

**Herbert PLASS (Wien), Lambert SCHUWIRTH (Maastricht),
Monika KILLER (Salzburg), Wolfgang PRODINGER
(Innsbruck), Gilbert REIBNEGGER (Graz), Wolfgang KÜNZEL,
Rudolf MALLINGER, Richard MARZ¹ (Wien)**

Assessment at Austrian Medical Schools – Results of a 2007 Survey

Abstract

Since 1995 the „Graz Conference on Medical Education“ has served as a discussion forum for curriculum development at Austrian medical schools. A main topic of the 11th Conference 2007 was the current status and developmental aspects of assessment. Answering 10 questions, officials of five universities outlined their approach to meet legal requirements as well as the internal needs of the institution. Although assessment methods and quality management differ, all attempt to implement methods that conform to international standards.

Keywords

Educational measurement, accreditation, medical school, professional competence

Prüfungswesen an österreichischen medizinischen Universitäten – Überblick 2007

Zusammenfassung

Seit 1995 war die „Grazer Konferenz - Qualität der Lehre“ Diskussionsforum für Curriculum-Entwicklung der österreichischen medizinischen Universitäten. Ein Schwerpunkt der 11. Konferenz 2007, war Stand und Entwicklung des Prüfungswesens. Anhand von 10 Fragen erläuterten Vertreter der Universitäten das Vorgehen um gesetzlichen Auflagen sowie lokalen Notwendigkeiten Rechnung zu tragen. Obwohl die Überprüfung des Studienfortschritts sowie die Qualitätskontrolle des Prüfungswesens unterschiedlich implementiert wurden, ist die Erfüllung internationaler Standards Ziel aller Universitäten.

Schlüsselwörter

Prüfungswesen, Akkreditierung, medizinische Universität, Fachkompetenz

¹ e-Mail: richard.maerz@meduniwien.ac.at

1 Introduction

In May 2007 the „Graz Conference on Medical Education“ (Qualität der Lehre in der Medizin) took place for the 11th time. Since 1995, when the conference was first held, its main aim has been to serve as a discussion forum for those interested in curriculum development at Austrian medical schools. This year's conference had the motto: Be prepared – teaching students what they need to know. It was held at the Paracelsus Medical University in Salzburg, Austria's only private medical school, which started accepting students in 2003.² The other Austrian participants came from the three much older federal Medical Universities (Vienna, Graz, Innsbruck) and from the University of Veterinary Medicine in Vienna. In addition there were participants from 9 other countries.

Starting with the first conference, examinations have been a topic of great interest. A change in the examination system was regarded as the key to meaningful change in the curriculum (SCHUWIRTH, 1996; BLOCH, 1998; ENGEL, 1999). Until the implementation of the new curriculum at Austrian Medical Universities the examination system mandated by law consisted of 23 separate examinations each covering an individual subject like physiology or internal medicine. The format of the examination was in almost all cases an oral held by an expert examiner. Thus the scope and depth of the material to be learned was under departmental control. This format thus precluded both centralized control over the curriculum as well as the implementation of meaningful quality control of the examination system.

Although under the new law each university had been responsible for developing their own new curriculum and examination system, a common element were written examinations covering several subject areas. For this years conference we decided that it was time to revisit the topic. A questionnaire regarding various aspects of the examination system of the type used for external evaluation or accreditation purposes was drawn up. It consists of ten questions.

2 Questionnaire

- 1: What instruments do you use for the assessment of the medical competence of your students? Can you give a rationale as to why these instruments were chosen?
- 2: How do you combine the results of various individual tests to determine whether a student can progress to a next semester / phase / year?
- 3: How is the pass-fail score for each of the instruments determined?
- 4: Is there any type of quality control (e.g. test item review panels) prior to test administrations? If so, please describe it.
- 5: Is there any type of quality control (e.g. item analyses) after the test administration? If so, please describe it.

