**Author's response to reviews**

**Title:** Observed Communication Skills: How Do They Relate To The Consultation Content? A nation-wide study of graduate medical students seeing a standardized patient for a first-time consultation in a general practice setting.

**Authors:**

- Tore Gude (tore.gude@medisin.uio.no)
- Per Vaglum (p.j.w.vaglum@medisin.uio.no)
- Tor Anvik (tor.anvik@ism.uit.no)
- Anders Baerheim (anders.barheim@isf.uib.no)
- Hilde Eide (Hilde.Eide@su.hio.no)
- Ole Fasmer (ole.fasmer@psyk.uib.no)
- Peter Graugaard (p.k.gravgaard@medisin.uio.no)
- Hilde Grimstad (hilde.grimstad@medisin.ntnu.no)
- Per Hjortdahl (per.hjortdahl@medisin.uio.no)
- Are Holen (are.holen@medisin.ntnu.no)
- Tone Nordoy (tone.nordoy@unn.no)
- Helge Skirbekk (helge.skirbekk@medisin.uio.no)
- Arnstein Finset (arnstein.finset@medisin.uio.no)

**Version:** 3  **Date:** 4 July 2007

**Author's response to reviews:** see over
Reviewer's report:

General - The opportunity to collaborate on an educational study across institutions is an excellent one and I commend the authors for taking this opportunity to examine their work in this careful way. With that said, I am concerned that the conclusions that are drawn go far beyond what can be measured by a single standardized patient case in terms of communication skills. Petrusa in his review article on Clinical Performance Assessments in the International Handbook of Research in Medical Education, 2002, cites early work by Stillman (1990) that a reliability of .68 was achieved only after 3.25 hours of testing in the area of interviewing. In our own research (Guiton G, Hodgson CS, Delandshere G, & Wilkerson L. Communication skills in standardized –patient assessment of final-year medical students: a psychometric study. Advances in Health Sciences Education, 9: 179-87, 2004) we believe we have demonstrated that communication skills are case specific; some communication skills are more important in one case than in another. In a study using only one case, the conclusions should be very narrowly limited to the situation in this one case.

Authors’ comment:

The concerns about concluding limitations with only one interviewed patient as assessment basis for observation of communication skills and the short time used for it (15 minutes as in an ordinary general practice consultation) is addressed in the introduction with references included and discussed later (Stillman’s 1991 is addressing the same issue as that one referred to by Petrusa).

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

pg. 5 Some further effort should be made to demonstrate that the students who volunteered were similar to others in the same institution using existing measures or at least demographics. The completion of a self-assessment survey by a previous class is not sufficient to suggest that the sample represents something other than a sample of convenience.

Authors’ comment:

We have now extended the check for representativity of the final-year 111 students who participated in the observational study with the whole final-year cohort (N = 320) they were recruited from. No differences in gender and age were found. As reported in our first draft, the scores on self-assessed communication skills among the same 320 students in their 5th year were similar to the scores among the 111 (of the 320) one year later when we carried out our observation study, i.e. it was not different classes, but the whole class and then part of it one year later (page 5).
You need to be sure that you state your conclusions very narrowly. In this one sample of final year medical students on this one case...

**Authors’ comment:**

We have modified our introductory remarks and conclusions in the discussion part.

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

Pg 7 I found the description of the consultation content indices and their combination into 8 subgroups to be very difficult to follow. I would suggest trying to describe this more clearly. You also need to report on the internal consistency of these indices just as you did on the ACIR. This will make it clear that you are aware of the limited generalizability of your findings. These same students may even behave differently in a similar case with these same dimensions but different medical content.

**Authors’ comment:**

We have argued for not constructing the content indices on the basis of a PCA and tried to clarify even more the construction of them including their internal consistency being aware of the moderate levels of Cronbach’s alpha (page 8).

Pg 7 The MAAS items were scored on a scale of 0-5 according to the text but appear in the Table 2 as yes/no items. This needs to be clarified.

**Authors’ comment:**

The MAAS items (nr. 6 and 9) used as a content index (Informing/planning) was dichotomized on their median in order to match the other indices based on dichotomized scores. Further, the construction of the possible eight combinations based on low/high counts on three of the four content indices (index two and three merged) is more clearly described (page 8).

Pg 8 The sorting of the ACIR items as content and non-content items needs to be tested for internal consistency just as you did for the entire set of items. Given the high internal consistency and principal component analysis results, I do not think you are justified in producing a new scale without conducting the same level of analysis.

**Authors’ comment:**

We have validated the ACIR scores against the MAAS scores and maintained the non-content partitioning by arguing for the way we selected the non-content items being consonant with the same procedure in another study (Aspegren, 2005). We have reported the internal consistency (.86) of the non-content mean (page 9).

Discretionary Revisions (which the author can choose to ignore)

Pg 4 Given the intensive communication skills training provided at one of the schools, I would suggest omitting curricular differences from your hypotheses and investigation.

**Authors’ comment:**

We have omitted the comparison between schools.
Pg 7 It would be more appropriate to formulate the indices based on a factor analysis.

Authors’ comment:
   We have argued for not constructing the content indices on the basis of a PCA (page 8).

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

Level of interest: An article of limited interest

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests: I declare that I have no competing interests.

Reviewer's report – 2.
Title: Observed Communication Skills: How Do They Relate To The Consultation Content? A nation-wide study of medical students with a simulated patient in a first general practice consultation.
Version: 1
Date: 9 May 2007
Reviewer: Nobutaro BAN
Reviewer's report:
General - We have an impression that students who are conscious about patient-centeredness often pay attention to communication skills, patients’ concern and psychosomatic aspects of the patients. This article is an very ambitious one trying to document this general impression.

------------------------------------------------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)
The major flaw of this paper however is the lack of validity of the ACIR that authors used as an independent instrument to measure communication skills. Authors well recognize this weakness and try to defend. However, unfortunately, it is not successful. Authors categorized item number 3, 4, 5, 6, 8, 10 and 11 of the ACIR as non-content-related. However, item number 1, 2, 12 are also non-content-related. In other words, the ACIR consists of most of the items of non-content-related. Thus it cannot defend the risk of tautology by comparing the result of difference when using the overall index in contrast with the non-content-related index of the ACIR.

Authors’ comment:
   We have validated the ACIR scores against the MAAS scores and maintained the non-content partitioning by arguing for the way we selected the non-content items being consonant with the same procedure in another study (Aspegren, 2005). We realize that such a partitioning can be viewed in different ways, but all
the way we find it reasonable to perform this procedure as one way of reducing the risk of tautology. We have reported the internal consistency (.86) of the non-content index mean (page 9).

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
There are several minor suggestions, such as inquiry about the superior result of traditional curriculum to integrated curriculum in depth, need for editing for English language usage, grammar and spelling and so on, but compared with the above mentioned major flaw, those would be a minor problems.

Authors’ comment:
We have omitted the comparison between schools (curricula). The revised manuscript has been reviewed by an English language consultant.

Discretionary Revisions (which the author can choose to ignore)
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: Needs some language corrections before being published.

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests: I declare that I have no competing interest