Author response

Fetal testosterone and autistic traits: A response to three fascinating commentaries

Simon Baron-Cohen, Bonnie Auyeung*, Emma Ashwin and Rebecca Knickmeyer
Autism Research Centre, University of Cambridge, Cambridge, UK

This article is an author response to three previous commentaries on ‘Fetal testosterone and autistic traits’ (Auyeung et al., 2009).

Response to Ami Klin (2009)

We are delighted that Professor Klin finds the fetal testosterone (fT) theory of autism both ‘bold’ and ‘far-reaching’ and thank him for his positive commentary. Here, we simply respond to a few points at issue:

First, Klin asks ‘is a child who obsesses about electronics and builds radios . . . of cardboard really learning the principles of electronics?’ Expressed differently, Klin is asking if this is just ‘rote learning, rather than rule deriving’. We think this is a deep and insightful question, but would answer it by taking a much more pragmatic approach to systemizing: if the person knows how to make or operate a system, then they must have derived the rules of that system. They may not wish to generalize these principles to other systems, because a good systemizer does not go beyond the data in drawing conclusions. The good systemizer knows how this specific system operates, but cannot assume the same rules apply to other systems until these have also been learnt or tested. So, we do not presume that, whenever a child with autism plays obsessively with an object, they have necessarily learnt the rules that govern that system, but the hyper-systemizing theory at least respects the possibility that the child is learning the rules about that object (system). In earlier decades, obsessive and repetitive behaviour in autism were often dismissed as ‘purposeless’, but in the hyper-systemizing theory (Baron-Cohen, 2006) the assumption is that such behaviour can be highly purposive, in the service of systemizing.

Second, Klin finds it surprising that the research on fT and autistic traits has been slow to come from other groups. We agree that there is a need for independent...
replication of the results showing a link between fT and autistic traits, and would encourage any scientists with stored amniotic samples to attempt a replication. However, we do not find it surprising at all that replications have been slow to happen, not least because to collect amniotic samples requires huge effort and patience to navigate hospital ethics committees. Furthermore, because by definition such studies require a longitudinal, follow-up design, this means that even if another research group collected a sufficient number of amniotic fluid samples (with maternal consent), one might need to wait at least 5 years before one could measure autistic traits in the offspring. For these reasons, we do not expect many independent replications in the short term.

Finally, Klin argues that if confirmed, the fT theory of autism could have treatment implications. We agree that this is possible in principle and indeed testosterone blockers already exist so that medically it is also possible in practice. However, we would discourage research or clinical practice from attempting such treatment until it is known what the unwanted side-effects of such hormonal treatments might be. It concerns us that Geier and Geier (2007) for example are using testosterone-blockers on already diagnosed children with autism (blocking current testosterone) as we are not aware that this has passed through the relevant safety checks and the drug they are using (Lupron) is a form of chemical castration, usually used for treating adult sex offenders. As such it seems wholly inappropriate to use it on children with developmental disabilities.

The idea of a hormonal treatment *in utero* also carries potential risks that should not be swept under the carpet. If reducing fT reduced social difficulties alone, this might be considered to be an overall benefit to the child. If however reducing fT also reduced attention to detail and systemizing, this might be considered a negative outcome of treatment, since such cognitive characteristics have value in their own right. Finally, if reducing fT changed a child’s sexual orientation or gender identity (which are theoretical possibilities and need to be tested), this might be seen as ethically unacceptable. We know that Professor Klin would share our ethical and clinical concerns about unevaluated treatments, so this point is not a criticism of his point but is intended to keep such debates on the agenda within the biomedical community. Furthermore, this discussion is putting the cart before the horse. The first step is to move from testing if fT affects autistic traits (which we reported is the case in the target article) to testing if fT is elevated in categorically diagnosed cases of autism (which will need a much larger sample size to test, and which we are in the process of studying). We thank Professor Klin for his very positive commentary.

