There is no correlation between prenatal testosterone exposure and aggression in women.

Abstract

Does prenatal testosterone exposure correlate with aggression in women? 15 female participants completed the Buss-Perry Aggression Questionnaire (Short Form). Their prenatal testosterone exposure was assessed by calculating their D2:D4 digit ratios. There was no correlation between digit ratios and aggressions scores (R=-0.014; 1 tailed p>0.05). This suggests that there is no relationship between prenatal testosterone exposure and aggression in women. However, there are doubts about the validity of the measurements of both aggression and digit ratio in this study.

Aim and background

The aim of this investigation is to find out if there is evidence of a link between testosterone exposure and aggression. It is claimed that testosterone is the cause of stereotypically male behaviours including aggression because it affects the structure and functioning of the brain. In this investigation D2:D4 digit ratio was used as a way of measuring testosterone exposure because a higher level of testosterone during prenatal development causes the fourth finger to grow longer relative to the index finger (Manning, 2002). The digit ratio is found by dividing the length of the index finger (D2) by the length of the fourth finger (D4). Aggression was measured using a questionnaire (the BPAQ-SF; Diamond, 2006) which gives a quantitative measurement of aggression, where a higher score indicates higher aggression. If prenatal testosterone is related to aggression then there should be a correlation between between D2:D4 digit ratios and aggression scores.

Hypotheses

H_1: There will be a negative correlation between D2:D4 digit ratios and BPAQ-SF aggression scores.

H_0: There will be no correlation between D2:D4 digit ratios and BPAQ-SF aggression scores.

Method

Design

This is a study with a correlational design. The two covariables were:

- Testosterone exposure, operationalised as the difference in length between D2 and D4, obtained by dividing D2 by D4.
- Aggression, operationalised as scores on on the Short-Form Buss-Perry Aggression Questionnaire.

Extraneous variables were controlled by standardising the procedure including the conditions under which the measurements were taken and the instructions given to the participants. A standardised procedure for measuring the participants' fingers was used.

Participants

The participants were an opportunity sample of 15 female psychology students from a selective school in the Midlands. Their mean age was 16 years 4 months. They came from a variety of ethnic backgrounds.

Materials

The BPAQ-SF (Diamond, 2006) is a 12-question instrument consisting of 5-point Likert scale items (1 = very unlike me; 5 = very like me). All items are scored positively. The overall score is calculated by summing the responses for each item. A ruler and a sheet of A3 paper were also required.

Procedure

Participants were recruited because they were in the classroom when the study took place. The nature of the study was explained and, following ethical briefing, students were invited to take part. Those that agreed to do so were given a copy of the BPAQ-SF and asked to complete it. They were given the following instructions:

> *This questionnaire is anonymous and the data you provide will not be linked to you individually. You do not have to complete the questionnaire and you may decide not to answer any of the individual questions once the study is complete. If you decide during the study that you do not wish to continue, then just stop. If, after the study, you decide you would rather not have taken part, please tell me and I will remove your data from the study.*
>
> *Please read each of the following statements and decide whether it describes you well. Don't think too hard about your answer, just give the rating that seems right to you.*

Participants completed the questionnaire simultaneously whilst sat in their usual seats. Once they had completed and scored the questionnaire, they were asked, one at a time, to come and sit at a table. They were asked to put their right hand, palm up with fingers together, on a sheet of A3 paper placed on the table. The researcher then used a ruler to measure the distance between the bottom of the first digital crease to the fingertip of D2. This was repeated for D4. Once all participants had been measured the nature of the study was explained and participants were thanked, given the opportunity to withdraw their data and asked if they had any questions.

Results

Results were analysed using Google Sheets software. Questionnaire responses that were incomplete were removed from the data. Digit ratios were calculated by dividing each participant's D2 measurement by their D4 measurement.

The mean aggression score was 26.13 with a standard deviation of 7.84, showing that there was a moderate amount of variation in the participants' aggression scores. The mean D2:D4 digit ratio was 0.97 with a standard deviation of 0.04. This shows that the participants' digit ratios were very similar to each other.

