



## Short Communication

## An endocrine basis for tomboy identity: The second-to-fourth digit ratio (2D:4D) in “tomboys”

Beth M. Atkinson<sup>a</sup>, Tom V. Smulders<sup>a</sup>, Joel C. Wallenberg<sup>a,b,\*</sup><sup>a</sup> Centre for Behaviour and Evolution, Institute of Neuroscience, Newcastle University, Newcastle Upon Tyne NE2 4HH, UK<sup>b</sup> School of English Literature, Language and Linguistics, Newcastle University, Newcastle Upon Tyne NE1 7RU, UK

## ARTICLE INFO

## Article history:

Received 13 December 2016

Received in revised form 27 January 2017

Accepted 27 January 2017

## Keywords:

Gender

Neuroendocrinology

Testosterone

2D:4D

Tomboys

## ABSTRACT

This paper investigates the relationship between organizational effects of pre-natal testosterone and the use of “tomboy” as a descriptor for young women. We show in a sample of 44 women that a woman’s right hand 2D:4D ratio is a significant predictor of whether they will be labeled as a “tomboy”, with a decrease in 2D:4D ratio corresponding to an increase in the probability of being called “tomboy”. Taking the right hand 2D:4D ratio as a proxy for the abundance of testosterone in the early life hormonal milieu, we propose that organizing effects of higher pre-natal T lead to increased masculine-typical behavior in childhood, which increases the likelihood some women will be referred to as tomboys. We suggest that the increase in masculine-typical behaviors is a result of how the organizing effects of T on the brain interact with children’s social modeling of male-coded and female-coded behaviors.

© 2017 Elsevier Ltd. All rights reserved.

## 1. Introduction

A growing field of behavioral neuroendocrinology has shown that gender- and sex-related behavior in both humans and non-humans is strongly affected by the perinatal hormonal milieu of the developing organism (see Balthazart (2011) for a cross-species review of sexual behavior; Cohen-Bendahan et al. (2005), *inter alia* on gendered behavior in humans). In particular, a number of studies have shown that gender identity and behavior in children is correlated with levels of pre-natal testosterone (Auyeung et al., 2009; Berenbaum and Bailey, 2003; Hines et al., 2016; Pasterski et al., 2005).

In studies of women with Congenital Adrenal Hyperplasia (CAH), childhood play behavior and gender identity were significantly influenced by prenatal androgens (Hines et al., 2004; Pasterski et al., 2005). CAH increases circulating testosterone levels in early life, which are standardly corrected by medical intervention after birth. Ehrhardt et al. (1968) also noted that CAH women were more likely to be described as “tomboy” by their mothers than non-CAH women. Salmon (2015) showed that “tomboy” status correlated with fiction choices in women, and suggested this result was

due to pre-natal masculinization. However, no studies to date have connected levels of prenatal androgens to the “tomboy” label in a non-clinical population. Our study fills this gap in the literature.

While it is difficult to access information about the early life hormonal milieu in non-clinical populations, proxies do exist. Here we use the 2nd-digit-to-4th-digit (2D:4D) length ratio, a proxy for early life androgen exposure which has been validated by experimental manipulation of T in rats (Talarovičová et al., 2009), and direct measures of the T/estradiol ratio in human amniotic fluid (Lutchmaya et al., 2004). We hypothesize that, within a population of women (assigned female at birth and female-identifying), an individual’s *in utero* exposure to testosterone is predictive of whether they would be called a “tomboy” in later life.

