Research in Autism Spectrum Disorders 57 (2019) 19–27

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

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

Some difficulties behind the concept of the 'Extreme male brain' in autism research. A theoretical review

Rosalind Ridley

Newnham College, Cambridge University, Sidgwick Avenue, Cambridge CB3 9DF, UK

ARTICLE INFO

Number of reviews completed is 2

Keywords:
Extreme male brain
Gender
Autism
Asperger's syndrome

ABSTRACT

The idea that autistic symptoms are produced by an 'extreme male brain' warrants further examination. The concept of the 'extreme male brain' derives from scores on the Autism Spectrum Questionnaire (AQ) i.e. it is defined behaviourally rather than anatomically. But if the concept of the 'extreme male brain' is to *explain* rather than *describe* autistic behaviour then evidence must come from a non-behavioural (e.g. physiological, biochemical or anatomical) source to avoid circularity of argument. The lack of a cognitive *intervening variable* linking autistic behavioural traits to brain activity casts doubt on the existence of a uni-dimensional spectrum of 'brain gender' for which autism is to be found at the 'extreme male' end. The inappropriate conflation of the *dependent variable* (score on the AQ) and the *independent variable* (brain anatomy of respondent) has led to the claim that a person with autism, even if female, has an 'extreme male brain'. This is comparable to the claim that, because on average men are taller than women, extremely tall women have 'extreme male height'. This stereotypical view of gender fails to recognize the overlapping diversity of cognitive styles found in males and females.

1. Introduction

In this article I want to explore the concept of the 'extreme male brain' and its usefulness in describing people with symptoms of autism. I will also consider the assumptions behind, and implications of, describing a person of either sex who displays supposedly male-typical behaviour, as having a 'male brain'. I will argue that the concept of the 'extreme male brain' raises complex issues about the meaning of 'extreme' in this context, about the meaning of 'male' when applied to persons, tissues or chemicals, and about whether there is any explanatory power in labelling 'behaviour' as 'brain'. It is only when we are clear about what the 'extreme male brain' is that we can evaluate the evidence intended to support or disprove the relationship of this 'extreme male brain' to autistic behaviour.[1]

Autism, Asperger's syndrome and the autism spectrum are important concepts in modern clinical and cognitive psychology. Starting from descriptions by Kanner (1943) and Asperger (1944) of children with social isolation and restricted cognitive interests, there have been many descriptions of people who exhibit this type of behaviour, and many hypotheses about the origins of this condition. Asperger (1944) was the first person to suggest that there is a relationship between autism and male-typical behaviour. This idea has been developed into the proposal that people with autism or Asperger's syndrome have an 'extreme male brain' (Baron-

E-mail address: rmr21@cam.ac.uk.

[1] I do not wish to consider here the evidence for or against sex differences in brain structure or function in the general (i.e. unselected) population since this is a substantial topic in its own right (e.g. Fine, 2010) or to consider in detail the evidence that the clinical features of autism arise from environmental effects on, or developmental changes in, brain and body structure since this has also been reviewed elsewhere (e.g. Teatero & Netley, 2013).

https://doi.org/10.1016/j.rasd.2018.09.007
Received 12 June 2018; Received in revised form 26 September 2018; Accepted 29 September 2018
Available online 09 October 2018
1750-9467/

Cohen & Hammer, 1997), that is to say, a brain which is both functionally and structurally male irrespective of the sex of the person concerned.

In order to assess the cognitive style in the general population that is associated with clinical autism, Baron-Cohen, Wheelwright, Skinner, Martin, and Clubley, (2001) developed the Autism Spectrum Quotient questionnaire (AQ) comprising 50 questions based on 5 cognitive domains: social skills, communication skills, imaginative flexibility, attention to detail and attention switching. These cognitive domains were chosen because of the cognitive features of autism (but excluding the motor abnormalities) that were described by Rutter (1978) and Wing and Gould (1979), although Baron-Cohen et al. (2001) stressed that the AQ is not an instrument for clinical assessment. Baron-Cohen et al. (2001) predicted and found that people with a diagnosis of high-functioning autism or Asperger's syndrome tended to score 'poor' in the domains of social and communication skills and imaginative flexibility, and 'good' in the domains of attention to detail and attentional focus. Subsequently, data on two separate questionnaires, the systemising quotient (SQ) (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003) and the empathising quotient (EQ) (Baron-Cohen & Wheelwright, 2004) were published as independent psychometric instruments which showed that people with high-functioning autism or Asperger's Syndrome tended to be high systemisers and low empathisers. These data, taken together, indicate a relationship between social and communication skills and imaginative flexibility in the AQ to the empathising skills of the EQ, and a relationship between attention to detail and attentional focus in the AQ to the systemizing skills of the SQ, although the overlap is neither theoretically or functionally total.

2. What is the 'extreme male brain'?

'The male brain is defined psychometrically as those individuals in whom systemising is significantly better than empathising, and the female brain is defined as the opposite cognitive profile. Using these definitions, autism can be considered as an extreme of the normal male profile. There is increasing psychological evidence for the extreme male brain theory of autism.' (Baron-Cohen, 2002).