**Response to David Skuse (2009)**

We feel honoured that Professor Skuse has written a commentary on our target article as he is the leading scientist addressing the fascinating topic of why autism affects males more than females and we have deep respect for his work. Skuse raises the interesting question about whether the fT theory of autism and the X-linked imprinting theory of autism that he espouses are mutually exclusive, as explanations of the uneven sex ratio in autism. We think it is too early to know if one is correct and the other is not, and we can even see how both might be correct. This is not only because there are genes on the X chromosome that influence fT levels, but also because fT only accounts for about 20% of the variance in autistic traits, so there must be other contributory factors. We note however that a recent study found no evidence that X-linked genes contribute to individual risk for an ASC (Gong et al., 2008).
Skuse raises some concerns over the phenotypic measures we used. He is right that the Childhood Autism Spectrum Test (CAST) is not normally distributed - it detects autism spectrum conditions (ASC) rather than being truly dimensional, such that the majority of typically developing children score low (under 6 on this scale). This is precisely why we used a second, convergent measure, the Child Autism Spectrum Quotient (AQ-Child). Skuse says he cannot comment on this scale as it is not yet published and we are pleased to correct him on this point as it was published in the *Journal of Autism and Developmental Disorders* (see Au yeung, Baron-Cohen, Wheelwright, & Allison, 2008). The AQ-Child is much more normally distributed so overcomes many of the psychometric concerns raised by the CAST. Skuse seems to think that 9% of boys score above the cut-off of 30 + on the AQ-Child but in fact the relevant percentage is less than 7%. If one moves the cut-off to 32 + (as on the AQ-Adult and AQ-Adolescent) the percentage of males scoring above this drops to approximately 4 and 0%, respectively (Baron-Cohen, Hoekstra, Knickmeyer, & Wheelwright, 2006; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), and it raises the possibility that whilst the prevalence of ASC is about 1% (Baird *et al.*, 2006), there may be undiagnosed cases in the population too (Baron-Cohen *et al.*, in press (a)). Skuse raises some other concerns about the AQ as an instrument, such as that the scale was constructed so that males would be overrepresented among the high scorers. We would politely disagree with this as it is not true. The AQ was constructed to be gender-neutral and it is simply an empirical finding that more males score at the higher end and that there is a sex difference on average on this scale.

Most importantly, he rightly says that the fT theory would be bolstered if it was the case that the relationship between fT and autistic traits was found not just within the males but also within the females. We are again pleased to draw his attention to the finding in the target paper that this was indeed the case. fT predicts AQ-Child irrespective of one's sex (see Table 3, p. 12). Skuse’s penultimate paragraph contains a series of points which we found a little confusing but which seem to hinge on the accusation of circularity. We do not think there is anything circular about the finding that fT predicts AQ-Child, since empirically, it could have been otherwise.

Finally, Skuse asks if - as Ronald, Happe, and Plomin (2005) argue - autism involves at least two (if not three) independent dimensions, why we think there may be a single dimension (total AQ score). We agree with both him and Ronald *et al.* (2005) that autism is likely to involve at least two dimensions. Ronald *et al.* (2005) refer to these as social and non-social and we refer to these as empathizing and systemizing. In our factor analysis of the AQ we find a single factor (total AQ score) as well as sub factors, and consider it useful to continue to analyse our data in both uni- and multidimensional ways. Once again, we are grateful to Professor Skuse for contributing to this debate and we look forward to studies that test if the fT and imprinting theories are related or independent of one another.

**Response to Elise Barbeau, Adriana Mendrek and Laurent Mottron (2009)**

We were delighted to see that Professor Mottron and his colleagues had also provided a commentary on our target article, not least because his laboratory in Montreal was the inspiration for our recent study of atypical sensory processing in autism (Ashwin, Ashwin, Rhydderch, Howells, & Baron-Cohen, 2009) and we regard him as having pioneered studies into perception in autism. Barbeau, Mendrek, and Mottron's
commentary raises some good questions but we would like to take this opportunity to point out that we think it is based on some incorrect assumptions.

