Hormones and Behavior 120 (2020) 104710

Contents lists available at ScienceDirect

Hormones and Behavior

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

Commentary

Facing off with the phalangeal phenomenon and editorial policies: A commentary on Swift-Gallant, Johnson, Di Rita and Breedlove (2020)☆

Cheryl M. McCormick[a,*], Justin M. Carré[b]

[a] Department of Psychology, Brock University, St Catharines, ON L2S 3A1, Canada
[b] Department of Psychology, Nipissing University, North Bay, ON, Canada

ARTICLE INFO

Keywords:
Digit ratio
2D:4D
Facial width-to-height ratio

ABSTRACT

Swift-Gallant et al. (2020) provide a thought-provoking perspective on the topic of digit ratio research, research that has had some prominence in the journal *Hormones and Behavior*, and is research that has garnered much controversy. In this commentary on their paper, we add to the discussion of why there is skepticism of the use of digit ratios as a measure of individual differences in prenatal androgens, we comment on the mis-use of the facial width-to-height ratio as a measure of individual differences in testosterone, the grey areas in the interpretation of evidence, and we address the concern raised in their article regarding editorial policies at *Hormones and Behavior* (spoiler alert: there are no secret policies).

S. Marc Breedlove's lab has made significant contributions to *Hormones and Behavior* with papers on several topics and as a member of the editorial board, and I (CMM) was thus delighted when Marc accepted my invitation to provide a review of the field of digit ratios for the 50th Anniversary Issue. Marc is a leader in such research, and it is research that is highly cited. There are 12 papers involving digit ratios in *Hormones and Behavior* since the first by Brown et al. (2002), and the papers include contributions from other labs and in species other than humans (ring-necked pheasant, Romano et al., 2005; Guinea baboons, Roney et al., 2004). Swift-Gallant et al.'s (2020) paper is not so much a thorough review of the digit ratio literature, but more of a perspective on this field of research, and provides much to consider regarding the basis for the sex difference in digit ratios and the extent to which digit ratios may be a reflection of prenatal androgenization. We have quibbles, however, with statements in the paper, some of which involve the role of Editor of *Hormones and Behavior* and others around the interpretation of evidence.

First, there is no, and as far as we know, there never has been, any policy of refusing to review, sending out for review, or publishing papers involving digit ratios in *Hormones and Behavior*. To wit, see: Nitschke and Bartz (2020), Lower digit ratio and higher endogenous testosterone are associated with lower empathic accuracy. *Hormones and Behavior*, *119*, 104648, a manuscript that was sent out for review by Kim Wallen, the previous editor of *Hormones and Behavior* shortly before my term (CMM) began, and a manuscript that I accepted for publication after the review process. There are no secret policies at *Hormones and Behavior*. A main consideration for any paper on any topic is whether the research presented is appropriate for the journal, given its scope. The main grey area has been in the decision of whether there is enough "hormones" in a paper or whether there is enough "behaviour" in a paper. As stated on the website, "*Hormones and Behavior* publishes original research articles, reviews and special issues concerning hormone-brain-behavior relationships, broadly defined." It is important that the journal reflects the interests of its society, the Society for Behavioral Neuroendocrinology, hence the grey area and hence the term "broadly defined".

Second, we don't think the review gets at a critical reason as to "why are digit ratios so maligned" (Swift-Gallant et al., 2020). We write this as researchers who have conducted their own research with an easy to measure and controversial marker, the facial width-to-height ratio, a.k.a. the face ratio (no paper on which has ever appeared in *Hormones and Behavior*, we might add, and likely rightly so). This marker has provided us with some of the biggest, most reliable effects that we have ever found in our research careers: the consistency with which observers rate faces in terms of various characteristics (notably aggressiveness), and the association of these ratings with the face ratio (see meta-analysis by Geniole et al., 2015). The extent to which this measure is predictive of behaviour is less strong and less reliable, nevertheless (Geniole et al., 2015). And the sex difference that propelled us to investigate this measure based on the initial report of this metric by

☆ This paper is part of the Virtual Special Issue, 50th Anniversary of Hormones and Behavior.

