LETTER

When big data aren't the answer

Kelsey Perrykkad[a,1] and Jakob Hohwy[a]

In PNAS, Greenberg, et al. (1) use data collected using 4 surveys from over half a million people to support the Extreme Male Brain (EMB) theory of autism and the Empathizing–Systematizing (E-S) theory of sex differences. Large sample sizes are—all other things being equal—better than small sample sizes. However, the most serious criticisms of these 2 theories (see ref. 2) are not addressed by increasing the sample size.

The questionnaires used by this study were all developed with reference to autism, and are measuring not independent, but interrelated, constructs (3–6). Historically, it has been taken as a given that there is increased prevalence of autism in males. Autism has also been defined based largely on characteristic social difficulties (read: differences in empathizing) and restricted interests in highly patterned stimuli (read: systematized thinking). The Autism Spectrum Quotient (AQ) was developed in the context of these assumptions, and the original paper on AQ took it as reassuring that both high autistic traits, as measured by the AQ, and clinical diagnoses of autism were found to have the same gender trends (5). However, evidence suggests that females have been systematically underdiagnosed and may present with a different clinical profile to their male counterparts (7). This is understandably not reflected by the AQ, given that it was calibrated to fit with the male-biased symptomatology at the time of its conception. So, it is by virtue of its design that male groups have disproportionately high AQ scores.

The 3 other measures [Sensory Perception Quotient (SPQ), Empathy Quotient (EQ), and Systemizing Quotient (SQ)] were all developed and validated with reference to their expected relationship with the AQ in diagnosed autistic populations and in the general population, and thus inherit the AQ's foundational design properties. In the supplemental information of ref. 1, Greenberg et al. state that the short versions of the measures were developed "independently of autism." However, they are a subset of the longer questionnaires, so taking a representative subset of questions cannot justify the claimed independence from autism and the AQ. By design, SPQ correlates with AQ, EQ is anticorrelated with AQ, and SQ is correlated with AQ.

While Greenberg et al. (1) acknowledge concerns about the "risks of convergence across measures," they also claim that "these limitations are offset by. . .big data, an independent replication cohort, and. . .using multiple measures in the same cohorts." Here, we have argued that the associations between scores on these questionnaires (and the participants' sex) should not come as a surprise—in big or small cohorts. Their correlation should also not lead us to believe that autism should be defined by its maleness, or that maleness should be defined by its high systematizing and low empathizing scores. The underlying construct measured by each questionnaire is either the same or very highly correlated, and more prevalent in males by design. Thus, these measures beg the question (in the philosophical sense), and big data don't get us out of this trap. Because researchers can now run large studies online with relative ease, we should be mindful that bigger sample sizes are no substitute for better measures.

1 D. M. Greenberg, V. Warrier, C. Allison, S. Baron-Cohen, Testing the Empathizing–Systemizing theory of sex differences and the Extreme Male Brain theory of autism in half a million people. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12152–12157 (2018).

2 R. Ridley, Some difficulties behind the concept of the 'Extreme male brain' in autism research. A theoretical review. *Res. Autism Spectr. Disord.* **57**, 19–27 (2019).

3 S. Baron-Cohen, J. Richler, D. Bisarya, N. Gurunathan, S. Wheelwright, The systemizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358**, 361–374 (2003).

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

[a]Cognition and Philosophy Lab, Philosophy Department, School of Philosophical, Historical, and International Studies, Monash University, Clayton, VIC 3800, Australia

Author contributions: K.P. and J.H. wrote the paper.

The authors declare no conflict of interest.

Published under the PNAS license.

[1]To whom correspondence may be addressed. Email: Kelsey.perrykkad@monash.edu.

Published online July 3, 2019.

5 S. Baron-Cohen, S. Wheelwright, R. Skinner, J. Martin, E. Clubley, The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* **31**, 5–17 (2001).
6 T. Tavassoli, R. A. Hoekstra, S. Baron-Cohen, The Sensory Perception Quotient (SPQ): Development and validation of a new sensory questionnaire for adults with and without autism. *Mol. Autism* **5**, 29 (2014).
7 M. Kirkovski, P. G. Enticott, P. B. Fitzgerald, A review of the role of female gender in autism spectrum disorders. *J. Autism Dev. Disord.* **43**, 2584–2603 (2013).