² The conference home page can be accessed under:
<http://www.mededu.meduniwien.ac.at/grako07/>

- 6: Do you use external examiners/accreditation system? If so, please describe it.
- 7: Is there any research with respect to the quality of your assessment system? If so, please describe it.
- 8: Do you use IT to support your test production, test administration or quality control? If so, please describe it.
- 9: What is your policy with respect to repeat examinations? Do you track students who have failed an examination? How soon and how often can they retake a failed attempt? Is there a time limit for the completion of their studies?
- 10: What are the major problems you face?

The questionnaire was sent to officials of the four Austrian medical schools and of the school of veterinary medicine. One session of the conference was devoted to the presentation of answers (given below) and discussion of the results.

3 Results

3.1 Instruments Used for Assessment

Question 1: What instruments do you use for the assessment of the medical competence of your students? Can you give a rationale as to why these instruments were chosen?

Paracelsus Medical University of Salzburg / PMU ³

- Multiple Choice Questions (MCQ) examinations are used primarily to assess basic factual information.
- Short answer papers help to assess the breadth of core factual knowledge.
- Essay, or modified essay papers assess in-depth knowledge, and understanding.
- Objective Structured Practical/Clinical Examinations (OSPE/OSCE⁴) will be used to assess practical skills and competencies in clinical procedures.
- Oral examinations are used to assess knowledge mainly of a higher domain. They are also used as an instrument in cases of doubt (repeat examinations).
- Portfolio-like assessment: Where a course is concerned with the development of professional practice, portfolio-like assessment is used.
- Attendance is required and monitored.
- Written assignments are used in those modules where the intended learning outcomes are limited to gaining knowledge. (Written assignments reflect the content and learning outcomes of the course.)

³ For PMU all questions were answered by Dr. Monika Killer, Dean of Student Affairs, m.killer@salk.at

⁴ See: http://en.wikipedia.org/wiki/Objective_Structured_Clinical_Examination

- The scientific training is assessed by presentation of a dissertation.

Medical University of Graz / MUG ⁵

- Generally, each of the first five years of study is divided into 6 modules of five weeks. Modules consist of lectures, seminars and laboratory courses / practical exercises. Students have to complete the seminars and laboratory courses / practical exercises belonging to a certain module before they are allowed to take the final examinations of the modules, the so called "Fachprüfungen".
- In small group courses like seminars and laboratory courses, lecturers are free to choose their preferred way of examination. Oral, written, and practical tests are used.
- The final examinations of the modules are mainly written exams consisting of multiple choice, short answer or short essay questions. In rare cases there are also oral examinations included. At the end of the fifth year of study as well as during their sixth year (clinical practical year), students have to participate in an objective structured clinical examination (OSCE).
- All instruments mentioned above should be used only after considering their educational impact.
- Multiple-choice exams are administered either as paper pencil or as online exams.

Medical University of Innsbruck / MUI ⁶

- The mainstays of the exam system are the four "SIPs", summative written integrative year-end exams (MCQ), which are spaced over the duration of study
- Each SIP (160-240 MCQ) has to be passed before students are allowed to progress to the next phase.
 - o These replace the 23 oral or oral + written exams of the old curriculum
- Practicals:
 - o Here passing grades can be largely obtained by attendance ...
 - o Alternatively, direct observation; written / oral / practical (SP) exams are used.
- OSCE and portfolio assessment will be introduced starting 2008.
- Formative written assessments:
 - o A set of MCQs are freely available for self-evaluation.
 - o Participation in the Berlin Progress Test will start in autumn 2007.