Baron-Cohen (2002) defines 'systemising' as 'the drive to analyse the variables in a system, to derive the underlying rules that govern the behaviour of a system' while he defines 'empathising' as 'the drive to identify another person's emotions and thoughts, and to respond to these with an appropriate emotion'. Identification of these cognitive styles arose from the Autism Spectrum Quotient (AQ) (Baron-Cohen et al., 2001) in which answering 'agree' to questions about attentional style, which have a systemizing component (e.g. I am fascinated by dates) scores positively while answering 'disagree' scores null. Answering 'disagree' to social questions, which may be related to empathising skills, (e.g. I find social situations easy) scores positively while answering 'agree' scores null. Thus a high AQ score will be achieved by a person who has a particular interest and skill in systemizing together with a low interest in social or emotional interactions. According to the definition above (Baron-Cohen, 2002), a person with a high AQ score will have a cognitive style which can be described as 'extreme male' or 'autism-like'.

In the SQ questionnaire, 'strongly agree' and 'slightly agree' answers score 2 or 1 points respectively on pro-systemising questions and 'strongly disagree' and 'slightly disagree' answers score 2 or 1 points on negatively-couched systemising questions (e.g. I rarely read about new technology) such that a high systemizing person scores highly on this questionnaire. A comparable scoring system is used in the EQ questionnaire, which counts as positive 'agree' scores on pro-empathising questions together with 'disagree' scores on negatively-couched empathising questions (e.g. seeing people cry doesn't really upset me) such that a highly empathetic person scores highly on this questionnaire.

To ascertain whether such measurements can justify the use of the term 'extreme male brain' to describe or explain autism we need to consider the nature of the questionnaires, how a score on these scales can be linked with the brain and with sex, and finally to ask how the 'extreme male brain' theory of autism can be tested.

2.1. Problems with the AQ

First, I want to consider the theoretical relationship between any cognitive attribute and the ways in which it can be measured and so I will start with an non-specific quality which I will refer to only as a 'particular cognitive style'. This gives us the freedom to think about the relationship between a psychological variable and a psychological metric in general before applying that understanding to the relationship of the AQ to cognition.

2.2. Externally-defined and internally-validated metrics

In order to explore any 'particular cognitive style' we would need to be able to measure it - we would need a metric. Some human attributes can be measured according to an *externally-defined* metric; for example, you can measure people's height in centimetres. But we would have no externally-defined metric for a 'particular cognitive style' we would just have perceived the possible existence a particular cognitive style. We could invent a questionnaire to measure the 'particular cognitive style' in individuals and then describe the distribution of that attribute within the general population. This would be an *internally-validated* metric - one in which the questions reflect the originally perceived variable such that the people who strike us as having that particular cognitive style turn out to have higher scores on the metric than those whom we did not so regard.

As we shall see in more detail later, externally-defined and internally-validated metrics have different relationships to the thing they are measuring. For an externally-defined metric, such as height, the scores on the metric need not be normally distributed – the scores will just be whatever they are found to be, although in practice externally-defined biological variables often do form a normal distribution. The situation is different for an internally-validated metric. In making a questionnaire, we would include questions that

we think measure aspects of that particular cognitive style and we would then validate our rating scale against our population. We could attempt to construct the questions so that the number of people with each score forms a normal distribution. This would help us to envisage the magnitude of the effects we were measuring and to undertake appropriate statistical analysis. But we would not have *discovered* that there is a normal distribution of that particular cognitive style because we would have created the rating scale to make it so. On the internally-validated AQ questionnaire the statistical normality of the distribution of people's scores has rarely been tested in individual studies, but a meta-analysis of studies using this metric found that the 'best fit' of AQ scores for people (men and women) was not quite a normal distribution (Ruzich et al., 2015). But this does not mean that this cognitive attribute is not normally distributed in the population. It means that some questions could be changed to make the resultant distribution more nearly normal, thereby improving the validity of any parametric analysis that was subsequently undertaken.

Our internally-validated cognitive questionnaire would also not have added to our understanding of what a 'particular cognitive style' is because we would have based our original questions on what we thought that cognitive style was like in the first place. It may be, for example, that asking people whether they prefer red to green is a highly discriminating question because people with this particular cognitive style almost always prefer red to green. But unless we ask that question we will never know the answer. Questions that are included in a questionnaire are therefore likely to be confirmatory of what we think that cognitive style is like and will not extend the breadth of our understanding of what that cognitive style includes. This limits the ability of the questionnaire to fully *describe* a particular cognitive style.

In an externally-defined metric, the variable is often measured on an *interval scale* - the gaps between the numerical scores are all of equal size. Because height is measured on an interval scale, it is meaningful to say that a 200 cm tall man is *twice* as tall as a 100 cm tall child. An internally-validated metric is often measured on an *ordinal* scale in which higher scores indicate only a greater amount of the thing being measured. There are no independent units for measuring many cognitive attributes which would be equivalent to using centimetres to measure height, so differences in cognitive attributes cannot be quantified on an interval scale. A cognitive attribute metric can be used to show that one group of people is different from another, or to place individual people at a particular point on a spectrum, but in neither case can it say by how much. Self-report and interviewer-report questionnaires are used extensively in psychometrics and great pains are taken to validate them against the population to which they are applied. They can be an extremely valuable research tool so long as the implications of a internally-validated rather than externally-defined metric are clear.

The Autism Spectrum Quotient (AQ) and associated questionnaire-based metrics are intended to measure a kind of 'non-clinical autism-ness', or A-ness, although the terms autistic, Asperger's syndrome and autism spectrum disorder refer to *clinically* defined conditions. As discussed later in Section 5, AQ is not a clinical assessment, and while it may have good discriminative power when cases of ASD are compared to healthy controls (Allison, Auyeung, & Baron-Cohen, 2012) it was found to have no discriminative power in identifying individuals who would subsequently receive a clinical diagnosis of ASD, compared to those who did not, in a group of individuals who had already been referred for psychiatric assessment with a suspicion of having ASD (Ashwood et al., 2016).

