Reviewer's report

Title: A Systematic Review of Tests of Empathy in Medicine

Version: 1 Date: 5 January 2007

Reviewer: James Wofford

Reviewer's report:

General

GENERAL – This manuscript is potentially publishable. However, most components of the manuscript should be improved.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

ABSTRACT
-- As the citation search is dated January 7th 2005, the search could be updated.
---The results section is thin. See comments below for candidates for inclusion in the abstract.

BACKGROUND
--- The explanation “conceptualizing empathy” could be improved. Other authors have separated empathy into multiple domains (emotive, moral, cognitive, behavioral) which allows the acknowledgement of intrinsic factors and factors that could be influenced by training. A good explanation of this approach is reflected in the reference (Suchman AL, Markakis K, Beckman HB, Frankel R. A model of empathic communication in the medical interview. JAMA 1997;277:678-80) and might improve the discussion of concepts.
--- After a good explanation of the three perspectives on empathy, the sentence on p. 5, line 28 “This systematic review was conducted to identify…” contains the phrase three types of empathy” that should be changed to be consistent with the previous discussion. Should it not be “three perspectives on empathy”?
--- “In essence, what we wanted to find out is whether, or to what extent, first person assessment correlated with second person assessment.” If this is going to be a complete systematic review of the subject, then why not address other aspects of measuring empathy. Why only first with second person? Why not first person with third person? Are there no studies? If so, they should be references in the text at least for completeness sake. Better yet would be to leave this sentence out, as it raises too many questions.

METHODS
--- The description of the search strategy and article selection, and the accompanying figure is clear.
--- The term “Search strategy” is more often used than is “inclusion procedure”. Change it.
--- If the criterion for choosing an article included the concept of emotional intelligence and quotient, the terms should have been explained as part of the introduction.
--- Where does the classification system of described in Box 1 come from? Should it be referenced?
--- Why not a box for reliability to match that for validity?
--- Shouldn’t “patient validation” be part of Box 1 – Classification of validity?
--- The term “patient validation” comes late in the article. Shouldn’t it be stated in the introduction where the three perspectives are outlined?

RESULTS
--- It should be more clearly stated how you got from the 33 citations (table 1 with 12 different tests) to the table of only 5 tests.
--- How many of the citations dealt with medical students, and how many were practicing doctors? This information should be somewhere in the results section, and preferably as a separate column in Table 1. This info should be included in the abstract results section.
--- p. 14 - The explanation of the different time intervals and their effect on the test results is more interpretation than results. Consider moving it to the discussion.
--- p. 14 - “All five tests had been reasonably well examined…” It is not clear whether this examination is part of each cited article or an examination that had occurred in the past.
--- p. 14 - “The closest assessment to this….” but this would be a third person perspective. Isn’t a patient validation a second person perspective?”
--- The results section is heavily weighted toward issues of reliability. Validity seems to be an afterthought, and the one paragraph devoted to it raises many questions.

DISCUSSION
--- If we understand correctly, you are demanding reliability and validity measures for the new MSAT. However, your analysis suggests that neither reliability nor validity measures are available or adequate for other tests that may be in use.
--- Providing the link is not necessary in the text. Cite it as a reference. The debate at the local level with these two universities is not necessarily worth the public accusations when the issue is really more global. Focus on the MSAT not the universities in order for an international audience to better understand the underlying issue.
--- With regard to the ‘gating” issue, if continuous measures of empathy have little validity, why would a less robust dichotomous variable be better at discerning empathy?
--- Why the paragraph about medics and nurses? It appears from what you have said that evidence about empathy testing for any health professional is inadequate/unavailable.
--- What is a medic”?

 Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

 Discretionary Revisions (which the author can choose to ignore)

Peter R. Lichstein
James L. Wofford

 What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

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

 Quality of written English: Acceptable

 Statistical review: No, the manuscript does not need to be seen by a statistician.

 Declaration of competing interests:

 I declare that I have no competing interests