⁵ For MUG all questions were answered by Prof. Dr. Gilbert Reibnegger, Vice-Rector for Teaching and Studies, gilbert.reibnegger@meduni-graz.at

⁶ For MUI all questions were answered by Prof. Dr. Wolfgang Prodinger, Chair of the Curriculum Commission, wolfgang.prodinger@i-med.ac.at

Medical University of Vienna / MUW ⁷

- Summative and formative assessments (SIP and FIP) at the end of each year (MCQ-Format, 125-230 questions per test).
- Small group courses are graded by the tutor (mostly non-standardized)
- Propedeutics for clerkship: OSCE-Format.
- Clinical Examination during 3rd (clinical) stage: CEX (Clinical Evaluation Exercise) -Format (e.g. <http://meded.utmb.edu/year4/icee/icee.htm>)

University of Veterinary Medicine Vienna / VUW ⁸

- Single subject examinations have been replaced by comprehensive examinations given by committees.
- Reduction in the number of examinations from 38 to 14.
- Implementation of Objective Structured Clinical Examinations (OSCE).
 - o This has led to an increase in the objectivity, transparency and reliability of examinations.
- Practical exercises held at external sites are now required.
- Required clinical skills are now defined and monitored by „log books“.

3.2 Decision Whether a Student can Progress

Question 2: How do you combine the results of various individual tests to determine whether a student can progress to a next semester / phase / year?

Paracelsus MU Salzburg

- Policy of Continuous Assessment:
 - o Attendance requirement of 85%.
 - o Student eligibility for examinations is determined by an individual's attendance, and performance.
- *All* examinations must be passed.
- Large subjects (e.g. Pathophysiology) may be covered by several examinations. In such cases the final grade is determined by using ECTS points as the weighting factor for individual results.

MU Graz

- Students must pass *all* required examinations of a particular stage (Studienabschnitt) of the curriculum, before they can progress to the next stage.

⁷ For MUI all questions were answered by Prof. Dr. Rudolf Mallinger, Vice-Rector for Teaching, rudolf.mallinger@meduniwien.ac.at

⁸ For VUW all questions were answered by Prof. Dr. Wolfgang Künzel, Vice-Rector for Teaching, Wolfgang.Kuenzel@vu-wien.ac.at

MU Innsbruck

- Practical exercises are prerequisite for taking the SIP.
- Progress to the next phase of the curriculum is based on the SIP score.

MU Vienna

- Progress to the next stage of the curriculum requires passing all exams of the current stage.
- Progress to next semester/year within one stage is possible even with negative test results.

Veterinary University Vienna

- Students have to pass all the required examinations in order to enter the next stage of the curriculum.
- For some courses there are separately defined parts that must all be passed.

3.3 Pass/fail Scores

Question 3: How is the pass/fail score for each of the instruments determined?

Paracelsus MU Salzburg

Pass/fail score for oral, written and practical exams, including OSCE:

(1) Sehr gut	91-100%	= excellent	performance far above average
(2) Gut	81-90,9%	= satisfactory	performance above average
(3) Befriedigend	71-80,9%	= satisfactory	performance on average
(4) Genügend	61-70,9%	= incomplete	performance just acceptable
(5) Nicht genügend	< 60,9%	= inadequate	performance not acceptable

Pass/fail score for MCQ:

(1) Sehr gut	92-100%
(2) Gut	83-91,9%
(3) Befriedigend	74-82,9%
(4) Genügend	66-73,9%
(5) Nicht genügend	< 65,9%

MU Graz

- There are no general rules concerning pass/fail scores, only recommendations.
- Except for multiple-choice exams - here in general 66 percent of the maximum score have to be obtained to pass an exam – otherwise the pass/fail score consists of at least 50 percent of the maximum score.

MU Innsbruck

- SIPs (end of year exams):

- Generally 60%.
- Some individual sections have also (lower) subscores which must be attained.
- Practical exercises:
 - Attendance (85%) prevails.
 - Oral > written > practical exams (preclinical phase).
 - Scores depending on format.
- OSCE and portfolio: to be decided in June 2008 (first application).

MU Vienna

- Summative assessment: Passing score is 66%. In future a change may take place based on analysis of previous exams.
 - All individual sections need a subscore of 60% for passing.
- Pass/fail score for other courses is set individually by lecturer. Failures are rare.