2.3. Systemising and empathising

As previously described, the types of questions asked in the AQ are functionally related to (but not theoretically identical to) the systemising questions and empathising questions found in the SQ and EQ questionnaires (Baron-Cohen et al., 2003) (Baron-Cohen & Wheelwright, 2004) where the former describes a cognitive style that is systematic, mathematical and mechanistic while the latter describes a cognitive style that is more concerned with the feelings and motives of other people. Although, for the numerical reasons explained above, people with a *very* high score on the AQ will be *high* systemisers and *low* empathisers, there are people in the general population who are both *high* systematisers and *high* empathisers, and there are people who score *low* on both these dimensions (Baron-Cohen et al., 2003). This indicates that A-ness, is a *collective* term rather than a *single* cognitive variable. Nonetheless these authors (Baron-Cohen et al., 2003) conflate systemising and empathising scores within one dimension of A-ness. This is an odd thing to do. What is the justification for conflating strong systemising and limited empathising in the general population? It could be justified if there were a clear correlation between the scores on the two metrics but the correlation between systemising and empathising in the general population, though negative and statistically significant, was low in that study ($r = -0.16$, $p < 0.01$) (Baron-Cohen et al., 2003). Other studies have found conflicting results including the possibility that the correlation between systemising and empathising is, itself, sex-dependent (Valla et al., 2010) and, in a factor analysis of component autistic traits in the general population, social abilities and detailed non-social interests formed separate clusters (Palmer, Paton, Enticott, & Hohwy, 2015). Conflating parameters could be justified if those parameters were part of a cluster i.e. scores were more closely correlated with each other than with other components of the overall collection. Combining scores on physics, chemistry and biology tests, for example, and calling the result 'the science score' makes sense because they are all science subjects conceptually related by the scientific method and functionally related as a cluster if the correlation in student performance is greater within science subjects than across other combinations of subjects. But combining the scores on German and maths tests, for example, and calling the product 'the Germath score' would not be justified, even where a correlation exits based solely on talented students scoring highly on all tests, because the combination of only German and maths would be arbitrary selection and not indicative of an underlying common mechanism. In summary, conflating systemizing and empathising scores in neurotypical people would only be justified if scores on these parameters were more closely correlated (albeit negatively) with each other than with other measures of cognitive style, and this has not been shown to be the case in the general population.

On the other hand, in the clinical situation, the strong co-occurrence in autism of poor performance on certain social tasks (Baron-Cohen, Leslie, & Frith, 1985; Tager-Flusberg, 2007) and superior performance on certain non-social tasks (Jolliffe & Baron-Cohen,

1997; Shah & Frith, 1983, 1993) warrants the combination of these symptoms to form a clinical entity because this combination of cognitive styles is disproportionately common in clinical cases. Indeed, it is often the co-occurrence of symptoms in a clinical condition that usually do *not* occur together in the general population that is crucial to a diagnosis. For example, the unusual co-occurrence of a movement disorder and spider-vein blemishes in the skin is important in the diagnosis of ataxia telangiectasia. Furthermore, the unusual co-occurrence of normally unrelated features is often an important clue to the underlying pathological mechanism of the clinical condition. The difference in the relationship between empathising and systemising skills in the clinical and general populations suggests that the biological determiner (cause) of clinical autism is unrelated to the separate biological determiners of empathising and systemising in the general population. Elucidation of the pathogenic mechanism specific to autism should, eventually, separate this clinical entity from normal human diversity and open up the possibility of therapeutic intervention.

2.4. Self-assessment

Scores on the AQ, SQ and EQ questionnaires are based on self-opinion rather than being the result of tests of ability. As such, the results are not necessarily an accurate measure of a behavioural trait. I may think I am good at personnel management. My staff may think differently. If you wanted to assess a cognitive skill such as memory you would conduct experimental memory tests. You would not ask people to tell you how good their memory was. Furthermore, you cannot explain *why* a person behaves in a particular way by asking them *whether* they behave in that particular way. Answers based on self-assessment therefore lack reductive explanatory power - at least some sort of putative mechanism is needed to explain why something happens.

Bishop and Seltzer (2012) have suggested that people with autism-like symptoms may be differentially poor at accurate self-assessment. People who have already received a diagnosis of an autism-like condition may have different personal biases when answering these questionnaires. Self-assessment *of* females is, of course, always done *by* females (and self-assessment *of* males *by* males). Differences in gender-based social expectations may produce a bias in self-assessment, e.g. women may think they ought to appear 'compassionate', men may wish to present themselves as 'unsentimental' (though neither of these examples fully epitomises empathy). Furthermore, male and female assessors may have different normative values for certain words and phrases in the AQ, such as 'easy to make friends' or 'keeping a conversation going', producing a gendered assessor-bias which would necessarily be confounded with gender of respondent.

3. Linking AQ to brain needs the bridge of cognition

3.1. Is the 'Extreme male brain' a description or an explanation of A-ness?

Crucial to understanding what is meant by the 'extreme male brain' is an understanding of the distinction between a *description* and an *explanation*. Since behaviour is instantiated (embodied) in brain (i.e. behavioural actions and neural activity occur in tandem), the 'soft' meaning of 'extreme male brain' is merely that the brain is being extremely male when extremely male-typical behaviour is occurring. The 'hard' meaning of 'extreme male brain' is that some physiological influence, determined for example by the Y chromosome or by fetal or subsequent androgen levels, has produced a brain with an 'extreme-male' *structure* that is now determining 'extreme male-typical' *behaviour*. This is an interesting idea but it only becomes an explanation when experiments have shown that such brains do indeed have an 'extreme male' structure. I will argue below that, in constructing the concept of the 'extreme male brain', *behaviour* has been relabelled *brain* and that this has no explanatory power and very little descriptive usefulness. Furthermore, the concept of an 'extreme-male brain' depends on there being such a thing as a structurally 'male brain' and a structurally 'female brain' and, while there are differences in the average size of different brain structure in males and females (Ruigrok et al., 2014), these differences do not make any specific brain size, shape, region or structure definitively male or female.

