Digit Ratio (2D:4D) in Lithuania Once and Now: Testing for Sex Differences, Relations with Eye and Hair Color, and a Possible Secular Change

Martin Voracek1, Albinas Bagdonas2 and Stefan G. Dressler1

1 Department of Basic Psychological Research, School of Psychology, University of Vienna, Vienna, Austria
2 Laboratory of Special Psychology, University of Vilnius, Vilnius, Lithuania

ABSTRACT

The second-to-fourth digit ratio (2D:4D) is a sexually dimorphic somatic trait and has been proposed as a biomarker for the organizational, i.e., permanent, effects of prenatal testosterone on the human brain. Accordingly, recent research has related 2D:4D to a variety of sex-dependent, hormonally influenced traits and phenotypes. The geographical variation in typical 2D:4D is marked and presently poorly understood. This study presents the first investigation into the 2D:4D ratio in a Baltic country. A contemporary sample of 109 Lithuanian men and women was compared with data from a historical sample of 100 Lithuanian men and women, collected and published in the 1880s and rediscovered only now. The findings included the following lines of evidence: (i) seen in an international perspective, the average 2D:4D in Lithuania is low; (ii) there was a sex difference in 2D:4D in the expected direction in both samples; (iii) a previously advanced hypothesis of an association of lighter eye and hair color with higher, i.e., more feminized, 2D:4D received no support in both samples; and (iv) the average 2D:4D in the contemporary sample was higher than in the historical sample. In view of a hypothesized increase in 2D:4D in modern populations, owing to increased environmental levels of endocrine disruptors such as xenoestrogens, this latter finding appears to be of particular notice. However, because finger-length measurement methods differed across the samples, it cannot be safely ruled out that the apparent time trend in Lithuanian 2D:4D in truth is an artifact. The puzzling geographical pattern seen in the 2D:4D ratio and the question of possible time trends therein deserve further investigations.

Key words: 2D:4D, digit ratio, sex differences, eye color, hair color, time trends

Introduction

The length ratio of the second digit (2D; index finger) to the fourth digit (4D; ring finger), hereafter digit ratio of the right (R2D:4D) or left hand (L2D:4D), is a sexually dimorphic somatic trait in humans and in a range of other primates as well as rodent and avian species (for a historical review of human digit ratio research, see Peters, Mackenzie, and Bryden1; a brief overview of cross-species digit ratio research is given in Voracek2). In humans, men on average present lower 2D:4D than women, i.e., have a longer 4D relative to their 2D. The 2D:4D ratio has been recently proposed as a biomarker for prenatal androgen exposure and sensitivity and its associated organizational (permanent) effects on the brain and behavior3. Presumably, prenatal testosterone promotes the growth of 4D and prenatal estrogen the growth of 2D.

Over the past few years, this hypothesis has generated much interdisciplinary research interest (see Schaefer, Fink, Mitteroecker, Neave, and Bookstein4 for a recent contribution to this journal) and the available evidence consistent with this hypothesis is now fairly strong. 2D:4D has been shown to be a correlate of a multitude of sex-dependent, hormonally influenced variables and phenotypes, including behavioral, cognitive, personality, and somatic traits, adult-onset disorders, and measures related to fertility and sexuality (for a review, see the monograph of Manning5).

Apart from the ubiquitously observed sex differences, geographical (population or ethnic) variation in typical 2D:4D measurements have been noted across several
multi-national investigations. Rather surprisingly, 2D:4D appears to vary somewhat erratically across nations and these cross-national differences in 2D:4D are noticeably larger than the sex differences found within nations. Presently, the international pattern seen in 2D:4D is still poorly understood and therefore subject of considerable research interest.