Veterinary University Vienna

- The pass score is generally defined as 60% with the exception of OSCE's where standard setting are set separately by examiner boards.

3.4 Quality Control Prior to Test Administrations

Question 4: Is there any type of quality control (e.g. test item review panels) prior to test administrations? If so, please describe it.

Paracelsus MU Salzburg

- Task of the dean for curriculum and his team is to evaluate all items prior to test administration.
- There is no item review panel – up to now.
- Exact information about the examination modality and dates are given at the beginning of each course.

MU Graz

- Multiple-choice questions are tracked within an electronic test management system. After the authors enter test items, formal quality control takes place at the Unit of Administration of Evaluations and Examinations. Questions are reviewed according to correct spelling and (technical) requirements of the test management system.

MU Innsbruck

- Pre-test multi-step review process:
 - Peer review.
 - Review by experts in methodology.
 - Test review committee (interdisciplinary) for each SIP.

MU Vienna

- SIP: Review by panels consisting of experts both in methodological reviewing as well as in the content reviewing. Each SIP has its own review panel. Workflow for reviewing and handling of rejected questions is defined.
- IT-supported yearly revision of all questions in the pool by coordinators and authors is in preparation.

3.5 Quality Control after Test Administration

Question 5: Is there any type of quality control (e.g. item analyses) after the test administration? If so, please describe it.

Paracelsus MU Salzburg

- Task of the dean for academic affairs and his team is to evaluate all courses; this includes tracking how well students are prepared for the examination.
 - o Evaluation at the end of each course.
 - o Two randomly selected students, the dean of academic affairs, and the teacher meet and discuss the evaluation results and the appropriateness of the examination.

MU Graz

- After administration and statistical evaluation item analyses are routinely conducted. Critical items (e.g. low or negative item-test correlation) are reported to the examiner and depending on his feedback they are either eliminated or adjusted for later use.

MU Innsbruck

- Two-step post-test item review process:
 - o 1. Student critique process + Statistical analysis.
 - o 2. Test committee decision and formal point-to-point reply to students.

MU Vienna

- Quality feedback procedure by a reporting system, which allows students to report about problems encountered with single questions during the exam.⁹
- Item analysis of all questions; author-feedback concerning questions with bad item statistics and/or student criticism; elimination of non-answerable questions (usually 0-5 questions per exam).
- Results of item analysis are reported to authors of MC-questions.

⁹ Students are encouraged to report suspect questions a member of the supervising staff, who enter the complaints in a protocol. These protocols are evaluated before the exam is scored and occasionally questions are eliminated.

Veterinary University Vienna

- Examination results are critically evaluated.
- Identify questions causing disproportional failing.

3.6 External Examiners/Accreditation System

Question 6: Do you use external examiners/accreditation system? If so, please describe it.

Paracelsus MU Salzburg

- External examiners (histology, pathology, anatomy)
- USLME (United States Medical Licensing Examination¹⁰) Step 1 after 3rd year
- The university is subject to control of the Austrian accreditation body.

MU Graz

- An external accreditation for the Diploma Program of Human Medicine by the German Quality Assurance Agency (ACQUIN)¹¹ is presently in progress.

MU Innsbruck

- No accreditation is currently taking place.
- Students can participate in external progress tests on a voluntary basis:
 - o Berliner Progress-Test (Start October 2007)

MU Vienna

- No

Veterinary University Vienna

- Accredited veterinary faculties are evaluated worldwide by the European Association of Establishments for Veterinary Education (EAEVE¹²; for Europe) or the American Veterinary Medical Association (AVMA¹³; for America, Australia, New Zealand). The University of Veterinary Medicine Vienna was last evaluated in 2006 and accredited in 2007.
- An international expert group does the evaluation.
 - o Visiting the faculty for 1 week.
 - o On the basis of a SER (Self Evaluation Report).
 - o Team consists of 6 members, including 1 practitioner.

