Author's response to reviews

Title: A generalizability study of the medical judgment vignettes interview to assess students' noncognitive attributes for medical school

Authors:

Tyrone Donnon (tldennon@ucalgary.ca)
Elizabeth Oddone Paolucci (eoddone@ucalgary.ca)

Version: 2 Date: 10 November 2008

Author's response to reviews: see over
The BioMed Central Editorial Team
Medical Education

MS: 1632297871214380 – A generalizability study of the medical judgment vignettes interview to assess students’ noncognitive attributes for medical school

Authors: Tyrone Donnon and Elizabeth Oddone Paolucci

To Whom It May Concern:

Below are the point-by-point highlighted responses to the comments made by the two referees. Please note that the same revisions outlined below are also highlighted in the text of the revised manuscript.

Version: 1 Date: 28 August 2008 Reviewer: Ortrun Meissner

Reviewer's report:

Limitations of study:
1. The sample size is fairly small.
2. The students interviewed have already been admitted to med school!

Near the end of the ‘Discussion’ section in the last paragraph, the following sentence on the limitations of the study address the limitations in 1 & 2 above:

“Although both the reliability and validity of the Medical Judgment Vignettes interview appears promising, the selection of participants for convenience and limited sample size used in this study needs to be addressed in subsequent research…”

Own observation: I noticed the fairly high mean age of subjects as compared with the age of our own first year medical students!

Although our first medical students are typically older in comparison to other medical schools in Canada, in addition to an undergraduate degree some of our first year medical students have graduate degrees at the MSc and PhD level.
Reviewer's report:
This paper is fascinating, and provides a valuable insight into the validity of a tool which could be an excellent addition to the armoury or techniques which we can use to select candidates for medical school. The value of the paper is that it indicates that there is a tool which does not require multiple judges per station. I particularly liked the use of Kohlberg’s theory of moral development to frame the marking of the vignettes.

I have only two comments which the authors might consider.

1) The authors themselves appeared to be the judges in this study. This is appropriate, but it might contribute to the low inter-rater variability. Since the authors have clearly planned the study carefully, and doubtless rehearsed the criteria, they have almost certainly gone through a “norming” process. Consequently it might be surprising if they didn’t often agree. The authors might care to address this possible confounder. In my opinion, the study would have been more informative, and applicable more widely, if other judges had been chosen. The authors do address the requirements for training and consistency in the use of the scoring rubric on p.12.

Near the end of the ‘Discussion’ section, the following sentences on the limitations of the study was added:
“…Although the scoring rubric provides a clear anchoring for marking the Medical Judgment Vignettes, the authors were well grounded in the use of the rubric and future examiners may not have the same depth of understanding or time needed to train them appropriately. In particular, the use of a detailed scoring rubric may reduce the variability between raters, but the time required to educate the examiners in its use may simply be too prohibitive.”

2) My biggest philosophical problem with the study is in the extent to which Kohlberg’s theory of moral development is predictive. By this I mean that, although I am confident that the vignettes can give a “snapshot” of the candidate’s moral development, I do not personally know the evidence which shows that it develops rather than regresses, or varies randomly over time. I am sure that the authors can reassure the reader.

At the end of the ‘Conclusions’ section, the effect size demonstrating the predictive validity of the vignettes was added so that the final sentence reads:
“Although the predictive validity of the Medical Judgment Vignettes has been shown to have moderate effect size measures ($r = .45$) across clinical performance measures in clerkship [15], further research on how well these noncognitive outcomes reflect physician practice in residency and beyond is needed.”

Sincerely,

Tyrone Donnon, PhD (Corresponding Author)
Elizabeth Oddone Paolucci, PhD