Up to now, the literature on human 2D:4D, comprised of more than 80 reports since the first account of Manning, contains samples from more than 25 different countries or ethnic groups from five continents. A full review of this growing body of evidence is outside the scope of the present contribution and may even be premature, but a brief sketch of the extant evidence regarding adjacent Northern European areas will illustrate some of the peculiarities pertaining to the population differences in typical 2D:4D measurements. In international perspective, digit ratio is rather high in the United Kingdom, amounting to 0.97 or 0.98 in most studies (all values listed here are for men). It is low in Sweden (0.95), very low in Finland (0.93), but exceptionally high in Denmark (1.02), and again rather high in Poland (0.99), much higher than in Germany (about 0.96). Since sample SDs for 2D:4D mostly range in the 0.030s, this constitutes quite marked group differences.

Digit ratio data from Norway, the Baltic countries of Estonia, Latvia, and Lithuania, and from Russia or Belarus have not yet been published. However, the serendipitous rediscovery of a forgotten work of late 19th-century anthropometry (Brennsohn, 1883) by one of us (M.V.) enabled us to add the first data from a Baltic population (Lithuanians) to the existing literature on population differences in 2D:4D. The account of Brennsohn, unearthed here, is a German-language, published medical dissertation of 1883 from the University of Dorpat (today’s Tartu, the second-largest city of Estonia). The content of his research was an extensive anthropometric investigation of a sample of ethnic Lithuanians. Luckily so, the raw data matrix of his investigation (i.e., all the variables measured across all study subjects) is given in an appendix to Brennsohn’s work, therefore making reanalysis of his data possible on the individual, i.e., non-aggregated, level. In tandem with our own collection of data from contemporary Lithuanians, this allowed us to test four hypotheses, which are outlined and summarized in the following.

The first hypothesis concerned the average 2D:4D of Lithuanians. Because of the proximity of Lithuania and Finland and because of the long-standing historical influences and traces of both Swedish and Germans in this area (apart from more recent Russian influences), the conjecture was that 2D:4D of Lithuanians would be similar to these populations (i.e., comparably low) rather than high (i.e., like in Polish). This was expected for both Brennsohn’s sample and our own data (hereafter referred to as the historical sample and the contemporary sample, respectively).

The second hypothesis tested for sex differences in 2D:4D. This is such an established finding that we expected to see the effect in both the historical and the contemporary sample alike.

The third hypothesis dealt with eye and hair color, which were also ascertained by Brennsohn. It has been hypothesized that prenatal exposure to estrogen, as indicated by 2D:4D, might be higher in individuals with light eyes and blond hair (Manning; cited in Frost, but see the earlier hypothesis of Manning, p. 37, which was in the opposite direction: according to this theorizing, blond hair may be associated with high prenatal testosterone, not with high prenatal estrogen). Currently, empirical evidence on this point is lacking, but it is of note in this context that a correspondence of lighter skin color to higher 2D:4D has been reported previously. Although eye, hair, and skin color are determined by different sets of genes, they often go together phenotypically as to suggest that the relation of lighter complexion to higher 2D:4D, as found for the skin, might also generalize to lighter pigmentation of the eyes and the hair.

The fourth hypothesis encompassed a comparison of the 2D:4D measurements in the historical and the contemporary sample. The collection of these data within the same region took place with a time interval of about 120 years or four human generations apart, thus enabling a unique opportunity to test for a secular change.

There is preliminary evidence for a moderate heritability of human 2D:4D as well as for assortative mating on this trait. Further, relations of 2D:4D with reproductive success have been found in various samples and it has been speculated that the expression of 2D:4D is dependent on sexually antagonistic genes. Taken together, over time and within populations these forces will alter the allele frequencies for the genetic base responsible for the variable expression of this trait.

Another relevant point in this context is animal-model evidence that endocrine disruptors can change the prenatal expression of 2D:4D. Xenoestrogens, i.e., organic compounds which mimic the effects of estrogen on the body, are a known class of endocrine disruptors. It has been speculated that they might be implicated in altered 2D:4D expression patterns in humans (p. 39). Found in common products, such as cosmetics, detergents, pesticides, and plastics, among others, environmental loads of xenoestrogens have markedly increased with the process of industrialization and the emergence of consumer societies, as seen over the course of the 20th century.