Explanations are by definition at a lower level on the reductionist axis than the thing to be explained. Behaviour is explained by being 'reduced' to some sort of physiological mechanism. But the level at which the explanation operates cannot be too far below the level of the thing to be explained. You cannot explain politics directly in terms of physiology. Although reflex behaviours may be adequately explained directly in terms of the anatomy and physiology of the nervous system, more complex behaviours need to be explained in terms of cognitive mechanisms before these mechanisms are themselves explained in terms of brain activity. For example, the pattern of memory impairments in people with dementia needs to be analysed in terms of different types of memory storage and retrieval mechanisms, which are themselves instantiated in different brain areas, before the pattern of impairments can be related to damage to the brain. Similarly, in order to explain A-ness, we need to explain A-like behaviour in terms of specific cognitive processes, and then explain these cognitive processes in terms of the brain mechanisms that instantiate them. This requires the use of specific cognitive tests that generate predictions about that behavioural output. Cognition can therefore be described as an *intervening variable* between brain mechanisms and behaviours and I will consider this next.

3.2. Intervening variables

It is common in psychology to seek an intervening variable that will simplify the relationship between various behaviours on the one hand, and various physiological conditions on the other, the supposition being that a simple psychological or cognitive variable will map onto a simple brain mechanism. Hunger, for example, acts as an intervening variable that links many behavioural factors, such as eating, cooking, and shopping for food, with many physiological factors, such as body weight, metabolic rate, level of exercise, stomach contents and so on. More importantly, hunger maps on to neural activity in the lateral hypothalamus (Anand &

Brobeck, 1951). Similarly A-ness comprises a complex set of behaviours and is found in association with a variety of clinical conditions (or none) and a unifying cognitive mechanism is needed that will eventually map onto brain activity in order to explain A-ness. One candidate cognitive mechanism is 'Theory of Mind' since this can be measured by experimental tests and may be impaired in autism (Baron-Cohen et al., 1985) but this is unlikely to be the cognitive mechanism behind the non-social aspects of A-ness. The lack of a putative intervening variable that could explain all the symptoms of A-ness has led some researchers to question whether A-ness should be regarded as a single entity at all (Happé & Ronald, 2008; Happé, Ronald, & Plomin, 2006).

4. Can we identify a 'male brain'?

The behaviour and cognitive styles typical of men and women differ in many ways and a propensity towards 'systemising' and 'empathising' may not be the only, the most substantial, or the most frequent way in which the sexes behave and think differently. Furthermore, the various ways in which men and women differ may not co-vary so that being 'extreme' on one measure may not entail being 'extreme' on another measure. Even a robust statistical difference between men and women does not *define* the nature of male and female. In the West, females have a greater life expectancy than men, but we do not choose to define males as 'shorter-lived'. The difference in longevity is regarded as true but neither definitive or 'essentialistic' where 'essentialism' is the philosophical position which claims that things have a set of characteristics that makes them what they are. In the current context, for example, essentialism would claim that there is such a thing as 'maleness' and that men deviating from this maleness are not truly male, rather than that such variation merely makes maleness more diverse.

'Extreme height' means being deviant from normal or average height, i.e. being extremely tall, or extremely short, because height is an externally-defined metric, that is to say, height and the people whose height is being measured are different things. But 'maleness' can only be defined in relation to the archetypal or average male. In other words 'maleness' is an internally-validated metric which merely describes itself. The 'average male' is a sort of Platonic form, an ideal that may not exist, in itself, but from which all examples of actual males deviate to some extent. A man cannot be more male than the Platonic form of maleness any more than a single hand-drawn circle can be more circular than the Platonic perfect circle from which all examples of circles deviate. 'Extreme male' should therefore mean 'extremely male-like' but, in practice, the term 'extreme male' is rarely being used in that way. This confusion arises because of failure to appreciate the difference between externally-defined metrics where 'extreme' is far away from the norm, and internally-validated metrics where 'extreme' can only mean extremely normal or average.

In a meta-analysis of many studies, Ruzich et al. (2015) found that, on the AQ scale from 0 to 50, neurotypical females had an average AQ of ~15 and neurotypical males had an average AQ of ~18, a difference of 3 points; females with a diagnosis of an autism spectrum condition had an average AQ of ~39 and males with an autism spectrum condition had an average AQ of ~36, a difference of 3 points. In the neurotypical cases the difference between males and females was statistically significant; effect size Hedge's $g = 0.40$, $p < 0.001$, $z = 3.36$ and in the autism spectrum cases the difference was not significant. The lack of statistical difference between the sexes in the clinical sample may be because the clinical diagnosis constrained the diversity of that sample, subtly different criteria of diagnosis may have been applied to males and females, a 'ceiling effect' may be operating at the extreme of the questionnaire, and the sample size was smaller than the sample of neurotypical cases. In comparison, the difference between the neurotypical sample of both sexes (~17) and the sample of both sexes with an autism spectrum condition (~35) was ~18 points; effect size Hedge's $g = 2.86$, $p < 0.0001$, $z = 26.42$. A female with autism therefore does not resemble a neurotypical man, she resembles an autistic man. That's because she has autism, not because she has a 'male brain', 'extreme' or otherwise.