First, Barbeau et al. point out that one problem for the view that fT levels are associated with systemizing and autistic traits is the finding that mental rotation is not correlated to 2D:4D (Falter, Arroyo, & Davis, 2006; Hooven, Chabris, Ellison, & Kosslyn, 2004) and that mental rotation is not superior in congenital adrenal hyperplasia (CAH; Hines et al., 2003). However, this may not be a problem for the fT theory of autism. First, this is because Falter et al. (2006) use the 2D:4D measure as a measure of fT but this is both indirect and needs far more validation as an index of fT. Second, Hooven et al. (2004) only tested current testosterone (in saliva) rather than fT. And third, mental rotation may not be the relevant measure as it may be an index of ‘spatial skills’ (possibly associated with postnatal testosterone (Hines et al., 2003)). The Embedded Figures Test may be a purer test of the excellent attention to detail required for good systemizing and associated with autistic traits (Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983). Barbeau et al. cite Falter et al.’s (2008) study as evidence that does not fit the fT or extreme male brain (EMB) theories of autism, but for some reason do not cite the response to it (Knickmeyer, Baron-Cohen, Auyeung, & Ashwin, 2008) which challenges the use of mental rotation and urges caution regarding their use of 2D:4D as an index for prenatal testosterone exposure.

Second, Barbeau et al. argue that, because the Empathy and Systemizing Quotients (EQ and SQ) are correlated not only with sex but also with what degree students study, this must be evidence of an environmental component to EQ and SQ scores. Whilst we do not dispute that there is likely to be an environmental component to empathy and systemizing (as with most human behaviour), we think this particular finding should not necessarily be interpreted as evidence for environmental influences. What this finding showed was that EQ and SQ are better predictors than is sex of what one studies at university (Billington, Baron-Cohen, & Wheelwright, 2007). Whilst this could reflect an effect of training in a particular subject area on EQ and SQ score, it could equally reflect that individuals with particular EQ and SQ profiles are attracted to different fields. A longitudinal study from the earliest years could clarify the direction of causality.

Third, Barbeau et al. point out that EQ and SQ are only weakly inversely correlated so must have largely independent mechanisms. We would agree with this point, since the neural regions involved in empathy (Baron-Cohen et al., 1999; Frith & Frith, 2001) and systemizing (Billington, Baron-Cohen, & Bor, 2008; Ring et al., 1999) are likely to be many and largely non-overlapping. However, fT appears to be involved in both (Auyeung et al., 2006; Chapman et al., 2006), in opposite directions, a finding that cannot simply be ignored and which may have aetiological significance.

Fourth, Barbeau et al. suggest there are some tasks on which females do better and on which people with autism show an extreme of the female pattern. They cite attribution of intentions, claiming that ‘every empirical study to date has shown that autistic individuals across a wide age range are capable of understanding the intentions of other people’s actions’. This claim is one we find surprising, in that the well-established Strange Stories test by Happe (1994) is just one example that reports marked deficits in recognizing intentions, and there are numerous others (Castelli, Happe, Frith, & Frith, 2000; Phillips, Baron-Cohen, & Rutter, 1995, 1998). The fact that children with autism outperform controls on the Meltzoff intentionality task may reflect that these involve copying actions performed on objects and can be passed simply by observation and folk physics skills. The idea that people with autism have excellent emotion recognition skills, as these authors claim, is another surprising statement, given the long
history of research showing the opposite (Golan & Baron-Cohen, 2006; Hobson, 1986). The idea that children with autism have intact lexical knowledge and therefore must be an extreme of the female profile again is also a statement that we would challenge. As Mottron and colleagues would not dispute, classic autism (unlike Asperger Syndrome) is diagnosed on the basis of delays in language, and in the typical population, girls show earlier language development and larger vocabulary size than boys (Fenson et al., 1994; Lutchmaya, Baron-Cohen, & Raggatt, 2002; Maccoby & Jacklin, 1974). The fact that when children do finally acquire language they have good lexical knowledge is likely to reflect their rather atypical language acquisition strategies, perhaps collecting precise word meanings (using semantics) rather than decoding speakers’ intentions (using pragmatics). This has been demonstrated in word-learning experiments (Baron-Cohen, Baldwin, & Crowson, 1997).

Fifth, Barbeau et al. argue that, if FT is having its effects through lateralization, the FT theory of autism should predict better global than local processing. Professor Mottron’s group was one of the first to highlight the opposite perceptual profile to this, in autism, and our own findings support his (Jolliffe & Baron-Cohen, 1997) as do others (Shah & Frith, 1983). We think it is premature to state how FT has its effects in the human brain, and that the laterality theory of FT is not the only contender. It is of interest that our new research shows a positive correlation with FT levels and attention to detail as measured on the Embedded Figures Test, suggesting that FT is correlated with better processing of local details (Auyeung et al., 2008).