Although the present research could not address these points in a direct way, it appeared to us more likely than not that the average 2D:4D of populations could change over time. Specifically, modern populations could be more estrogenized than their historical counterparts. Therefore, thinking of a possible impact of xenoestrogens on contemporary populations, we tested for a secular increase in 2D:4D, i.e., whether the historical (1880s) Lithuanian sample presented lower 2D:4D than the contemporary (2000s) sample.
Materials and Methods

Participants

The historical sample comprised 100 participants (60 men and 40 women), aged 18 to 70 years (M=36.9 and SD=12.6 years), from a rural population background. The contemporary sample comprised 109 psychology students from the University of Vilnius (23 men and 86 women), aged 18 to 43 years (M=22.4 and SD=4.3 years). All participants were ethnic Lithuanians.

Measures

For the historical sample, the experimenter recorded participants’ eye and hair color. This procedure yielded 12 categories for eye color (blue, steel-blue, bluish grey, greyish blue, grey, greenish grey, greyish green, greyish brown, light brown, brown, brownish grey, and dark brown) and six categories for hair color (blond, dark blond, light brown, brown, dark brown, and black). Two men and one woman with grey hair were discarded for the present analyses involving hair color. We recoded the first seven eye color categories (blue to greyish green) as «light» and the latter five categories (greyish brown to dark brown) as «dark». For hair color, we recoded blond and dark blond as «light» and the latter four categories as «dark».

For the contemporary sample, the survey instrument used included items on participants’ self-reported eye color and their natural, adult hair color, according to slightly modified gradings, as previously used in related research16. The categories were as follows: light blue, dark blue, green (any shade of green), greenish blue or bluish green (any mixed shade of green and blue), grey, light brown, and dark brown for eye color; and fair (or light) blond, dark blond, fair (or light) red, red, fair (or light) brown, medium brown, dark brown (but not black), and black for hair color. Responses from two women were missing for the eye color variable. We recoded the first five (out of seven) eye color categories and the first two (out of eight) hair color categories as «light» and the remainder as «dark». In essence, the dichotomizations for the eye and hair color gradings applied in the two samples consistently separated people with either non-brown eyes or blond hair from the remaining, darker pigmented phenotypes.

Procedure

The data of the historical sample were gathered during summer and autumn 1882 in various villages and manors in the Kaunas district, an area located in central Lithuania. The experimenter directly measured finger length, using standard procedures and instruments of 19th-century anthropometry, with a sliding caliper to the nearest one mm. The published source is silent on which method was measured, but, considering the historically strong and widespread cultural prejudice against left-handedness17, in all likelihood it must have been the fingers of the right hand. Measurement landmarks were the metacarpophalangeal joint, identified by palpation on the dorsal side of the hand, and the finger tip.

In the contemporary sample, photocopies of the right and left palm were taken. Finger lengths were measured twice by a trained investigator (S.G.D.) with a digital vernier caliper measuring to 0.01 mm (Mitutoyo Ltd., Andover, Hampshire, U.K.; Model 500-191U). The second measurements were made after a two-week interval, blind to the first measurement series, and all finger-length measurements were made blind to the other data collected. Measurement landmarks were the ventrally located proximal-most (boundary) metacarpophalangeal flexion crease that divides the finger from the palm region and the finger tip, for the latter one excluding possibly protruding fingernails.

Data analysis

Intraobserver repeatabilities of finger-length measurements for the contemporary sample were assessed with one type of intraclass correlation coefficient (ICC), employing a two-way mixed-effects model with absolute-agreement definition (Case 3 ICC in McGraw and Wong18). The dual measurements for individual traits (2D and 4D) were averaged before calculating 2D:4D. Further, DR–L, the directional asymmetry in digit ratios (2D and 4D) were averaged before calculating 2D:4D. Measurement concordance was high and reflected real differences between individuals. Brennsohn’s data were single measurements, so repeatability figures were not available for the historical sample.