Adding 'brain' to 'extreme male' sounds reductionist because it appeals to the spurious notion that physiological attributes have more explanatory power than psychological attributes. If brain tissue were specifically male (or female) this would, potentially, be the way to exit from the circular arguments alluded to previously, but the maleness of brain would have to be definable and verifiable at an appropriate level on the reductionist axis. There are, potentially, several levels at which brain might be said to be male or female some of which are described in the next sub-sections. In neurotypical and psychotypical people, the various indicators of sex are not in conflict and the concept of sex is not challenged. But the suggestion that a female may have a male–brain, extreme or otherwise, requires clarity of definition of male/female before evidence for or against the concept can be adequately interpreted.

4.1. Genetics

At the *genetic* level, a male brain might be defined as one in which the cells carry a Y chromosome, but this is unhelpful in resolving complications when chromosomal and other genetic abnormalities are identified, or where genetics are at odds with behaviour or self-identified gender. There is no logical reason why either self-identification by a person or the genetic make-up of cells should be privileged, the one over the other, when describing the gender of functioning brains.

4.2. Anatomy

Can brain, at the *anatomical* level, be identified as male or female? Average brain size comparison in males and females is confounded by a difference in average body size in the two samples. In addition there may be differences in the average size of some brain regions in males and females (Ruigrok et al., 2014) but this doesn't make any particular size or shape of brain areas definitively male or female. Overall, the overlap between size and shape of brain areas across sex is so complicated that almost no one has a brain with characteristics confined to those of the typical values of one sex and most brains might be better described as a *mosaic* of cell groups some of which err towards the average male size and shape and some of which err towards the average female size and shape

irrespective of the sex (or gender) of the person concerned (Joel et al., 2015).

4.3. Endocrinology

At the *hormonal* level, is a brain male if it has been exposed to male-typical levels of androgens *in utero* or after birth? Testosterone is often referred to as a male hormone because it plays a crucial role in physical and behavioural development in men, but testosterone is also present in, and crucial for, neurotypical development and behaviour in women (Davison & Davis, 2003). It is the resulting *phenotype*, not the *hormone*, that is male or female. A brain exposed to male-typical levels of testosterone will be described as male if it occurs in a body that is genetically and phenotypically male. But it need not be the testosterone *per se* which made it male-typical in these circumstances because there are many other sex-determining effects in cells containing the Y chromosome. The brain of a woman who has been exposed to male-typical levels of testosterone will not be described as male unless she exhibits male-typical behaviour. On the other hand, a woman who exhibits male-typical behaviour may be *described* as having a male brain but without independent anatomical, or hormonally-based evidence this has no *explanatory* power over 'male-typical behaviour'.

4.4. Function

Can brain, at the level of a *functioning* organ, be defined as male or female? Since behaviour is instantiated (i.e. embodied) in the brain (and body) of the person who exhibits that behaviour, there is no explanatory power in saying that a person is acting in a male-like way because they have a male-like brain.

4.5. Personhood

At the *person* level, a male brain is one that is part of a person who is generally regarded as male. This is acceptable as a definition but has little explanatory power. It is also dependent on it being possible to define what a male person is and this can be problematic in unusual cases. It is particularly unhelpful in cases where a person who is chromosomally and anatomically male, and who exhibits male-typical cognitive skills and interests but who believes they are female, undergoes gender re-assignment and whose sexual orientation is towards men.

5. Testing the 'extreme male brain' hypothesis

5.1. Dependent and independent variables

In an experimental analysis, it is usual to measure a *dependent* variable (something that you think may vary) in relation to an *independent* variable - something different that varies, or can be varied, in an orderly fashion. If you were to compare height in men and women, the independent variable could be sex and the dependent variable would then be the heights of the people in the sample measured in, say, centimetres, perhaps displayed as two bar charts side by side. This would show that the *average* height of men is about 14 cm centimetres higher than the *average* height of women in the United Kingdom[2]. Alternatively you could plot a distribution curve in which height would the independent variable and the number of people of each height would then be counted to form the dependent variable. Two such distribution curves could be plotted, one for men and one for women. This would show that the *modal* height (the most frequent height) of men was a few centimetres higher than the *modal* height for women, that most of the tallest people were men and that most of the shortest people were women. Scientists interested in the characteristics of men and women would want to look at the bar chart, although they may want to check the characteristics of the distribution curve as well. A manufacturer of trousers would want to look at the distribution curves to work out how many trousers of each length to make in male and female fashions. In either case, the axes of the bar charts or of the distribution curve are orthogonal (i.e. the axes of the dependent and independent variable are arranged at 90° to each other).

The same relationship between variables applies to people and A-ness, where A-ness is comparable to height. Scores on the AQ could be plotted as bar charts for men and women, or the number of men and women with each possible score on the AQ could be plotted to produce distribution curves for men and women. But in both cases A-ness will be plotted orthogonally to sex. By re-naming scores on the AQ as 'male brain-ness', as Baron-Cohen (2002) has done, the dependent and independent variables have been conflated, and the relationship between A-ness and sex has been obscured. The variables are no longer orthogonal and we are no longer able to make non-tautological statements about what is going on. We can conclude only that, on average, men are more 'male-brained' than women. If we were to re-label extreme tallness as 'extreme male-height' we would conclude that men have, on average, more 'male-height' than women. This is a vacuous statement. If a taller than average man was described as having 'extreme male-height', then a woman of the same height would also have 'extreme male-height'. No, she wouldn't. She would just be an extremely tall woman. A woman who is taller than an average man would be more likely than a man of the same height to have a pathological condition e.g. Marfan syndrome, because she would be more deviant from the average height for women than a man of the same height would be deviant from the average man, but to suppose that this makes her height, in some way, 'male' is an essentialistic fallacy.

