Reviewer's report

Title: Item response theory analysis of cognitive tests in people with dementia: a systematic review

Version: 2 Date: 22 October 2013

Reviewer: hans wouters

Reviewer's report:

Thank you for your reply. Most answers were very clear to me.

Major Compulsory Revisions:

1. I still find it difficult to accept that non-English articles were excluded. Does this decision imply that everybody in Europe except for researchers in the UK should stop studying cognitive decline in older age because their non-English instruments are flawed? Excluding non-English articles is not justified for another reason: the authors did not pool the data but rather 'restrict' themselves to a narrative review. Even after excluding non-English articles they still report that at least 5 MMSE items are different for patients who live in the US compared to patients who live in the UK. Apparently, there is a language bias between American-English and British-English? Therefore, a better approach would be to discuss all relevant papers whether or not these are English or non-English and then to mention the caveat to study language bias using Differential Item Functioning (DIF) in the way this has been done by Teresi and co-workers.

2. Also of 384 articles, only four studies were selected. To me, excluding so many studies seems only appropriate when a quantitative meta-analysis is performed. However, the authors did not analyse their results quantitatively, but rather 'restrict' themselves to a narrative review.

3. Furthermore, at least the CAMCOG should be included.

Minor Essential Revision:

On page 21: I do not understand the phrase “Examining difficulty hierarchies of the error categories within the items revealed some disordered item categories, for example committing two errors being less difficult than committing one error”. Committing two errors being less difficult than committing one error actually reflects neatly ordered categories?

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable
Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.