Results

Repeatabilities of finger-length measurements

Intraobserver measurement repeatabilities in the contemporary sample were as follows (all ps<0.001, df1= df2=108 for the F ratios): ICC=0.997 for right 2D (F=889.8), 0.998 for right 4D (F=1399.5), 0.997 for left 2D (F=823.4), 0.997 for left 4D (F=724.1), 0.984 for R2D:4D (F=121.0), 0.977 for L2D:4D (F=85.9), and 0.946 for D2D:4D (F=55.6). Measurement concordance was high and measurement error trivially small relative to the inter-individual differences measured. The measurements thus reflected real differences between individuals. Brennsohn’s data were single measurements, so repeatability figures were not available for the historical sample.

Sex differences in 2D:4D

Descriptive statistics and the results of unpaired t tests for sex differences in 2D:4D for both samples are displayed in Table 1. All sex effects were in the expected direction, except for that there was no sex difference on DR–L in the contemporary sample. As for the historical sample, the mean 2D:4D for men was lower than for women, albeit this group difference fell short of statistical significance. In the contemporary sample, R2D:4D and L2D:4D were comparably strongly associated for men and women, rs=0.73 and 0.65, respectively (ps<0.001).
Associations of 2D:4D with eye and hair color

For this set of analyses, digit ratio was standardized (z score method) within each sex. In the historical sample, 50 out of 60 men (83.3%) and 31 out of 40 women (77.5%) had light eyes and 37 out of 58 men (63.8%) and 19 out of 39 women (48.7%) had light hair. Light-eyed people in this sample presented somewhat higher standardized digit ratios than dark-eyed people, M (and SDs) being 0.04 (1.00) versus –0.14 (1.00), respectively, but this difference was statistically not significant (t\textsubscript{58}=0.70). Fair-haired individuals showed somewhat lower standardized digit ratios than dark-haired people, with M (and SDs) of –0.02 (1.09) versus 0.07 (0.83), but this difference also was statistically not significant (t\textsubscript{58}=–0.44).

In the contemporary sample, 22 out of 23 men (95.7%) and 65 out of 84 women (77.4%) had light eyes and five out of 23 men (21.7%) and 22 out of 86 women (25.6%) had light hair. Light-eyed individuals in this sample showed somewhat higher standardized R2D:4D than dark-eyed individuals, M (and SDs) 0.02 (1.00) versus –0.05 (1.06), whereas the results for standardized L2D:4D were almost indistinguishable, M (and SDs), in above order, –0.01 (1.00) versus ~0.02 (1.04). Both group comparisons were statistically insignificant, t\textsubscript{195}=0.30 and 0.01, respectively. Fair-haired individuals had somewhat lower standardized R2D:4D than dark-haired individuals, –0.07 (1.03) versus 0.02 (0.99), as well as lower standardized L2D:4D, –0.27 (0.99) versus 0.09 (0.99), respectively. These group differences were also statistically not reliable, t\textsubscript{197}=–0.39 and –1.61.

Sample differences in 2D:4D

To test for sample differences in digit ratio, R2D:4D and L2D:4D in the contemporary sample were averaged. A two-way factorial analysis of variance, with sample and sex as the between-subjects factors, indicated both significant sex differences (F\textsubscript{1,205}=8.57, p=0.004, partial \(\eta^2=0.04\)) and significant sample differences in digit ratio (F\textsubscript{1,205}=26.78, p<0.001, partial \(\eta^2=0.12\)), whereas the interaction between these two factors was negligible (F\textsubscript{1,205} <1, ns, partial \(\eta^2<0.01\)).

Discussion

The present study investigated patterns of 2D:4D in two cohorts of Lithuanian men and women temporally separated by more than a century. We here reported the first digit ratio data from a Baltic country and also undertook the first cross-temporal test of human digit ratio, by comparing within the same geographical locale a historical sample with a contemporary one. We tested a total of four hypotheses, the evidence from which is discussed below.