Sixth, Barbeau et al. ask, if FT is associated with cognitive characteristics, why not physical ones too? This is an excellent question and goes to the heart of future research in this area: is there any evidence of somatic hyper-masculinization in autism? We doubt that the relevant physical characteristics are likely to be athletic ability, as this may be more related to postnatal testosterone, but we are open-minded about whether FT affects other physical characteristics. We are interested by the recent findings (Ingudomnukul, Baron-Cohen, Wheelwright, & Knickmeyer, 2007) that women with autism (and their mothers) have higher rates of polycystic ovary syndrome (PCOS), which is known to be testosterone linked, for example. More research is needed into testosterone-related physical characteristics in autism.

Seventh, Barbeau et al. ask why we did not find an effect of FT on block design. This is a good question, as one interpretation of block design is that it involves segmentation, as does the Embedded Figures Test, on which we have found an FT effect (Auyeung et al., 2008). This needs more exploration, and in this kind of research there are many factors that can contribute to not seeing an FT effect, including statistical power.

Eighth, Barbeau et al. pose a forward-looking question: does the AQ correlate with neural anomalies in autism such as ratio of grey to white matter? We are pleased to report that this is a question we are currently testing using MRI and DTI in different samples. They also ask if other sex steroid hormones (such as oestriodial) might be relevant to predicting AQ and again, we are testing this at present, since we found FT only accounts for 20% the variance in AQ.

Ninth, Barbeau et al. find the anomaly of smaller amygdalae in girls with CAH to be at odds with the findings of enlarged amygdalae in autism, if both conditions involve elevated FT. However, CAH cannot be characterized simply as a disorder of elevated prenatal testosterone levels as prior to treatment such individuals also have a glucocorticoid deficiency. The reduced amygdala size is most likely to reflect the glucocorticoid deficiency as it is also observed in males with CAH, whose prenatal androgen exposure is thought to be in the typical male range. It is possible that the
effects of the glucocorticoid deficiency may outweigh the effects of testosterone on the amygdala. We would also refer them to the studies in autism showing either smaller or larger amygdalae (Abell et al., 1999; Aylward et al., 1999; Baron-Cohen et al., 2000; Howard et al., 2000; Nacewicz et al., 2006; Rojas et al., 2004; Salmond et al., 2005; Sparks et al., 2002) and suggest that more studies are needed into whether and how fT exerts effects on amygdala volume.

Finally, in their conclusion, Barbeau et al. argue that we have only found a correlation between autistic traits in a general population rather than a correlation between fT and autism. We agree with this conclusion, which restates what we did and what we found. As we pointed out in the target article, to test for a link between fT and autism, an amniotic sample size of literally thousands would be needed, given the rate of autism is 1% (Baird et al., 2006; Baron-Cohen et al., in press (a)). Because our amniotic sample size is only in the hundreds, we can only test for links between fT and autistic traits. That we found such a link is we suggest a huge step forwards. We are now testing the link between fT and categorically diagnosed autism by using the Danish Biobank, which does have sufficient sample size to test this hypothesis definitively.

We thank Professor Mottron and his colleagues for their willingness to debate these questions, not least because we see many potential ways to integrate his ‘enhanced perceptual functioning’ hypothesis of autism (Mottron, Dawson, Soulieres, Hubert, & Burack, 2006) with hyper-systemizing in autism (Baron-Cohen, Tavassoli, Ashwin, Ashwin, & Chakrabarti, in press (b)).

We conclude that the finding of a link between fT and autistic traits is consistent with a converging set of findings from this unique longitudinal project (fT inversely predicting eye-contact, vocabulary, empathy, and positively predicting systemizing, autistic traits, and embedded figures performance). Any one of these findings can be questioned, as is true of any study. But the way these findings line up together suggests this may be fruitful avenue for future research.

References


Received 13 November 2008; revised version received 13 November 2008