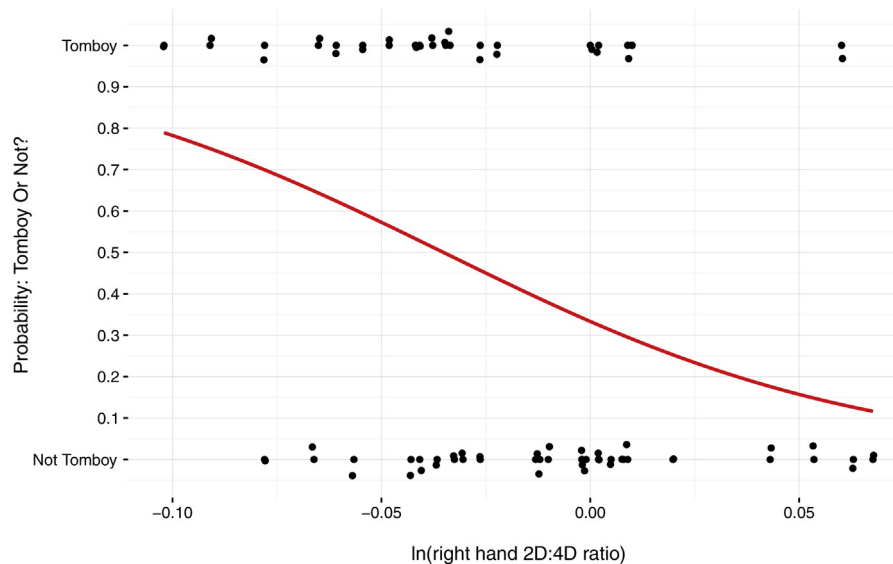
## 2. Methods

## 2.1. Participants

The participants for this study were recruited primarily in two ways: an advert was distributed via Newcastle University’s Institute of Neuroscience database of volunteers, and posters were displayed on both Newcastle and Northumbria University campuses. For this reason, participants were largely in university or post-university. The advert mentioned that the study related to gender and hormones, and asked for participants with intact fingers and no endocrine conditions. (The advert did not contain the word “tomboy”.) Before participation in the study, participants

\* Corresponding author at: Percy Building, Newcastle University, Newcastle Upon Tyne NE1 7RU, UK.

E-mail addresses: [bethatkinson1008@gmail.com](mailto:bethatkinson1008@gmail.com) (B.M. Atkinson), [tom.smulders@ncl.ac.uk](mailto:tom.smulders@ncl.ac.uk) (T.V. Smulders), [joel.wallenberg@ncl.ac.uk](mailto:joel.wallenberg@ncl.ac.uk) (J.C. Wallenberg).



**Fig. 1.** Relationship between the probability of “tomboy” label and log of the right hand 2D:4D ratio, modeled by the logistic. (Actual data points, which are binary, are shown on the “Tomboy” and “Not Tomboy” lines.)

were briefed on the procedure, and informed of our policies of confidentiality and data protection. A consent form was then signed by both the subject and experimenter.

The data presented here is based on 44 female participants, with ages 18–68 (mean = 31.5, “tomboys” mean = 31, non-tomboys = 31.9), a sub-study of a larger project on gender with 96 participants. 2 participants were excluded because of unreadable hand-scans, 5 because of incomplete questionnaire answers, and the remaining 45 in the larger project were men, with 3 excluded for incomplete information and unreadable hand-scans (mean age = 31.7). (Though these remaining 42 men are not the focus of this study, we discuss them for comparison below.) Participants were also asked their ethnicity: the 44 women included 35 white British, 5 Chinese, 7 South Asian, 1 who identified as “mixed”, and 2 who did not report ethnicity; the 42 men were 41 white and 1 South Asian.

## 2.2. Measuring 2D:4D ratio

Participants were asked to place their right hand face down on a “Doxie portable scanner” (resolution: 600 dpi). A scan was taken, and this was repeated for the left hand. These images were saved as PNG files (lossless compression), and the scans were assessed for their suitability: the basal creases of the fingers had to be in frame and clear. Digit length was measured with digital calipers in the Gnu Image Manipulation Program (GIMP). Allaway et al. (2009) have shown digital measurement from scans to be the most consistent and accurate method of determining the 2D:4D ratio.

While we measured the participants’ 2D and 4D lengths, they completed the written questionnaire described below.