Firstly, seen in an international perspective, 2D:4D in Lithuania apparently is low. The average 2D:4D for contemporary Lithuanian women found here was not higher than typically encountered measurements for English men. Lithuanian 2D:4D measurements are more similar to the ones found for Finland, Germany, or Sweden than to those found in Poland, Lithuania’s southwesterly neighbor. This finding adds to the currently unsolved puzzle of strong geographical and population differences in 2D:4D worldwide and also within Europe. It is now increasingly evident that these differences are not explainable with simple gradient models, since geographically adjoining countries do not necessarily present comparable 2D:4D. This question definitely deserves further inquiry. Specifically, it would be interesting to see 2D:4D data from the other Baltic countries (Estonia and Latvia) as well as from Russia and Belarus.

Secondly, noticeable sex differences in both right and left 2D:4D were observed within the contemporary Lithuanian sample and also, to a lesser degree, thus failing to achieve the nominal statistical significance level, in the historical sample. We believe this nonsignificant finding should not be disregarded. The sex effect in Brennsohn’s historical sample which we reanalyzed, amounted to nearly 0.30 SD units, which is not uncommonly small: several samples with sex differences in 2D:4D ranging in the 0.20s or 0.30s in terms of SD units have been reported in the literature\(^5\) (p. 21). Moreover, the joint analysis of sex differences in 2D:4D across both samples through analysis of variance indicated a significant sex effect which was not qualified by a sex-by-sample interaction effect. Lastly, it is well known that sexually dimorphic differences tend to be minimized in stressed populations.
which may well have been the case with the historical rural population. We therefore conclude that sex differences in 2D:4D are not only cross-sectionally ubiquitous across many ethnic groups, but also generalize across generations and time.

Thirdly, there was no support for the hypothesis of a relation between non-brown eyes or blond hair and higher 2D:4D in both data sets. We suggest further investigations into this hypothesis with larger samples, since these effects might be small and therefore might not reliably show up in small data sets, such as those analyzed here. In addition to the limited sample size available for the present analysis, several further limitations of the present study apply in regards to the test of this third hypothesis.

Although the sex ratio in the contemporary sample faithfully reflected the student sex ratio at the study enrollment site, it nevertheless was too skewed to permit sex-specific analyses, owing to sparseness of men in the sample. And although we believe that our dichotomized recoding for the eye and hair color gradings consistently separated individuals with either non-brown eyes or blond hair from the remainder in both the historical and the contemporary sample, the gradings itself, as used across these samples, were not exactly identical. Much larger samples would be needed to analyze possible associations of eye and hair color on 2D:4D along a rank-ordered (instead via a dichotomized) scheme. Further, the source of the historical sample gives no clue on the full categorical scheme and the exact procedure used by the experimenter in his assessment of people’s eye and hair color. Only the results of Brennsohn’s eye and hair color evaluations are accessible and these, admittedly, sometimes give an idiosyncratic impression (e.g., what is the exact difference between blue and steel-blue eyes or how can greyish brown eyes be kept apart from brownish grey ones?).

Another shortcoming was that in the contemporary sample we relied on participants’ self-reported eye and hair color. Although the coding schemes seem perfectly clear and were adopted from previous research, a margin of error or bias introduced through the self-report mode nevertheless cannot be completely ruled out. Objective measurement of constitutional skin pigmentation has been used for validating the hypothesized association of lighter skin color with higher 2D:4D. We therefore suggest that future studies on the relations of eye and hair color with 2D:4D should aim to objectively measure these phenotypes. It would also be interesting to test specific combinations of eye and hair color.

One final point pertaining to this discussion is that fair hair frequently tends to darken during the adolescent phase and in the younger adult years and this darkening trend seems to begin already in childhood. It would therefore be advantageous to ask adult individuals not about their current, natural hair color, but rather about their hair color as a child, or alternatively, to assess it from their child color photos or to directly test this in child samples. Since there are additional sex and individual differences in the darkening trend, it would also be beneficial to conduct studies with individuals who are at least 25 years old.