[2] www.worlddata.info/average-bodyheight.php

5.2. Measuring the 'extreme male brain'

Although the *definition* of A-ness as 'extreme male brain' prevents experimental assessment of the relationship between these two factors, we could regard the 'extreme male brain' as an *hypothesis* about the nature of the neurological basis of A-ness and then proceed with experimental investigation. Baron-Cohen and colleagues (Auyeung et al., 2009) claim that the sex of brain tissue depends irrevocably on its developmental history, specifically that exposure to prenatal testosterone renders a brain 'male' and that excess testosterone produces an 'extreme male brain' (see **3.3** above). There are at least three types of experimental design by which this could be tested, depending on the sample of people being assessed. A detailed review of the data yielded by experiments of these types is beyond the scope of the conceptual issues discussed here but three examples of these experimental designs are given below as an indication of the difficulties that may be encountered in the collection and interpretation of results.

5.2.1. Scores on the AQ could be correlated with a marker or measure of prenatal testosterone in a sample of people from the general population

Since pre-natal testosterone is greater in male fetuses than in female fetuses, and average scores on the AQ are somewhat higher in men than women, a positive correlation between prenatal testosterone and score on the AQ across both sexes may occur without pre-natal testosterone causing A-ness. Some other sex-related factor may be related both to fetal testosterone and to A-ness. Analyzing data from males and females separately can overcome this problem. Auyeung et al. (2009) reported a correlation between pre-natal testosterone, measured by amniocentesis, and autistic traits in children who had not had a prior diagnosis of clinical autism, when measured in each sex separately, as well as in a combined analysis. But in later studies, no correlation was found between pre-natal testosterone and parentally assessed measures of autistic traits in young children (Kung et al., 2016) or between testosterone levels in umbilical cord blood taken at birth and the subsequent AQ scores in either males or females assessed separately (Whitehouse et al., 2012).

In these examples, the participants were unlikely to have been drawn *randomly* from the general population because pregnancies that are subject to amniocentesis or other invasive procedures are likely to have already given rise to special concerns and the sample sizes were small. Nevertheless, the power of this type of analysis is that the participants were not selected because of any prior association with autism and this reduces potential *post-hoc* sampling biases.

5.2.2. Markers of pre-natal testosterone could be compared statistically between a group of people with extreme A-ness or autism and a group of neurotypical people from the general population

Higher scores on a composite measure of five hormones, including sex hormones (e.g. testosterone) and non-sex hormones (e.g. cortisol), were recorded from the amniotic fluid surrounding males who subsequently developed autism, Asperger's syndrome, or developmental disorders with certain autistic symptoms, compared to other males (Baron-Cohen et al., 2015). Testosterone, on its own, was not, however, elevated in males with these conditions compared to neurotypical males. Results for women with or without autism were not included in this study.

5.2.3. A sample of people with a clinical diagnosis of a syndrome involving exposure to high concentrations of testosterone pre-natally could be assessed for A-ness or autism later in life

Females with congenital adrenal hyperplasia, who had been exposed to extremely high levels of pre-natal testosterone, were found to have only a minor increase in scores on the AQ, from average female to average male scores; males with congenital adrenal hyperplasia did not differ in score on the AQ from neurotypical males (Knickmeyer et al., 2006), although this condition does not elevate testosterone levels in males *in utero*. In a second study, scores on other tests of autistic traits were similarly not found to be elevated in females (or males) with congenital adrenal hyperplasia (Kung et al., 2016). These results suggest that some other, non-testosterone, effect is driving abnormal development in autism. Even if *typical* pre-natal testosterone levels in a male fetus do contribute to the genesis of *typical* male characteristics in adult males [though the evidence for this has been critically reviewed (Grossi & Fine, 2012)], this does not mean that *excessive* fetal testosterone will produce *excessive* male characteristics in either men or women.

6. Clinical entities and normal variation

6.1. Clinical entities

When I refer to greater A-ness I am referring to the numerically high end of the AQ spectrum. Although the phrase 'being on the spectrum' is popularly used to refer to people with high but sub-clinical scores on the AQ, when a questionnaire is administered to a sample of the general population, everybody has a score that places them somewhere on that spectrum. There are difficulties, to which I shall return later, in describing the extreme end of *normal* variation as *clinical*. A taller than average person may have a diagnosable, pathological and potentially treatable condition such as Marfan syndrome, or they may just be extremely tall. Something other than an extreme position on any metric is needed to define a condition as clinical.

The distinction between extreme A-ness on the one hand, and clinical autism and Asperger's syndrome on the other, does not depend on the score on the AQ *per se*. Two people may achieve the same score on the AQ while only one warrants a clinical diagnosis. Making a clinical diagnosis requires that the condition be seen as deleterious to the person, or occasionally as deleterious to other people, as in some cases of serious psychosis. Strong legal safeguards are required when making a clinical diagnosis in a person who does not regard themselves as ill. Making a clinical diagnosis also presumes that there is a pathological cause at a lower reductive

level than the symptomatology. This poses a substantial difficulty for the 'extreme male brain' hypothesis even though Baron-Cohen is rightly clear that the AQ is not diagnostic of a clinical condition (Baron-Cohen et al., 2001). A person from the general population who gains a score of > 32 on the AQ may be regarded as rather unusual but lacking an underlying pathology, while a person with a score of > 32 and a clinical diagnosis of an autism-associated condition based on other diagnostic criteria, is presumed to have a neurological condition, even if the nature of that neurological condition is still unclear. Brain abnormalities which might occur in a clinical population cannot be assumed to exist in people who merely have a score > 32.

