Usage of the MMSE obscures the science of dementia

Eugen Tarnow*

Avalon Business Systems, Inc. USA

Abstract

The 12th World Congress on Controversies in Neurology (CONy 2018) is contemplating the topic: "The term Alzheimer's disease should be dropped as it is impeding future research". This article suggests discarding the MMSE might be similarly helpful.

Introduction

The 12th World Congress on Controversies in Neurology (CONy 2018) is contemplating the topic: "The term Alzheimer's disease should be dropped as it is impeding future research". This article considers whether removing the use of the Mini-Mental State Exam (MMSE) would similarly be a good idea.

There is a tendency in clinical practice, due to a combination of the difficulty of doing the research, strong financial interests, and personal credit needs, to never settle scientific issues. One of these unsettled issues is the use of the MMSE. A Google search on ClinicalTrials.gov shows that it is used in more than 1200 clinical trials of Alzheimer's disease, yet there are critics [1-3].

The MMSE was created in an evening and consists of a laundry list of test items. The scoring is ill defined [4] there are no alternate forms, and test-retest reliability for the total score is not great – 5 points out of 30 and the standard error goes to infinity above a score of 24 [1]. The test-retest reliability of individual items is much worse because errors tend to cancel out in the total score.

That leads to the question – is there anything special about the MMSE? Why not just use a test, for example, in which the subject is asked to assemble a piece of IKEA furniture? As a patient descends into dementia the ability to put together the furniture will diminish just like the MMSE score. It tests hand-eye coordination, spatial thinking, ability to read and understand instructions, etc, and it certainly is an important measure of daily living for many of us. (And the provider can get all his furniture put together for free.) And, need I remind the reader, the inventor of these instructions is one of the richest men in the world – you can't argue with success! And, presumably, the instructions had much more standardization in them and took a lot longer to construct than an evening (or did they?).

There is no good answer why the MMSE should be preferred over the IKEA assembly because there is no good reason for using the MMSE. Drug companies use it, presumably, as a way to play roulette: perhaps the drug in question might show a slight MMSE edge on the twentieth trial (the only one published)? And the FDA plays along. Providers use the MMSE because it is quick (why spend time trying to figure out whether somebody is suffering from dementia) and almost free. If drug companies actually believed in their drugs and had some humility surely, they would use a highly reliable test of the very specific brain function they think the drug can fix rather than

the hodge-podge result of a fun evening? Perhaps they are expecting to hit a jackpot so large that it cancels out all the noise present in the MMSE and seemingly assume that all the knowledge destroyed by the disease would be miraculously added back into the brain if the neural structures started to repair themselves.

Consider a classic experiment, the word free recall experiments of Murdock [5]. In these experiments fifteen psychology college student subjects are presented with 80-word lists of ten words each. The words are read out using a metronome and exactly two seconds pass between each word. The words were selected from the 4000 most common English words [6]. Even under these highly standardized circumstances, it takes ten lists to reach an internally consistent result. I plotted Cronbach's alpha, a measure of internal consistency as the number of subtests are increased in Figure 1; an "acceptable" alpha is 0.8 or higher. In other words, college student subjects respond to the test items somewhat differently and in order to rate their memory for a standard set of words it takes a total of 70 words under perfect experimental conditions. Notice that for only two subtests (twenty words) Cronbach's alpha is a paltry 0.15.

What if we simplify the test items by removing item meaning to try to remove individual differences? I constructed the Tarnow Unchunkable Test, consisting of subtests of three double-integer test items designed to be unchunkable – i.e. to have no special relationships The test was administered to 500 Russian college students from a teaching college near Moscow [7] and, in a different experiment, elderly subjects entering a New York memory clinic [8]. The corresponding Cronbach's alphas are displayed in Figure 2. This shows that it is very hard to separate out college students, easier to separate out old people coming into a memory clinic where the range of performance is larger but it still takes a well-defined test with several similar subtests.

It is known that the subtest of the MMSE that is most sensitive to dementia is the three-word delayed recall. Three words. 1, 2, 3. Not 20 with its limited internal consistency as in Murdock [5] but three.