And fourthly, the results from the comparison of 2D:4D in the historical versus the contemporary sample were consistent with a possible increase of 2D:4D over time in Lithuanians. Any combination of the effects of assortative mating on 2D:4D, of sexually antagonistic genes, of sex-specific relations of 2D:4D with fertility and reproductive success, or of environmental xenoestrogens, or all of them acting together to varying degrees, could in principle produce such a shift in typical 2D:4D ratios within a population. In this respect, the data base available for this cross-temporal test was not informative, since it did not permit us to disentangle the suspected causative factors.

There is yet another reason for why one should not put too much confidence on the present finding of an apparent secular increase in 2D:4D in Lithuania. This originates from the different methods of finger-length measurement used in the data sets. While for Brennsohn’s historical sample the proximal measurement landmark was the metacarpophalangeal joint (identified by palpation on the dorsal side of the hand), for the contemporary sample, according to the method now applied by the majority of 2D:4D researchers, the measurements were made from photocopies of the palm, where the proximal landmark point was the ventrally located boundary (proximal-most) metacarpophalangeal flexion crease which divides the finger base from the palm. Thus the two studies were methodologically distinct. We deemed it futile to try to mimic the method Brennsohn used, because his text does not give further methodological details beyond those presented here.

On the other hand, it should be noted that at present there is no evidence for a non-equivalence of precisely these two ascertainment methods for 2D:4D. Quite contrary, it has been stated that «measurement of digit length can be performed in a number of ways». There is evidence that the 2D:4D ratio «does not arise from metacarpal length but is an independent characteristic of the digits» and that 2D:4D values obtained from photocopies tally to those obtained from radiographs of the hand, «despite the fact that the soft-tissue measurements of the digits were taken from approximately halfway along the proximal phalanx whereas bone measurements began at the proximal end of the phalanx».

Even measurements from hand outline drawings yield 2D:4D values equivalent to the photocopy method we used, as well as 2D:4D values obtained via inked prints are equivalent to those obtained via photocopies. Further on, owing to the narrow sampling approaches, both the historical and the contemporary sample probably were not representative for the 2D:4D distribution in those regional populations, which could impair the validity of this finding. Although there is no strong evidence for assuming that the two samples could stem from subpopulations with differing gene pools.
REFERENCES


M. Voracek
Department of Basic Psychological Research, School of Psychology, University of Vienna, Liebiggasse 5, Rm 03-42, A-1010 Vienna, Austria
e-mail: martin.voracek@univie.ac.at

OMJER DULJINE PRSTIJU (2D:4D) U LITVI NEKAD I DANAS: TESTIRANJE SPOLNIH RAZLIKA, POVEZANOST S BOJOM OČIJU I KOSE, I MOGUĆE DUGOTRAJNE PROMJENE

SAŽETAK

Omjer drugog i četvrtog prsta (2D:4D) je spolno dimorfno somatsko svojstvo pa se smatra biološkim markerom za organizacije, tj. trajne učinke prenatalnog testosterona na ljudski mozak. U skladu s tim u nedavnom je istraživanju omjer 2D:4D povezan s različitim spolno-specifičnim, hormonalno utjecanim svojstvima i fenotipovima. Geografska varijacija u tipičnom 2D:4D omjeru je znatna i zasad još slabo razumijevana. Ovo istraživanje omjera 2D:4D predstavlja primjer posljedice naših evolucijskih korijena i ima značajno biološko, geneticko i socijalno značenje. U skladu s tim u nedavnom je istraživanju omjer 2D:4D podržava da je moguće da se istraživanje omeđuje na primjerom pojedinog gena ili gena koji utječu na omjer 2D:4D. Međutim, zbog eksperimentalnih tehnika koje se koriste u istraživanju, nije moguće utvrditi značajnost ove nezavisnosti na ovoj pečat. Zbog toga je potrebno da se istraživanje omeđuje na primjerom pojedinog gena ili gena koji utječu na omjer 2D:4D. Međutim, zbog eksperimentalnih tehnika koje se koriste u istraživanju, nije moguće utvrditi značajnost ove nezavisnosti na ovoj pečat. Zbog toga je potrebno da se istraživanje omeđuje na primjerom pojedinog gena ili gena koji utječu na omjer 2D:4D.