¹⁰ <http://www.usmle.org/>

¹¹ <http://www.acquin.org/acquincms/>

¹² <http://www.eaeve.org/>

¹³ <http://www.avma.org/>

3.7 Research into Quality of Assessment System

Question 7: Is there any research with respect to the quality of your assessment system? If so, please describe it.

Paracelsus MU Salzburg

- Not up to now.

MU Graz

- Not at the moment. Exceptions are investigations concerning the utility of computer-assisted learning in comparison with traditional learning.

MU Innsbruck

- In particular for the SIP1 (end of year 1 exam):
 - o Learning-styles and performance.
 - o Admission and performance.
 - o Concept of the SIP 1 (authored by Hans-Georg Kraft (AMEE¹⁴ contributions, MME thesis etc.).

MU Vienna

- Outcome-analysis of MC tests.
- Analysis of learning behaviour imposed by the assessment system.

Veterinary University Vienna

- The centre of study affairs does research on quality.

3.8 IT Support

Question 8: Do you use IT to support your test production, test administration or quality control? If so, please describe it.

Paracelsus MU Salzburg

- Moodle.

MU Graz

- Multiple-choice questions are managed within an electronic test management system (Perception¹⁵). This system is also employed for post-examination processing and item analysis.
- Furthermore, the possibility to run on-line MC assessments (fully computer-based examinations) has been established and is being used with great success.

¹⁴ Association for Medical Education in Europe, <http://www.amee.org/index.asp>

¹⁵ Perception ist a product of Questionmark: www.questionmark.com.

MU Innsbruck

- Commercial Product (LogiExam® 1.6¹⁶):
 - o Central database (secured).
 - o Front-end(s) used for producing and administering tests.
 - o Test sheets are scanned (OCR).
- Additional statistical item analysis (in house).

MU Vienna

- Database with MC questions.¹⁷
- Approved blueprints (blocks, chapters, themes) for composing test.
- IT-based administration.
- Automatic scanning of exam-sheets.¹⁸
- IT-based calculation of grading.
- IT-based production of individual feedback-reports with online-access for the students.

Veterinary University Vienna

- 2005 an IT room for on-line examinations was set up (52 work stations).
- Since 2006 especially MC based examinations run computer based.
- In 2007 the use of the eLearning platform will be extended to examinations for self-control as well as for formal examinations¹⁹.

3.9 Policy with Respect to Repeat Examination

Question 9: What is your policy with respect to repeat examinations? Do you track students who have failed an examination? How soon and how often can they retake a failed attempt? Is there a time limit for the completion of their studies?

Explanatory note: While in many countries this is a legitimate question to ask universities, in Austria this must indeed be passed on to politicians. Current law decrees that a student must be given the chance of a minimum of three retakes (four attempts). Only the Paracelsus Medical University as a private institution has here some leeway.

¹⁶ <http://www.logiexam.de/>

¹⁷ Currently in use: „Question Bank“ (<http://qb.soc.surrey.ac.uk/>) by Speedwell, UK. Will be replaced end of 2007 by an in-house development: „m3e-exam“.

¹⁸ Currently in use: „multiquest“ (<http://www.wmin.ac.uk/page-8367>) by Speedwell, UK. Will be replaced by „Forms5“ from Kaiser data (<http://www.kaiserdata.at/index.htm?programme.htm~main>).

¹⁹ The VUW uses Blackboard as ist eLearning platform: http://www.vu-wien.ac.at/elearning/content/index_ger.html

Paracelsus MU Salzburg

- If the exam is “failed” or “not completed without excuse”, a maximum of 3 repetitions is allowed:
 - o Repetition 1: oral or written – within 14 days.
 - o Repetition 2: oral – as soon as possible, latest 1 month before the start of the next semester.
 - o Repetition 3: “kommissionelle Prüfung” i.e. Examination with the minimum of 2-3 expert examiners and the Dean of student affairs. Latest date: last day of the academic year.

