Intelligence 46 (2014) 341–344

Contents lists available at ScienceDirect

Intelligence

Yes, but flaws remain

Richard J. Haier [a,*], Sherif Karama [b,c], Roberto Colom [d], Rex Jung [e], Wendy Johnson [f]

[a] School of Medicine (Emeritus), University of California, Irvine, CA, USA
[b] McConnell Brain Imaging Center, Montreal Neurological Institute, Montreal, Canada
[c] Douglas Mental Health University Institute, Department of Psychiatry, McGill University, Montreal, Canada
[d] Universidad Autónoma de Madrid-Fundación CIEN/Fundación Reina Sofía, Spain
[e] Department of Neurosurgery, University of New Mexico, Albuquerque, NM, USA
[f] MRC Centre for Cognitive Ageing and Cognitive Epidemiology and Department of Psychology, University of Edinburgh, UK

ARTICLE INFO

Article history:
Received 9 May 2014
Received in revised form 12 May 2014
Accepted 13 May 2014
Available online 12 June 2014

Keywords:
Intelligence
g-Factor
Factor analysis
Peer review
Brain imaging

ABSTRACT

Hampshire and Owen maintain that their original paper was flawless, but doubts remain about their factor analysis methods and related assumptions. Failure to cite relevant papers, poor sampling and restricted ranges also remain problematic for the definitive conclusions they drew. The editorial review process for investigating the serious issues we raised prior to publication in Neuron remains a mystery. We stand by the opinion expressed in our preview: the Hampshire et al. paper is an interesting but flawed exercise and their conclusions are not as definitive, or original, as they believe.

Contents

Hampshire and Owen appear eager to critique our alleged beliefs rather than clarify important factual aspects of their paper. Their comments misrepresent our concerns and misdirect attention from the points we raised:

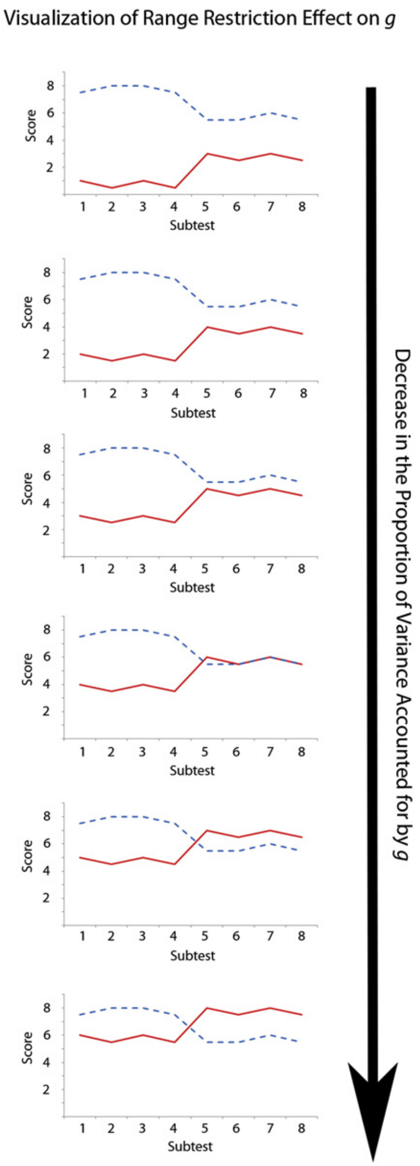
1. The key point of our critique is that the method of factor analysis Hampshire et al. (Hampshire, Highfield, Parkin, & Owen, 2012) chose forces independent factors, so it is possible that the conclusion reached is an inevitable consequence of the method rather than a new discovery about the relations among test scores and brain activities that have implications for understanding the nature of the *g*-factor. Other commentators have expressed similar concerns to ours on this issue (Ashton, Lee, & Visser, 2014a, 2014b, 2014c; Hampshire, Parkin, Highfield, & Owen, 2014a, 2014b). In their response to our comment, Hampshire and Owen maintain their view that one method of factor analysis, if carefully chosen, can reveal a "ground structure" in the data that has only one reasonable interpretation, even if that interpretation contradicts much other evidence. We do not find their rationale a compelling exception to the general understanding that factor analysis solutions are always based on assumptions that, at best, indicate various possible interpretations. We believe this could be clarified if Hampshire and Owen reanalyzed their data using oblique

* Corresponding author.
E-mail address: rich.haier@gmail.com (R.J. Haier).

http://dx.doi.org/10.1016/j.intell.2014.05.010

rotations in the factor analyses of both the behavioral and brain data. That way, the question of whether the resulting factors are correlated is addressed empirically rather than being relegated to assumptions.

