results of direct measurements. The investigator printed out the images of hands on A4 paper, and measured finger lengths with a Vernier calliper, following the procedure of direct measurements. In order that the first and second digit measurements were independent, the latter was made blind to the former and the second measurement was made at least 24 h after the first.

The repeatabilities or intra-class correlation coefficients of 2D:4D's were calculated from repeated measures ANOVAs. Product-moment correlations were computed for the associations between d2D:4D and i2D:4D, and sex differences in mean 2D:4D were tested with t-tests. Paired t-tests were used to examine the mean differences between d2D:4D and i2D:4D. The influence of sex and age on the directional differences in d2D:4D and i2D:4D was tested using two-factor ANOVAs. Cohen's "rule of thumb" for the magnitude of effect sizes was used. For correlation coefficients this was $r = 0.1$ small, $r = 0.3$ medium and $r = 0.5$ large. For differences in means it was $d = 0.2$ small, $d = 0.5$ medium and $d = 0.8$ large.

The ethical committee at the Department of Psychology at the University of Goettingen (Germany), protocol #192, approved the protocol.

3. Results

3.1. Descriptive statistics

There were 80 participants, 21 adults and 59 children. The mean age (± SD) of the former was 49.2 ± 10.6 years and of the latter 7.9 ± 1.3 years. With regard to the adult sample, there were five men (mean age 48.2 ± 6.7) and 16 women (mean age 49.5 ± 11.8). The children's sample comprised 24 boys (mean age 8.0 ± 1.4) and 35 girls (mean age 7.9 ± 1.2).

3.2. Repeatability

With regard to direct measurements, the intra-class correlation coefficient or ICC (r_1) of the first and second measurements of 2D:4D was high for both right ($r_1 = 0.84$, $F[1,79]$, $p < .0001$) and left hands ($r_1 = 0.84$, $F[1,79]$, $p < .0001$). For indirect measurements, the ICC's were higher (right: $r_1 = 0.94$, $F[1,79] = 33.41$, $p < .0001$; left: $r_1 = 0.96$, $F[1,79] = 51.52$, $p < .0001$) and the CI's showed the difference was significant (95%CI for direct right and left = 0.76–0.89; 95%CI for indirect right and left = 0.91–0.96 and 0.94–0.97, respectively). Thus, for both d2D:4D and i2D:4D, the between-differences in 2D:4D were greater than the differences caused by measurement error. Therefore, we calculated the mean d2D:4D and the mean i2D:4D and used these in all subsequent comparisons.