Correlation between D2:D4 digit ration and BPAQ-SF scores
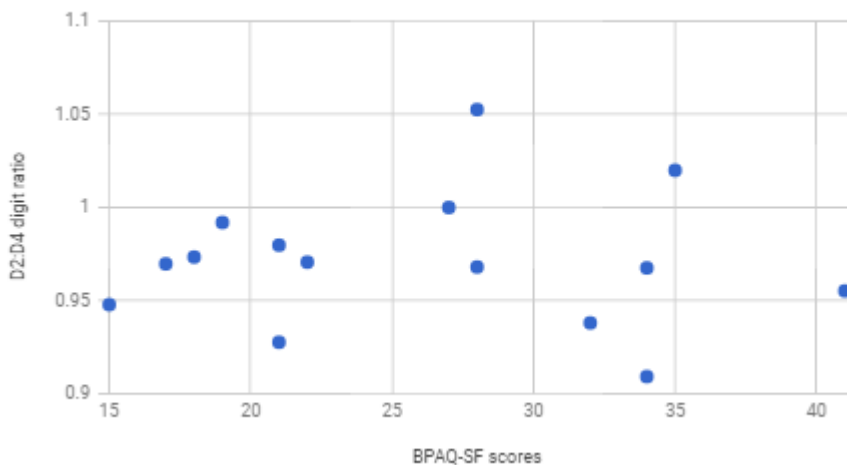
The scatterplot shows no correlation between D2:D4 digit ratio and aggression scores. To check this, Spearman's rank order correlation was used to calculate the correlation coefficient between the two variables. This is a suitable test because both variables (BFAQ scores and D2:D4 ratios) are at least ordinal level as they can be ranked in order. The observed value of R was -0.014. The critical value for significance (1 tailed $p < 0.05$, N=15) is 0.521. Since the observed value of R is less than the critical value the null hypothesis was accepted. There is no correlation between D2:D4 digit ratios and BPAQ-SF scores.

Discussion

These results suggest that prenatal exposure to testosterone does not influence aggression in women. This is consistent with other research that has found a relationship between D2:D4 digit ratio in men but not women (Bailey and Hurd, 2005). However, it is also possible that the failure to find a significant correlation in this study was due to a methodological problem.

The generalisability of this study is fairly low. The sample was small and very restricted by age and educational history. It may be that repeating the study with a larger, more representative sample would produce different results. If the study was repeated, a stratified sample of women of different ages and backgrounds could be obtained. This would make the findings more representative of the wider population and increase generalisability. In terms of ecological validity, the participants in this study were in their natural environment, so it is unlikely that the setting affected their behaviour.

The reliability of this study is likely to be high, because extraneous variables were well controlled. A standardised procedure and standardised instructions were used, so all the participants were told to fill in the questionnaire the same way. The researcher used a standardised method for measure digit length, but always measuring from the bottom of the digital crease to the fingertip whilst the fingers were extended. In addition, a standardised measure of aggression was used. All this means that the study could be replicated by other

researchers. In addition, the results of this study were consistent with other findings (e.g. Bailey and Hurd, 2005), which suggests that the result is reliable.

The objectivity of this study is moderate. The researcher gathering the data was aware of the aim of the study, which may raise the danger of researcher bias. In future, the data should be collected by a researcher who is unaware of the aim. This would mean that judgements of aggression or digit ratio would not be affected by bias. On the other hand, a standardised measure of aggression was used, so it was not necessary for the researchers to judge the participants' aggression levels. This makes the study more objective. There is a possibility of bias in the measurement of digit length because the method used lacked precision and the researcher had to judge the D2:D4 measurements, which could reduce objectivity.

The validity of this study is low. The BPAQ-SF may not be a valid measure of aggression in this sample. During completion, some participants asked for clarification of the meaning of some scale items. This would suggest that the questionnaire was not fully understood by all, which would affect the validity of aggression scores. The problem is that the BPAQ was created and standardised with people from the US, and uses some phrases that are not widely used in the UK. Future studies should use an aggression measure tested on UK respondents. In addition, the BPAQ-SF is a self-report questionnaire, and some participants might have under-reported their aggression due to social desirability bias. It might be better for future studies to use an alternative measure of aggression, for example, peer ratings.

In addition, the validity of the D2:D4 measurement is doubtful. The used of a ruler to measure digit length lacked precision, as measurements could only be made to within 0.5mm. This could have masked the actual variability in the D2:D4 measurements, meaning the correlation would not be apparent. In future, digit measurements could be made by using digital calipers or by scanning participants' hands and measuring digit length electronically. This would make the assessment of the correlation more valid.

The ethical acceptability of this study is high. All the participants were fully informed of relevant considerations when they gave consent and the fact that one participant chose to leave the study during the procedure suggests that the participants understood their rights before the study began. The participants were debriefed at the end of the study. It is unlikely that any elements of the study affected the participants' well-being.