Braak & Braak [9] is cited frequently. If that paper is correct, surely with today's fMRI imaging techniques it should be a cinch to create a
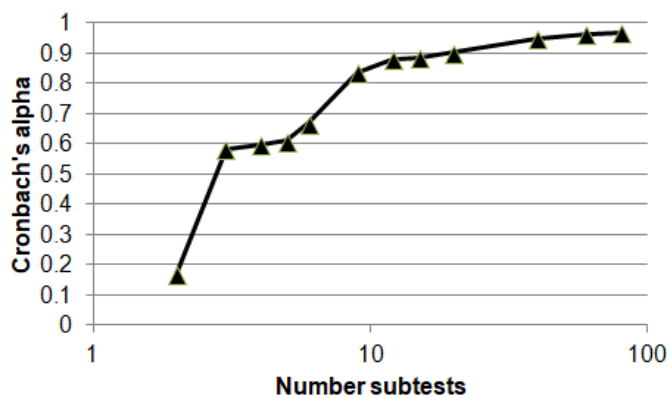

Figure 1. Cronbach's alpha as a function of the number of subtests for the classic Murdock (1962) 10-2 free recall test.
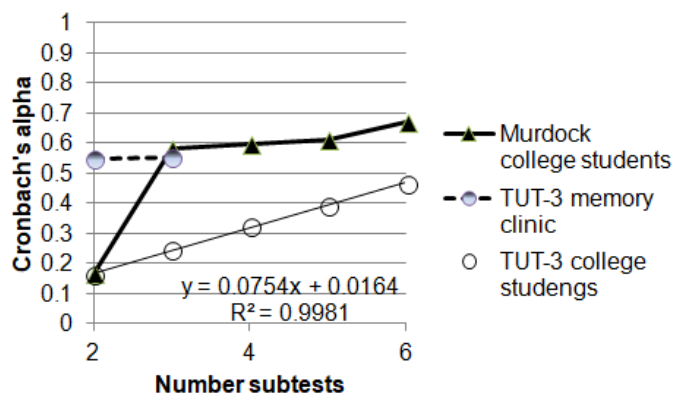

Figure 2. Cronbach's alpha as a function of subtests for the 3-item TUT tests of college students [7], the 3-item TUT test on an old population [8] and using the data from the classic Murdock 10-2 free recall test [5].

corresponding sequence of valid and reliable tests? But perhaps we do not believe Braak & Braak [9] – yet another issue we choose not to settle.

At some point we might ask ourselves, as the CONy 2018 conference does, since we are getting nowhere with AD - does AD actually exist (see, for example, Lam et all, [10])? And we might answer that, perhaps, AD should be relabeled "MMSE<24".

References

1. Ashford JW, Schmitt FA (2001) Modeling the time-course of Alzheimer dementia. *Curr Psychiatry Rep* 3: 20-28. [Crossref]

2. Buschke H, Kuslansky G, Katz M, Stewart WF, Sliwinski MJ, et al. (1999) Screening for dementia with the memory impairment screen. *Neurology* 52: 231-238. [Crossref]

3. Cullen B, O'Neill B, Evans JJ, Coen RF, Lawlor BA (2007) A review of screening tests for cognitive impairment. *Journal of Neurology, Neurosurgery & Psychiatry* 78: 790-799.

4. Molloy DW, Alemayehu E, Roberts R (1991) Reliability of a standardized mini-mental state examination compared with the traditional mini-mental state examination. *Am J Psychiatry* 148: 102-105.

5. Murdock B (1962) The serial position effect of free recall. *Journal of Experimental Psychology* 64: 482-488.

6. Thorndike EL, Lorge I (1944) The teacher's wordbook of 30,000 words. New York: Columbia University, Teachers College.

7. Ershova R, Tarnow E (2016) Working Memory Capacity Test Reveals Subjects' Difficulty Managing Limited Capacity, RUDN *Journal of Psychology and Pedagogics*: 14-19.

8. Tarnow E (2017) Preliminary Evidence: Diagnosed Alzheimer's Disease but not MCI Affects Working Memory: 0.6 of 2.6 Memory Pointers Lost, *J Alzheimers Dis Parkinsonism* 7: 315.

9. Braak H, Braak E (1995) Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol Aging* 16: 271-278. [Crossref]

10. Lam B, Masellis M, Freedman M, Stuss DT, Black SE (2013) Clinical, imaging, and pathological heterogeneity of the Alzheimer's disease syndrome. *Alzheimer's research & therapy* 5: 1.