At present there are some reports of anatomical differences between the brains of autistic and non-autistic people (e.g. Hazlett et al., 2017) but it would be difficult to describe those differences as evidence of an 'extreme male brain' because, as described above, the anatomical nature of the 'male brain' and the 'female brain' has not been established.

6.2. Is sex relevant to autism?

Women, on average, score about 3 points lower on the AQ than do men (Ruzich et al., 2015). Does this mean that sex holds the clue to the nature of the biological mechanisms that lead to autistic behaviour? Sex influences almost everything about our physiological and mental constitution to the extent that it can be quite difficult to find a human variable on which sex does not have an influence. Against this background, the 3 point difference in mean score on the AQ between men and women in the general population, while true, might not be particularly specific or important. Men, on average, have greater food intake than do women but this does not mean that maleness is a major driver of over-eating. Similarly, many clinical conditions have a sex difference, with the number of people affected, severity of symptoms and mean age of onset showing a greater vulnerability in one sex, but this does not mean that sex holds the key to understanding the aetiology of the disease or warrant the disease being labelled as a disorder of sex. Parkinson's disease, for example, is a disorder of older age and is more common in men than women (Wooten, Currie, Bovbjerg, Lee, & Patrie, 2004) but it is not helpful to label Parkinson's disease as 'extreme male-aging', especially in cases when the disease occurs in women. The average score on the AQ for people (male and female) with a diagnosis of autism is ~18 points higher than the average score for the general population of men and women, whereas the average the score on the AQ for neurotypical men is only ~3 points higher than for neurotypical women (Ruzich et al., 2015). Whatever is driving the brains of some people towards autism, it would seem to be rather more powerful than the effect of sex on AQ scores in the general population.

6.3. Implications

The implication of this critical review of theory is that the concept of 'the extreme male brain' as a description or explanation of autism should be abandoned and replaced with the recognition that scores at any point on the AQ can be the product of a brain within either a male of female person. This allows the search for the biological basis, developmental causes and clinical correlates of autism to be broadened beyond that of differences between the sexes.

6.4. Conclusion

As was pointed out in Section 3, 'essentialism' is the philosophical position which claims that things have a set of characteristics that makes them what they are. In the current context, essentialism would claim that there is such a thing as 'maleness' and that men deviating from this maleness are not truly male, rather than that such variation merely makes maleness more diverse. It is an essentialist perspective that supposes that a highly systemising female computer programmer has an 'extreme male brain'. An essentialistic perspective might also claim that Shakespeare, who had an exceptional interest in, and understanding of, human psychological vulnerability, had an 'extreme female brain'. Identifying the systemising/empathising distinction as the 'essence' of the 'extreme male brain' is conceptually narrow. There are many ways in which males and females differ psychologically that are not captured by this dichotomy. Men and women may deviate widely in the extent to which they enjoy, for example, rough physical sports but the position of individual men and women on this spectrum may not match their score on the AQ, making it impossible to define what constitutes an 'extreme male' or an 'extreme female'.

The use of the term 'extreme male brain' is problematic both as a concept and as a reductionist explanation of autistic behaviour. The Autism Spectrum Questionnaire is an internally-validated metric, based on self-assessment, using an ordinal rating scale. It is useful in placing individuals within a spectrum of autistic behaviours but is limited in its power to identify previously unrecognised autistic traits. The somewhat higher scores of men compared to women on the Autism Spectrum Questionnaire does not warrant the attribution of 'extreme maleness' to the high end of the spectrum rather than to the average score for men i.e. the Platonic form of maleness. Although research has identified components of the autistic syndrome, for example, good systemising and limited empathising skills, there is currently no cohesive, autism-related cognitive intervening-variable that can be mapped onto a specific brain mechanism (Happé et al., 2006). Given that the brain and the body determine behaviour, the description of male-typical behaviour as being the product of a male brain is true but trivial in a male and conceptually problematic in a female. In neither case, does equating behaviour with brain have explanatory power.

The concepts involved in the borderland of neurology and psychiatry warrant particularly careful scrutiny because they concern things about which we understand very little. Of particular importance is the dissemination of ideas to the public, which includes people with, or concerned about, neurological and psychiatric conditions. Despite a greater awareness of the complexity of gender in recent times, many people are still essentialist and stereotypical in their thinking and attitudes. To equate systemising with maleness and empathising with femaleness encourages these views and does little to help us understand autism and related disorders.

Funding

The author received no direct funding but gratefully acknowledges the Library and Facilities of Cambridge University.

Conflict of interest

The author declares no conflict of interest.

References

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief “red flags” for autism screening: The Short Autism Spectrum Quotient and the Short Quantitative Checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*, 202–212.

Anand, B. K., & Brobeck, J. R. (1951). Hypothalamic control of food intake in rats and cats. *The Yale Journal of Biology and Medicine, 24*, 123–140.

Ashwood, K. L., Gillan, N., Horder, J., Hayward, H., Woodhouse, E., McEwen, F. S., et al. (2016). Predicting the diagnosis of autism in adults using the Autism-Spectrum Quotient (AQ) questionnaire. *Psychological Medicine, 46*, 2595–2604.