2. In a related issue, Hampshire and Owen write: "… if individual differences in ability were driven entirely by spatially 'diffuse factors', then it would be highly unlikely that the task-network loadings and the task-behavioral component loadings would correlate." Referring to this, they add: "… this relationship provides relatively unambiguous evidence in support of our selection of rotated factor orientations." Finding that task-network loading and task-behavioral components correlate is not a proof that the factor solution is the best possible one. Oblique rotations (in both imaging and behavioral datasets) could have produced an ever better fit between task-network loadings and task-behavioral component loadings.



3. Apparently, the original three Neuron reviewers felt that a paper seeking to explain individual differences in the *g*-factor did not need to describe the subjects with respect to age or sex or, critically, the range of mental ability in the 16 "young healthy" subjects who participated in the imaging study at Cambridge University. In their response, Hampshire and Owen downplay this problem, but the small sample they used could have substantially affected their ability to reveal neural underpinnings compatible with the existence of a general factor of intelligence. This is not simply a secondary point to be addressed in future research as Hampshire and Owen suggested. Ignoring the effects of sample selectivity is a classic error in psychometrics. For example, Spearman, using data from a general population, found a clear general factor. Thurstone, using data from a highly educated population, did not (Thurstone, 1938). When Thurstone subsequently examined data from a representative sample of the population, he found a *g*-factor and conceded the discrepancy was due to range restriction in his educated sample (Thurstone & Thurstone, 1941). A range restriction effect in Hampshire et al.'s work would likely impact both behavioral and brain data and could give the spurious appearance of independent factors. Fig. 1 shows how range restriction can affect the ability to detect *g*.

4. Another crucial issue that was disregarded in Hampshire and Owen's response is the need for a test that provides a measure of cognitive ability that correlates highly with the gold standards in the field. The fact that their composite test battery correlated 0.65 with the Cattell Culture Fair Test of Intelligence is not a demonstration that their battery is a valid measure of *g*. Not only is 0.65 a low correlation for confirming the validity of a test, but the Cattell test itself, due to its restricted sampling of abilities, is not as good a measure of *g* as previously thought (Johnson, te Nijenhuis, & Bouchard, 2008). To draw any conclusions about *g*, one needs appropriate samples, a demonstration that one is using a valid test of general cognitive ability, and a sample

Fig. 1. In this illustration, the y-axis shows scores on 8 cognitive ability tests (each represented on the x-axis). The first 4 tests tap one cognitive domain; the last 4 tap another. Although only two subjects (dotted blue; solid red) are shown for clarity, each represents many subjects with similar kinds of relative ability patterns at different overall levels. The dotted blue subject represents a subsample of the highest scoring subjects on all the tests. Starting with the top graph and moving down, each graph represents a situation where a greater proportion of the lowest-performing subjects have been removed and a 'new' solid red subject is selected as representative of the remaining lower-scoring individuals. In the top graph, there is a clear 'general intelligence' effect where the dotted blue subject performs better than the solid red subject on all subtests. Further, certain degrees of specific-domain abilities are observed (the dotted blue subject has relatively stronger performance in the first domain while the solid red subject has relatively stronger performance in the second domain). As range restriction becomes more intense from top to bottom (i.e. more lower performing subjects are removed from analysis), the general factor accounts for decreasing proportions of the variance in their performances. In the bottom graph, g is no longer observable and the two subjects differ only in specific abilities. The bottom graph essentially represents a situation where range restriction on g is complete (i.e. all subjects have the same level of g). Because, in this extreme situation, there is essentially no variance in g, such variance cannot be observed by any statistical analysis. Which graph fits best with Hampshire et al.'s samples is not clear but it appears almost certainly not to be the top graph. This is why their failure to provide the test means and standard deviations for their samples is a critical omission.