* Corresponding author.
E-mail address: cmccormick@brocku.ca (C.M. McCormick).

https://doi.org/10.1016/j.yhbeh.2020.104710
Received 8 January 2020; Accepted 10 January 2020

Weston et al. (2007) is even more dubious (yet, we still maintain that the within-sex variation may be meaningful). We speculated early on that this marker might be related to pubertal androgens in men, but we had, and still have, no evidence for this speculation. It was simply speculation. Yet the literature began using the measure as a marker of testosterone (prenatally, pubertally, adult levels, depending on the study), often citing our paper (Carré and McCormick, 2008) as the basis for the assertion without acknowledging that we had made a speculation, not an assertion, and that we had no data in this paper or any other to make such an assertion. We have strongly called out such claims when acting as a reviewer for such a paper, but we are not always a reviewer of such papers. See for example the statement in one abstract: "To measure testosterone exposure, we apply the facial width-to-height metric…" (Kozlov et al., 2018). Similar statements are abundant in the digit ratio literature, with the digit ratio cast as a proxy for prenatal testosterone exposure.

We are aware of the skepticism researchers have for the face ratio – partly because of these types of unsubstantiated claims regarding testosterone and because soon the face ratio became associated with just about everything under the sun and often without any theoretical motivation, as have digit ratios. A few random examples for digit ratios:

Rogers, P., Caswell, N., Brewer, G. (2017). 2D: 4D digit ratio and types of adult paranormal belief: An attempted replication and extension of Voracek (2009) with a UK sample. *Personality and Individual Differences*, *104*, 92–97.

Friedl, A., Neyse, L., Schmidt, U. (2018). Payment scheme changes and effort adjustment: the role of 2D: 4D digit ratio. *Journal of Behavioral and Experimental Economics*, *72*, 86–94.

Huh, H. (2012). Born to be a Marine: Digit ratios and military service. *Personality and Individual Differences*, *53*(3), 166–168.

Hopp, R. N., de Moraes, J. P., Jorge, J. (2012). Digit ratio and academic performance in dentistry students. *Personality and Individual Differences*, *52*(5), 643–646.

Aycinena, D., Rentschler, L. (2018). Discounting and digit ratio: Low 2D: 4D predicts patience for a sample of females. *Frontiers in Behavioral Neuroscience*, *11*, 257.

Nepomuceno, M. V., Saad, G., Stenstrom, E., Mendenhall, Z., & Iglesias, F. (2016). Testosterone at your fingertips: Digit ratios (2D: 4D and rel2) as predictors of courtship-related consumption intended to acquire and retain mates. *Journal of Consumer Psychology*, *26*(2), 231–244.

This small sample is representative of many digit ratio research papers. It might be notable that most of this research is found in social science journals that typically do not involve neuroendocrine measures. It is this vast literature that underlies much of the skepticism regarding the digit ratio. Nevertheless, let us assume that many of the reported relationships are valid and reliable. Let us also assume that there is a relationship between prenatal testosterone and digit ratios, which, consistent with the Swift-Gallant et al. (2020) review, is likely weak. On that basis, any variance in the digit ratio other than that related to prenatal testosterone could be the actual basis of any relationship between the digit ratio and another measure. Thus, any paper involving the digit ratio cannot conclude that a relationship between the digit ratio and a measure is an indication that there is a relationship between prenatal testosterone and the measure. And thus, any paper involving the digit ratio and a measure will fall in the grey area: Is there enough "hormones" in the research to be in *Hormones and Behavior*? When the decision is yes (e.g., as in Nitschke and Bartz, 2020, which also included measures of testosterone), we would not allow a paper to matter-of-factly state that digit ratios are a marker of individual differences in fetal testosterone exposure, we would request that authors note the discrepancies in the literature and the failures to replicate alongside the replications, and we would expect that the conclusions would be based on the evidence presented and distinct from speculation.