Asperger, H. (1944). Die ‘Autistischen psychopathen’ im kindesalter. *Archiv für Psychiatrie und Nervenkrankheiten, 117*, 76–136 Translated and annotated by U. Frith (1991) in *Autism and Asperger’s Sydrome* (ed. U. Frith) pp37–92. Cambridge University Press.

Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., & Hackett, G. (2009). Fetal testosterone and autistic traits. *British Journal of Psychology, 100*, 1–22.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Science, 6*, 248–254.

Baron-Cohen, S., & Hammer, J. (1997). Is autism an extreme form of the "male brain"? *Advances in Infancy Research, 11*, 193–217.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*, 163–175.

Baron-Cohen, S., Auyeung, B., Nørgaard-Pedersen, B., Hougaard, D. M., Abdallah, M. W., Melgaard, L., et al. (2015). Elevated fetal steroidogenic activity in autism. *Molecular Psychiatry, 20*, 369–376.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a ‘theory of mind’? *Cognition, 21*, 37–46.

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systematizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 358*, 361–374.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17.

Bishop, S. L., & Seltzer, M. M. (2012). Self-reported autism symptoms in adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*, 2354–2363.

Davison, S. L., & Davis, S. R. (2003). Androgens in women. *The Journal of Steroid Biochemistry and Molecular Biology, 85*, 363–366.

Fine, C. (2010). *Delusions of gender.* Icon Books Ltd.

Grossi, G., & Fine, C. (2012). The role of fetal testosterone in the development of the ‘essential difference’ between the sexes: Some essential issues. In R. Bluhm, A. J. Jacobson, & H. L. Maibom (Eds.). *Neurofeminism. Issues at the intersection of feminist theory and cognitive science* (pp. 73–104). Palgrave Macmillan.

Happé, F., & Ronald, A. (2008). The ‘Fractionable autism triad’: A review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychological Review, 18*, 287–304.

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience, 9*, 1218–1220.

Hazlett, H. C., Gu, H., Munsell, B. C., Kim, S. H., Styner, M., Wolff, J. J., et al. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature, 542*, 348–351.

Joel, D., Berman, Z., Tavor, I., Wexler, N., Gaber, O., Stein, Y., et al. (2015). Sex beyond the genitalia: The human brain mosaic. *Proceedings of the National Association of Science, 112*, 15468–15473.

Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? *Journal of Child Psychology and Psychiatry, 38*, 527–534.

Kanner, L. (1943). Autistic disturbances of affective contact. *The Nervous Child, 2*, 217–250.

Knickmeyer, R., Baron-Cohen, S., Fane, B. A., Wheelwright, S., Mathews, G. A., Conway, G. S., et al. (2006). Androgens and autistic traits: A study of individuals with congenital adrenal hyperplasia. *Hormones and Behavior, 50*, 148–153.

Kung, K. T. F., Spencer, D., Pasterski, V., Neufeld, S., Glover, V., O’Connor, T. G., et al. (2016). No relationship between prenatal androgen exposure and autistic traits: Convergent evidence from studies of children with congenital adrenal hyperplasia and of amniotic testosterone concentrations in typically developing children. *Journal of Child Psychology and Psychiatry, 57*, 1455–1462.

Palmer, C. J., Paton, B., Enticott, P. G., & Hohwy, J. (2015). ‘Subtypes’ in the presentation of autistic traits in the general adult population. *Journal of Autism and Developmental Disorders, 45*, 1291–1301.

Ruigrok, A. N. V., Salimi-Khorshidi, G., Lai, M.-C., Baron-Cohen, S., Lombardo, M. V., Tait, R. J., et al. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience and Biobehavioural Reviews, 39*, 34–50.

Rutter, M. (1978). Diagnosis and definition. In M. Rutter, & E. Schopler (Eds.). *Autism: A reappraisal of concepts and treatment* (pp. 1–25). New York: Plenum Press.

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., et al. (2015). Measuring autistic traits in the general population: A systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism, 6* 1-12. Erratum: Molecular Autism 6, 45.

Shah, A., & Frith, U. (1983). An islet of ability in autistic children: A research note. *Journal of Child Psychology and Psychiatry, 24*, 613–620.

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry, 34*, 1351–1364.

Tager-Flusberg, H. (2007). Evaluating the Theory-of-Mind hypothesis of autism. *Current Directions in Psychological Science, 16*, 311–315.

Teatero, M. L., & Netley, C. (2013). A critical review of the research on the extreme male brain theory and digit ratio (2D:4D). *Journal of Autism and Developmental Disorders, 43*, 2664–2676.

Valla, J. M., Ganzel, B. L., Yoder, K. J., Chen, G. M., Lyman, L. T., Sidari, A. P., Keller, A. E., Maendel, J. W., Perlman, J. E., Wong, S. K. L., & Belmonte, M. W. (2010). More than maths and mindreading: Sex differences in Empathizing/Systemizing covarainace. *Autism Research, 3*, 174–184.

Whitehouse, A. J. O., Mattes, E., Maybery, M. T., Dissanayake, C., Sawyer, M., Jones, R. M., et al. (2012). Perinatal testosterone exposure and autistic-like traits in the general population: A longitudinal pregnancy-cohort study. *Journal of Neurodevelopmental Disorders, 4* article 25.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders, 9*, 11–29.

Wooten, G. F., Currie, L. J., Bovbjerg, V. E., Lee, J. K., & Patrie, J. (2004). Are men at greater risk for Parkinson’s disease than women? *Journal of Neurology, Neurosurgery, and Psychiatry, 75*, 637–639.