## 2.3. In-house questionnaire

Participants completed a short questionnaire containing a gender identity assessment, 8 additional questions on gender and sexuality, and some basic demographic information. A researcher was available to answer any questions, but she did not read the questionnaire as the participants were writing, to ensure confidentiality.

The first question in the questionnaire was ‘If female, were you ever described as a tomboy?’. We asked this question first so that there would be little risk of priming from the other questions. Due

to the gendered nature of the term “tomboy”, we only considered women for this study.

Each participant was assigned a random number for anonymity, which was written on their questionnaire with the date and time of appointment. Finally, a debrief was administered and all participants were given £5.

## 2.4. Analysis

Statistical analyses were made using R, and figures were created in R using the ggplot2 package.

## 3. Results

In our sub-study of 44 participants (19 “tomboys” and 25 non-tomboys), the results clearly show that lower 2D:4D digit ratios correspond to an increased probability of being called a tomboy (mean right hand 2D:4D ratio of 0.969 for tomboy group, 0.994 for non-tomboy group; Fig. 2).

We modeled whether a person was labeled a tomboy as a binary outcome in a logistic regression with the right hand 2D:4D ratio as a continuous predictor. A model including the 2D:4D term fits significantly better than the null model ( $p = 0.019$ ). (Note: the model used  $\ln(2D:4D)$  rather than raw ratios to correct for skew.) The slope of the model is  $-19.7$ , which is reliably different from zero ( $p = 0.03$ ), and the relationship between the predictor and response variables is in the hypothesized negative direction: lower 2D:4D ratios correspond to a higher probability of the tomboy label being applied to an individual (Fig. 1). An estimated odds ratio of 1.22, derived from the slope divided by 100 (for scale), provides a sense of the size of effect: every decrease of 0.01 in  $\ln(2D:4D)$  multiplies the odds of being a tomboy 1.22 times. A Mann–Whitney–Wilcoxon test confirms this ( $U = 328$ ,  $p = 0.03$ ). No significant effect holds for the 2D:4D ratios of the participants’ left hands.

Age of the participants did not have an effect, based on model comparison ( $p = 0.99$ ), and the interaction between Age and right hand 2D:4D was not significant ( $p = 0.13$ ). Additionally, the Akaike and Bayesian Information Criteria (AIC and BIC, measures of information lost in a model) were lower for a model without Age, so we removed Age from the final logistic model reported above.

Similarly, ethnicity of the participants did not show any significant effects ( $p = 0.29$  for interaction with 2D:4d,  $p = 0.44$  for main

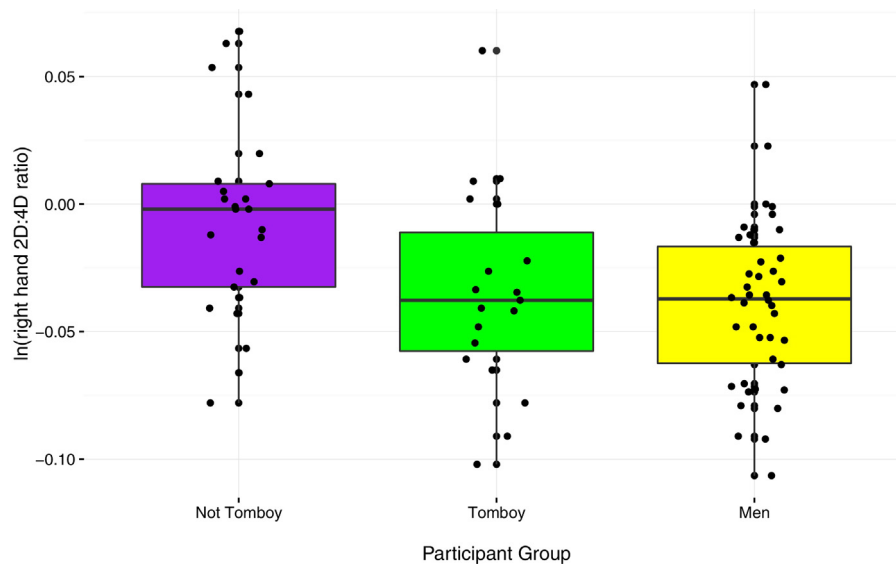


Fig. 2. Mean right hand ratios for the tomboy and non-tomboy participants.

