The Emperor Has No Clothes

A Review of

*Reliability in Cognitive Neuroscience: A Meta-Meta-Analysis*

by William R. Uttal


http://dx.doi.org/10.1037/a0037026

Reviewed by

Harry A. Whitaker

William Uttal begins *Reliability in Cognitive Neuroscience: A Meta-Meta-Analysis* with a thoughtful short chapter explaining the raison d’être for meta-analyses of a body of experimental data, in this case focusing on the cognitive neuroscience literature. Those who are familiar with modern statistical methods will understand; those not-so-well-versed neuroscientists would do well to review carefully the statistical issues. Uttal’s next statistical step into meta-meta-analysis is a bit controversial, although no one could fault his goal, which is to attempt to reduce the variability in the data from brain-imaging experiments. He provides a clear exposition of the methodology of meta-meta-analysis in the second chapter of this five-chapter book. Next is the heart of Uttal’s book, a selective but fair discussion and review of contemporary neuroscience studies that use a variety of cognitive tasks.

Considering the amount of this research as well as the media attention it has garnered, the book’s conclusion is disappointing. Meta-analysis and meta-meta-analysis increase the variability, leading Uttal to conclude that the answer to the question implicit in the title of his book is, no, the data are unreliable. Uttal’s point should be taken seriously by anyone engaged in research that uses changes in brain images as the dependent variable in a cognitive experiment. Uttal frames his question thus: “A key question asked throughout this book is: are nodes (i.e. activation peaks) real or merely artifacts of our . . . method?” (p. 195).

Section 3.2.1 of *Reliability in Cognitive Neuroscience* (pp. 84–91) introduces an aspect of cognitive neuroscience research in particular, and psychological research in general, that could easily have been given much more space: individual differences. Uttal discusses some recent studies that directly assess intersubject and intrasubject variability; the intrasubject variability is worrisome, but the more pronounced intersubject variability raises the question, what’s wrong?

Chapters 4 and 5 address this question from a variety of theoretical and methodological perspectives, some more persuasive than others. For example, pointing out that we do not have the means to integrate data from all of the neurons in the brain at the same time is true enough but not very interesting. On the other hand, Uttal’s discussion of whether the
independent variables in the typical cognitive neuroscience experiment are well conceived is
definitely on target. He cites a nice article by Poldrack (2006) in this regard. Cognitive
science has made a substantial contribution to brain-imaging research.

Optimistically, as the theoretical constructs derived from psychology and linguistics improve,
many of us would expect the data from brain-imaging studies to make more sense. Uttal is
pessimistic that this will happen; given the current state of the field, his unenthusiastic
outlook is excusable.

Missing from Uttal’s panoply of troublesome issues that contribute to the unreliability of
research in cognitive neuroscience are at least two methodological problems that he does
not analyze; both, in fact, support his conclusions. Were these two problems to be
adequately addressed, it is conceivable that one might have an account of some of the root
causes of the variability in brain-imaging data.

The first problem is the observation that the geographic location of an activation node (in
this case, evoked potentials) varies according to the emotional state of the subject. In a
psychophysical paradigm, a single subject viewed a series of business logos with letters or
words, for example, the well-known red-and-white Coca-Cola disk, and passively named
them as they randomly appeared on-screen. The three conditions were being in (a) a happy
mood, (b) an angry mood, and (c) a neutral mood. Over a series of 500–600 presentations
in each mood state, the centers of electroencephalographic activity at 250 msec after
stimulus onset varied systematically within and across the left and right hemisphere (Yockey
& Whitaker, 2012). The simple message is that a subject’s emotional state alters where
(and probably how) his or her brain processes simple symbolic visual information. Note that
the majority of studies in cognitive neuroscience do not mention whether subjects’
emotional state was controlled for, let alone even assessed.

The second problem was pointed out nearly two decades ago (Whitaker & Hochman, 1995).
The pictures derived from functional magnetic resonance imaging (fMRI) studies are
commonly described as showing some aspect of the mind in action. It is said that such
images reveal where mental arithmetic occurs, where words are formed, where lies are
generated, perhaps even where religious sentiments are located. What, of course, is actually
seen is a computer-generated image of changes in the oxygen concentration in the blood of
some veins within the venous network that are presumably draining tissue regions in the
brain that have recently been active (and, by the way, whose distance or geometrical
relationship to the actual site of neuronal activation is unknown).

What are usually discussed in the conclusions of fMRI- or positron-emission-tomography
(PET)-based imaging studies are the “hot spots”—activity that is above a cut-off threshold.
However, the nodes of activation, which are the dependent measures in most brain-imaging
studies, are only half of the data obtained in these experiments. A more comprehensive
analysis shows that typically many areas of the brain are simultaneously activated at low,
subthreshold levels; surprisingly, these “cold spots” are statistically correlated with the
independent variables of the experiment. Therefore, there is no statistical reason why the
hot spot data are any more germane to the hypothesis than are the cold spot data; such a
determination can be made only by extraexperimental, a priori information.

Whether the overall reliability of cognitive neuroscience experiments would be improved, or
whether the variability that plagues interpretations of these experiments would be reduced
if the cold spot data were to be incorporated into the interpretation of the results, is of
course, moot (Whitaker & Hochman, 1995). The obvious suggestion is to follow Uttal’s recommendations: better theories, better methods.

Uttal is well known for a series of moderate-sized monographs, starting in 1989, that are critical of modern cognitive science, particularly cognitive neuroscience. His 2001 *The New Phrenology* was one of the more popular of these essays. It is reasonable to ask to what degree his critiques have been taken notice of in mainstream cognitive neuroscience journals. A cursory Internet search in three of the most highly regarded journals in this field will disappoint those of us who are convinced that the critical issues in *Reliability in Cognitive Neuroscience* should be carefully addressed. *Human Brain Mapping* was launched in 1993; there was a citation to Uttal’s criticisms in a 2013 article. *NeuroImage* was launched in 1992; there have been eight articles mentioning Uttal’s books, published between 2003 (one citation) and 2012 (one citation). Three articles in the *Journal of Cognitive Neuroscience* between 2001 and 2008 cite one of more of his books.

As Uttal himself points out, on the face of it, the cognitive neuroscience paradigm is flourishing, without, it seems, much concern about the linked problems of variability and reliability. A notable exception is Michael Miller, whose laboratory group has been studying individual differences for over a decade (Miller et al., 2002). Reading *Reliability in Cognitive Neuroscience* will convince you that too much variability and not enough reliability are to be taken seriously.

**References**