Swift-Gallant et al.'s (2020) interpretation of the digit ratio evidence is theoretically motivated, and they are restricting their perspective to evidence of a sex difference in the digit ratio and a relationship with sexual orientation. Interpretation of evidence, however, involves another grey zone and hence is part of the controversy. They interpret the evidence, primarily from individuals with disorders of sexual development, to indicate that prenatal levels of androgens affect human digit ratios. Yet, based on the effect sizes, this same evidence also argues against the use of digit ratios as a measure of individual differences in prenatal androgens (e.g., Berenbaum et al., 2009; Van Hemmen et al., 2017), and a reason as to why papers using them in such a way would be difficult to publish in *Hormones and Behavior*. Further, the evidence for a sex difference in the digit ratio and the size of any such effect also is a matter of much debate. See, for example, Forstmeier (2018) for an attempt to resolve the statistical problems inherent in assessing whether sex differences are an artefact of allometry. Meta-analyses are useful tools for determining replicability and effect sizes, yet they too exist in a grey area of interpretation. The statement that the finding that lesbians have smaller digit ratios than straight women has been confirmed by meta-analysis is debatable. The meta-analysis by Grimbos and colleagues (2010) cited by SwiftGallant and colleagues (2020) is far from definitive. Specifically, this meta-analysis reported that lesbians had smaller (more 'masculinized') right and left hand digit ratios relative to straight women (Hedges $gs = 0.29$ and 0.23, respectively). However, after correcting for publication bias using the trim-and-fill technique, effects sizes decreased substantially for right and left hand digit ratios (Hedges $gs = 0.13$ and 0.07, respectively) and are no longer significant. Despite the presence of publication bias, the authors of the meta-analysis argued that their "fail-safe" analysis suggested that more than 40 null findings would need to exist to effectively render the difference in digit ratio between lesbian and straight women statistically non-significant, implying that publication bias in this instance was not a huge problem. However, Rosenthal's (1979) fail-safe method has been widely criticized on methodological grounds (e.g., Scargle, 2000) and some authors have suggested that it "gives the meta-analytic researcher a false sense of security" regarding the stability and trustworthiness of meta-analytic estimates (Ferguson and Heene, 2012). Thus, although Swift-Gallant et al. (2020) are convinced that digit ratios have some link to prenatal hormones in humans, many won't be, based on the available evidence.

We thank the authors for providing us the opportunity to address the topic of editorial policies at the journal. We don't have formal policies, but we do have a set of best practices to which we try to adhere: to provide a fair and constructive review process in a timely manner and to publish high-quality research in behavioural neuroendocrinology.

References

Berenbaum, S.A., Bryk, K.K., Nowak, N., Quigley, C.A., Moffat, S., 2009. Fingers as a marker of prenatal androgen exposure. Endocrinology 150, 5119–5124.

Brown, W.M., Hines, M., Fane, B.A., Breedlove, S.M., 2002. Masculinized finger length patterns in human males and females with congenital adrenal hyperplasia. Horm. Behav. 42, 380–386.

Carré, J.M., McCormick, C.M., 2008. In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. Proc. R. Soc. B 275, 2651–2656.

Ferguson, C.J., Heene, M., 2012. A vast graveyard of undead theories: publication bias and psychological science's aversion to the null. Perspect. Psychol. Sci. 7, 555–561.

Forstmeier, W., 2018. Avoiding misinterpretation of regression lines in allometry: is sexual dimorphism in digit ratio spurious? BioRxiv 298786.

Geniole, S.N., Denson, T.F., Dixson, B.J., Carré, J.M., McCormick, C.M., 2015. Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. PLoS One 10 (7), e0132726.

Grimbos, T., Dawood, K., Buriss, R.P., Zucker, K.J., Puts, D.A., 2010. Sexual orientation and the second to fourth finger length ratio: a meta-analysis in men and women. Behavioral Neuroscience 124, 278–287.

Kozlov, V., Libman, A., Schultz, A., 2018. Testosterone and repression in non-democracies: evidence from a sample of Russian governors. Kyklos 71, 244–278.

Nitschke, J.P., Bartz, J.A., 2020. Lower digit ratio and higher endogenous testosterone are associated with lower empathic accuracy. Horm. Behav. 119, 104648.

Romano, M., Rubolini, D., Martinelli, R., Alquati, A.B., Saino, N., 2005. Experimental manipulation of yolk testosterone affects digit length ratios in the ring-necked pheasant (Phasianus colchicus). Horm. Behav. 48, 342–346.
Roney, J.R., Whitham, J.C., Leoni, M., Bellem, A., Wielebnowski, N., Maestripieri, D., 2004. Relative digit lengths and testosterone levels in Guinea baboons. Horm. Behav. 45, 285–290.
Rosenthal, R., 1979. The file drawer problem and tolerance for null results. Psychol. Bull. 86, 638–641.
Scargle, J.D., 2000. Publication bias (the "file-drawer problem") in scientific inference. J. Sci. Explor. 14, 91–106.
Swift-Gallant, A., Johnson, B.A., Di Rita, V., Breedlove, S.M., 2020. Through a glass, darkly: human digit ratios reflect prenatal androgens. Horm. Behav. 120, 104686 (in press).
Van Hemmen, J., Cohen-Kettenis, P.T., Steensma, T.D., Veltman, D.J., Bakker, J., 2017. Do sex differences in CEOAEs and 2D: 4D ratios reflect androgen exposure? A study in women with complete androgen insensitivity syndrome. Biol. Sex Diff. 8, 11.
Weston, E.M., Friday, A.E., Liò, P., 2007. Biometric evidence that sexual selection has shaped the hominin face. PLoS One 2, e710.