effect of ethnicity only, and models including ethnicity had higher AIC and BIC); thus, we have removed ethnicity from the final model reported above. However, for completeness, we also modeled the subset of the data with only white participants, and again found the significant effect of  $\ln(\text{right-2D:4D})$ :  $p = 0.004$ , with an estimated slope of  $-28.7$  (reliably non-zero,  $p = 0.019$ ), odds ratio of 1.33 per hundredth of  $\ln(2\text{D:4D})$ .

To verify that our measuring technique was valid, we also replicated the well-established male/female sex/gender difference in 2D:4D ratio, using the 86 participants in our larger sample: men,  $N = 42$ , mean right-2D:4D = 0.962; all women,  $N = 44$ , mean right-2D:4D = 0.982; Mann-Whitney  $U = 636.5$ ,  $p = 0.013$  (men also shown in Fig. 2). Modeling male vs. female as a binary outcome with a logistic, the sex factor (significant at  $p < 0.001$ ) has an estimated slope of  $-24.2$  (different from zero,  $p < 0.01$ ), showing that the odds of being a man goes down 1.27 times per hundredth of  $\ln(2\text{D:4D})$ ; compare to 1.22 for tomboys above.

#### 4. Discussion

The results above show that the right hand 2D:4D ratio is a predictor of whether a woman (or female-presenting individual) will be labeled a tomboy during their childhood, which we take to mean that the organizational effects of *in utero* androgens plausibly bias individuals toward the more masculine-typical behaviors that society associates with the tomboy label. This finding is in accord with the literature supporting the general organizational/activational hypothesis, and particularly the idea that androgens can act on the brain during a critical period in early life and influence later life behaviors (Arnold, 2009), such as play in childhood (Auyeung et al., 2009; Hines et al., 2004; Pasterski et al., 2005).

Our finding could have a number of causes. One possibility is that behaviors related to the tomboy label are directly driven by pre-natal T levels. Play-styles and toy choice have been found to be sexually dimorphic in non-humans, and toy choice especially is not easily explained by social factors in animals without a toy culture (e.g. Alexander and Hines (2002) for vervet monkeys).

The effect could also stem from a lower tendency on the part of tomboys to emulate female-coded behaviors (Hines et al., 2016). Since participants answered “Were you ever described as a tomboy?”, rather than a question more focused on internal gender

identity, we can infer that the tomboys engaged disproportionately in behaviors that others in the surrounding society perceived as masculine-coded. Hines et al. (2016) showed CAH girls to have a lower desire/ability to learn and/or emulate female-coded behaviors than non-CAH girls. When a young woman displays a higher proportion of male-coded behaviors (or lower of female-coded behaviors) than the average for their age-cohort in a given society, then they are likely to be labeled “tomboy”.

Importantly, this mechanism can explain the application of labels like “tomboy” as a function of hormonal organizing effects independently of how binarized the male-female expectations of a given societal context are. In other words, we’d predict this effect to hold both in societies where many behaviors are strictly male-coded and female-coded, and in societies which do not have such strict expectations about gendered behavior (assuming they both have a tomboy label). The fact that Age was not a significant factor in our sample might indicate that this prediction is correct: the effect of 2D:4D on tomboy labeling was plausibly constant across the ages of our participants, even though the participants in their 60s grew up in a society with different gender expectations than the one our 20-year-olds grew up in. This prediction could be tested more robustly by comparing societies with a tomboy label, but with different gender-role expectations.

