

research requires a common language in order to facilitate discovery – I need to know that what I am measuring is equivalent to something of the same name measured in another country by another researcher at another time. Thus, instruments such as the ADI-R and ADOS (considered gold standards in ASD research) are vital. They also allow for meaningful comparison of groups – it is only when the cut-offs are applied rigidly that they become less useful, and this was not the case in our research.

However, a tremendous amount of heterogeneity exists among people with ASD (those who meet current categorical definitions), let alone those who have some symptoms but do not meet full diagnostic criteria (the broader autism phenotype). The challenge, then, is what to do with individuals who lie in different places along the spectra that comprise a standard ASD conceptualisation. As Fitzgerald rightly points out, it is those individuals who have independently learned or been supported to cope with their differences in a way that allows them to function in a ‘neurotypical’ society who are missed and excluded by the diagnostic categories that the ADOS and ADI-R conform to. However, simply because they appear to be functioning well does not mean that they are not experiencing difficulties – I agree with Fitzgerald. It was my experience conducting this research that, for many participants, it was actually because of the pressure of coping, or because they were not recognised as struggling, that many individuals got into difficult circumstances that precipitated the onset of psychosis or other serious mental health problems – an observation that is unsurprising for many clinicians, I’m sure.

The challenge for research and clinical practice, then, is to find a way to bridge the gap between rigid diagnostic categories and representative samples. This is a problem for psychiatry as a whole, not just those interested in certain conditions, which makes innovations such as the research domain criteria initiative from the National Institute for Mental Health so relevant and interesting.²

- 1 Larson FV, Wagner AP, Jones PB, Tantam D, Meng-Chuan L, Baron-Cohen S, et al. Psychosis in autism: comparison of the features of both conditions in a dually affected cohort. *Br J Psychiatry* 2017; **210**: 269–75.
- 2 Sanislow CA, Pine DS, Quinn KJ, Kozak MJ, Garvey MA, Heinssen RK, et al. Developing constructs for psychopathology research: research domain criteria. *J Abnorm Psychol*, 2010; **119**: 631–9.

Felicity Larson, PhD, Department of Psychology (Clinical Psychology Section), Frankland Building, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. Email: fxl437@bham.ac.uk

doi: 10.1192/bjp.211.2.117a

The evidence base and readability of Freeman *et al* on virtual reality for treating delusions

I commend Freeman *et al*¹ for their innovative research using virtual reality in the treatment of persecutory delusions. The rather ‘soft’ finding – probably valid, but hardly surprising – is: ‘Cognitive therapy using virtual reality could prove highly effective in treating delusions’.

I have concerns about the evidence base of the study, that is the design, the data, the sample and the statistical methodology. Each of these concerns interrogates the validity and reproducibility of the study.^{2,3}

First, the sample size is extremely small – 30 participants. The consequences of this include overestimates of effect size and low reproducibility of results.

Second, 27 of the 30 participants were unemployed. There is little point to any research if one cannot extrapolate from one’s sample to some broader reference population. For this, the sample should mimic the population in important ways. However, there is no discussion about the sampling, or the reference population to which extrapolation might be extended.

Third, neither the patients nor the researchers were masked to the randomisation allocations. This, surely, is a fundamental flaw of the experiment. A double-blind experiment should be used to ensure impartiality, and avoid bias, such as, for example, the Hawthorne effect.⁴

Fourth, the main outcomes – comparing the delusional conviction of the two groups at the beginning and end of testing, as well as their distress – were tested using ANCOVA. But in the results section the authors report: ‘For ratings of conviction in paranoia, a gradual reduction across the scenarios for the threat belief testing group can be seen, whereas the conviction scores remain stable in the exposure group’ (p. 64). This suggests that the two groups diverge over time, having different slopes, rather than the assumed homogeneous slopes in the ANCOVA model.

Fifth, the term ‘repeated measures mixed model’ covers a wide range of possible models, and leads one to expect a single model incorporating the repeated measures and random effects, not ten models as are presented in the online supplement. Further, none of the models is clearly articulated in mathematical form.

Sixth, there are no graphs to display the data or statistical results. Tay *et al*⁵ propose the use of graphical descriptives to enhance research rigour, especially in psychology.

It appears that the article is written on two levels. The introduction and method sections, describing participants, design and virtual reality, are clear and lucid. By contrast, the evidence base of the article, discussing the data, models, analysis and results, is almost unintelligible. Further, the small sample size, sampling bias, lack of randomisation masking, lack of model specification and lack of statistical graphics, seriously undermine the study.

The phrase ‘evidence-based research’ has become popular in psychology. Thus, it is incumbent on readers, authors and journal editors to ‘raise the bar’ and demand higher standards of the evidence base of research studies.

- 1 Freeman D, Bradley J, Antley A, Bourke E, De Weever N, Evans N, et al. Virtual reality in the treatment of persecutory delusions: randomised controlled experimental study testing how to reduce delusional conviction. *Br J Psychiatry* 2016; **209**: 62–7.
- 2 Gardenier JS, Resnik DB. The misuse of statistics: concepts, tools, and a research agenda. *Account Res* 2002; **9**: 65–74.
- 3 McNutt M. Reproducibility. *Science*, 2014; **343**: 229.
- 4 McCahey R, Warner J, Illiffe S, van Haselen R, Griffin M, Fisher P. The Hawthorne Effect: a randomised, controlled trial. *BMC Med Res Method* 2007; **7**: 30.
- 5 Tay L, Parrigon S, Huang Q, LeBreton JM. Graphical descriptives: a way to improve data transparency and methodological rigor in psychology. *Perspect Psychol Sci* 2016; **11**: 692–701.

John L. Fresen, Statistician, IT Services, University of Oxford, Oxford, UK. Email: john.fresen@it.ox.ac.uk

doi: 10.1192/bjp.211.2.118

Community treatment orders and capacity to consent

We welcome Newton-Howes & Ryan’s plea for a more restrictive use of community treatment orders (CTOs).¹ They have a heavy