Cognitive Symptoms in US Government Personnel in Cuba: The mending is worse than the hole

Sergio Della Sala\textsuperscript{a}, Robert D. McIntosh\textsuperscript{a}, Roberto Cubelli\textsuperscript{b},
Jason A. Kacmarski\textsuperscript{c}, Holly M. Miskey\textsuperscript{d} & Robert D. Shura\textsuperscript{d}

\textsuperscript{a}) Human Cognitive Neuroscience, Psychology, University of Edinburgh, Edinburgh, UK
\textsuperscript{b}) Department of Psychology and Cognitive Sciences, University of Trento, Rovereto, Italy
\textsuperscript{c}) Health Psychology Section, Veterans Affairs Eastern Colorado Health Care System, Denver, CO, USA
\textsuperscript{d}) Mental Health & Behavioral Science Service Line, Salisbury Veterans Affairs Medical Center, Salisbury, North Carolina, USA

Disclaimer: The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the U.S. Government.
A recent paper published in JAMA reported on a case-series of 21 people with cognitive complaints compatible with post-concussive symptoms (Swanson et al., 2018). These individuals were US government personnel who had been working in Havana, Cuba, and the proposed cause of their symptoms was unusual auditory phenomena of an unknown origin. Percentile scores from a battery of 37 neuropsychological tests were provided for six patients in the study, and each patient was classified as cognitively impaired if scoring below the 40th percentile on at least one of these tests. We have strongly criticised these methods as inconsistent with any recognised practice for evidence-based neuropsychology (Della Sala & Cubelli, 2018; Shura, Kacmarski, & Miskey, 2018), and with statistical logic (Della Sala & McIntosh, 2018). We are not surprised that Swanson and colleagues found that all six patients were impaired: when using the 40th percentile, 40% of people will fail each test and the chances of anyone passing 37 tests without an impairment being diagnosed are vanishingly small.

Remarkably, in a subsequent rebuttal to this and other criticisms, the authors did not defend their idiosyncratic choice of a 40th percentile threshold, but implied that they used some other criterion, as follows: “Within-individual deviations from an average performance are considered signs of brain dysfunction. Percentile scores in our report showed that all impaired patients had several scores that deviated by more than 1 SD from their respective means, some exceeding 2 SDs, which translates to more than 40 percentile points below their means (below 10th percentile relative to their average performance). This meets standard criteria for neuropsychological impairment…” (Hampton, Swanson, & Smith, 2018, p. 604). We are unsure what this confusing, ambiguous statement means. We can imagine two possible interpretations, each of which is problematic, and we consider them briefly in turn.
Hampton and colleagues explicitly refer to “within-individual” deviations from “their respective means,” suggesting that they calculated each individual’s mean (percentile) score across the 37 tests, and then expressed each test score in terms of how many standard deviations it fell from this mean (i.e., a within-subject z-score). Within-subject profiling can be a valid approach to understanding variations in a person’s abilities across domains, but such profiles can only be evaluated by reference to the distribution of between-test differences in a normative sample (e.g., Binder, 2009; Oakes et al, 2013). The within-subject approach that Hampton and colleagues’ statement implies has no normative point of reference, and would be meaningless.

As noted by Oakes et al. (2013), people with higher intellectual ability, as one may suspect the government employees were by the virtue of their positions and education, generally demonstrate greater variability in their abilities: “Individuals with high intellectual ability show much more variability in their performance than individuals with average function” (pp. 141-142). Even without these interpretive problems, a within-subject z-score would only be meaningful to calculate if the within-subject distribution of test scores satisfies the parametric assumption of normality, but this assumption was violated in four of the six individuals reported by Swanson and colleagues (2018). Regardless, the relevant z-scores are not reported, nor indeed are they mentioned, in Swanson and colleagues' original paper.

A simpler interpretation of their statement, though less compatible with the language of “within-individual deviations”, would be that Hampton and colleagues are implying that the actual cut-off per-test was 1 standard deviation below the normative mean. This would be equivalent to a lower cut-off at the 16th percentile or a T-score of 40. While this too is a liberal cut-off, it is not entirely without precedent (see Strauss et al, 2006, p. 91). Further, a 2 standard deviation cut-off would be approximately equivalent to the 2nd percentile, and a much less
controversial standard for impairment (although it should not be applied across multiple tests without considering the number of test scores examined). By this latter criterion, one of the patients tested (case 6), did perform more poorly than would be expected by chance, scoring at or below the second percentile on six of 37 tests (exact binomial test, \( p < .0001 \)), but the other five performed well within normal limits (case 2 failed two tests, and no other patient failed any).

An old Venetian saying seems very apt here: “Xe pèso el tacòn del buso” - the mending is worse than the hole. Hampton and colleagues have tried to patch over an unjustifiable threshold for impairment reported in the original paper with an even less cogent statement of their actual criterion in the rebuttal. Only two things are clear in this murky matter: that the crucial criterion for cognitive impairment was mis-stated in the original paper, and that the neuropsychological data presented do not support the conclusion that whatever happened in Cuba resulted in persistent cognitive decline.

It is imperative that neuropsychology as a field be presented in an accurate manner, especially in such influential and respected outlets as JAMA. From a clinical standpoint, misdiagnosing patients as impaired can even have iatrogenic effects in which the patients can come to behave in ways consistent with the alleged impairment. Researchers have an ethical responsibility to avoid misdiagnosis, as well as a scientific duty to be clear and accurate. We contend that Swanson et al. (2018) have fallen short on both counts.

References