The social learning account above is not a mutually exclusive with the idea that some of these directly influenced behaviors are responsible for tomboy-labeling; we leave open the possibility that both could be responsible for the effect observed here.

The fact that the effect is only significant for the right hand 2D:4D of the participants is in line with studies showing stronger behavioral, sex, or gender identity effects for right than for left hand 2D:4D (Cohen-Bendahan et al., 2005). Manning et al. (1998) show that sexual dimorphism on a number of gonad-related biological measures are predicted by the right hand 2D:4D and not the left hand. Additionally, Lutchmaya et al. (2004) found a significant relationship between the fetal T/estradiol ratios of participants (measured in amniotic fluid) and the participants’ 2D:4D ratios after they were born, but the relationship was **only** significant for the right hand ratios. Thus, the right hand 2D:4D ratio seems to be a more reliable proxy for the *in utero* hormonal milieu than the left hand ratio.

## 5. Limitations of the study

One clear limitation of this study is the use of the 2D:4D ratio, which is not uncontroversial as a proxy for the pre-natal hormonal milieu. As mentioned above, Lutchmaya et al. (2004) found a relationship between the T/E ratio in amniotic fluid and 2D:4D, not the straightforward relationship between amniotic T and 2D:4D that Talarovičová et al. (2009) found in rats, indicating that human 2D:4D could develop from a somewhat different mechanism. See also the review of 2D:4D results in Cohen-Bendahan et al. (2005) for some inconsistent results regarding 2D:4D and behavior, and discussion in Wong and Hines (2016), who conclude that 2D:4D is a weak marker of pre-natal androgens.

Another limitation of the study is the sample for the tomboy analysis includes only women. Ideally the study would have included men as well, and shown a 2D:4D effect for the male equivalent of “tomboy”; unfortunately, this was not possible, as there is no equivalent term for men which does not carry a much stronger stigma, making a straightforward comparison impracticable. (Note, however, that we did find the expected sex difference in 2D:4D in our larger study, as reported in Section 3.) The sample is also smaller than ideal, given e.g. the sample of 207 gay women Brown et al. (2002) used to show a 2D:4D difference in “butch” vs. “femme” women. Finally, though the participants were drawn from a population that is overwhelmingly post-university and so of relatively similar socioeconomic status (SES), we did not explicitly collect SES information on participants. We acknowledge this as a shortcoming of the study, though we have no reason to believe that the small amount of variation was not randomly distributed in our sample.

## 6. Conclusions

We have shown a significant relationship between the right hand 2D:4D ratio in women and their likelihood of being called a “tomboy” by others around them. We interpreted this effect to mean that organizational effects of higher levels of *in utero* T exposure, which correspond to lower right hand 2D:4D ratios, lead to more masculine-typical behaviors in some women. This in turn makes them more likely candidates for being called tomboys. Though the “tomboy” label has been noted in various studies as an important predictor (or reflection) of various behaviors in women, this is the first study to show a direct link between that label and a measure of pre-natal T exposure in a non-clinical population.

## Ethics

Ethical approval for the study was obtained from the Newcastle University's Faculty of Humanities and Social Sciences Committee on Research Ethics.

## Conflicts of interest

None.

## Data, code and materials

Data set of measurements and questionnaire responses available at: <https://github.com/joelcw/gendocrine/blob/master/datasets/tomboys.csv> <https://github.com/joelcw/gendocrine/blob/master/datasets/men.csv> and R code at: <https://github.com/joelcw/gendocrine/blob/master/R/tomboyAnalysis.R>. The questionnaire can be found at: <https://github.com/joelcw/gendocrine/blob/master/forms/questionnaire.pdf>.

## Author contributions

BA conducted primary collection of data, made hand measurements, and conducted some statistical analysis. TS contributed to the study design, supervised data collection, and helped draft the article. JW conceived the study, supervised data collection, and conducted some statistical analyses. All authors helped draft the article.

## Funding

Funded by the Faculty of Medical Sciences at Newcastle University as part of BA's MRes studies.

## Acknowledgements

We thank our participants, and Jacques Balthazart for helpful discussion, as well as two anonymous reviewers.

## References

- Alexander, G.M., Hines, M., 2002. Sex differences in response to children's toys in nonhuman primates (*cercopithecus aethiops sabaeus*). *Evolut. Hum. Behav.* 23 (6), 467–479.
- Allaway, H.C., Blois, T.G., Pierson, R.A., Lujan, M.E., 2009. Digit ratios (2D:4D) determined by computer-assisted analysis are more reliable than those using physical measurements, photocopies, and printed scans. *Am. J. Hum. Biol.* 21 (3), 365–370.
- Arnold, A.P., 2009. The organizational–activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues. *Horm. Behav.* 55 (5), 570–578.
- Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., Hackett, G., Hines, M., 2009. Fetal testosterone predicts sexually differentiated childhood behavior in girls and in boys. *Psychol. Sci.* 20 (2), 144–148.
- Balthazart, J., 2011. *The Biology of Homosexuality*. Oxford University Press, Oxford.
- Berenbaum, S.A., Bailey, J.M., 2003. Effects on gender identity of prenatal androgens and genital appearance: evidence from girls with congenital adrenal hyperplasia. *J. Clin. Endocrinol. Metab.* 88 (3), 1102–1106.
- Brown, W.M., Finn, C.J., Cooke, B.M., Breedlove, S.M., 2002. Differences in finger length ratios between self-identified “butch” and “femme” lesbians. *Arch. Sex. Behav.* 31 (1), 123–127.
- Cohen-Bendahan, C.C., van de Beek, C., Berenbaum, S.A., 2005. Prenatal sex hormone effects on child and adult sex-typed behavior: methods and findings. *Neurosci. Biobehav. Rev.* 29 (2), 353–384.
- Ehrhardt, A.A., Epstein, R., Money, J., 1968. Fetal androgens and female gender identity in the early-treated adrenogenital syndrome. *Johns Hopkins Med. J.* 122 (3), 160.
- Hines, M., Brook, C., Conway, G.S., 2004. Androgen and psychosexual development: core gender identity, sexual orientation, and recalled childhood gender role behavior in women and men with congenital adrenal hyperplasia (CAH). *J. Sex Res.* 41 (1), 75–81.
- Hines, M., Pasterski, V., Spencer, D., Neufeld, S., Patalay, P., Hindmarsh, P.C., Hughes, I.A., Acerini, C.L., 2016. Prenatal androgen exposure alters girls' responses to information indicating gender-appropriate behaviour. *Philos. Trans. R. Soc. B: Biol. Sci.* 371 (1688), 20150125.
- Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., Manning, J.T., 2004. 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Hum. Dev.* 77 (1), 23–28.
- 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. *Hum. Reprod.* 13 (11), 3000–3004.
- Pasterski, V.L., Geffner, M.E., Brain, C., Hindmarsh, P., Brook, C., Hines, M., 2005. Prenatal hormones and postnatal socialization by parents as determinants of male-typical toy play in girls with congenital adrenal hyperplasia. *Child Dev.* 76 (1), 264–278.
- Salmon, C., 2015. The impact of prenatal testosterone on female interest in slash fiction. *Evolut. Behav. Sci.* 9 (3), 161.
- Talarovičová, A., Kršková, L., Blažeková, J., 2009. Testosterone enhancement during pregnancy influences the 2D:4D ratio and open field motor activity of rat siblings in adulthood. *Horm. Behav.* 55 (1), 235–239.
- Wong, W.I., Hines, M., 2016. Interpreting digit ratio (2D:4D)-behavior correlations: 2D:4D sex difference, stability, and behavioral correlates and their replicability in young children. *Horm. Behav.* 78, 86–94.