MU Graz

- With regard to the repeat of examinations we are following the Austrian University Act (Universitätsgesetz) 2002. We are tracking students who failed, because there are a limited number of attempts, but students are allowed to take the examination at the next possible date. There is no time limit for graduation.

MU Innsbruck

- 3 retakes (no additional retakes were allowed by the University Senate).
- No time limit to complete the programme.
- We do analyse the exam progress on a cohort level (anonymously).
- We do not follow (and interview etc.) individual students.

MU Vienna

- 4 retakes (one additional retake was allowed by the University Senate).
- No time limit to complete the programme.
- No tracking (mentoring).

Veterinary University Vienna

- 3 retakes in the first stage of the degree programme.
- 4 retakes in the second and third stage.
- No time barrier for the next retake.
- No time limit to complete the programme.

3.10 Major Problems

Question 10: What are the major problems you face?

Paracelsus MU Salzburg

- The packed curriculum does not always leave enough time for extra learning and repetition.
- Standardization of examinations in practicals.

MU Graz

- Frequency of examinations (6 per year).
- Copying by students of MC questions.

MU Innsbruck

- Diversity of assessment methods not yet established (particularly lacking: practical exams and longitudinal assessments).
- Quality of the MC questions provided by the faculty.
- Preparation of four sets of each MCQ exam per year exhausts the MCQ pool.
- Unrealistic ideas about traditional oral examinations among teachers and students.

MU Vienna

- Legally defined high frequency of examinations.
- Collecting a sufficient amount of new MC-questions.
- Standardization of examinations held in small groups.

Veterinary University Vienna

- Examinations have to be held while the semester is in progress.

4 Discussion

While plenty of problems remain, some resulting from Austrian law but some decidedly home-grown, the sense of the discussion was clearly that the progress made over the last years has been remarkable. Formative examinations have been introduced. The number of examinations has decreased. As a consequence the material covered by each examination has increased thus limiting the cramming / forgetting cycles. More integrative examinations have replaced those covering a single-subject. Oral exams are being replaced by more objective formats. Examinations are no longer limited to knowledge. Most important there is now an awareness that the examination system is not a given but the most important element of the curriculum which must be subjected to continuous development and rigorous quality control.

Yet there are always possibilities to improve. Relating to the questions some observations and advices are:

The choice and use of instruments

In the answers of the medical schools considerable emphasis was placed on the format of the assessment, rather than on the content. However, the literature suggests that the form of the examination is far less important in determining what the test actually measures than the content (WARD, 1982; MAATSCH & HUANG, 1986; NORMAN et.al., 1987; SCHUWIRTH et.al., 2001). There is much belief that open-ended questions and oral examinations test higher-order

cognitive skills, yet if the content of such examinations is reviewed, often low level factual knowledge is asked, which could better have been asked using written multiple-choice type examinations.

Combination of results

Many schools use a highly conjunctive way of combining result; requiring that all elements are passed before the student can go on to the next phase. It is logical that from a measurement point of view a compensatory method leads to higher reliabilities, because a conjunctive model adds up false-negative results (HOFSTEE, 1983).

Standard setting

Many different methods of standard setting have been described in the literature (CUSIMANO, 1996). This indicates that there is no single best method for all types of tests, and that the optimal method should be chosen according to the specific type and content of the examination. The literature converges on the notion that a randomly chosen standard (e.g. 60% for all examinations) is poor practice.

Quality control

Although the empirical literature on this topic is limited, it is generally held that quality control is an essential part of high quality education and assessment. Review panels that check all raw test material on content, wording and relevancy before the test, item analyses and student comments after the test, have worldwide been shown to be of great benefit for the quality of examinations.

Accreditation

Accreditation serves a double purpose. Most people see accreditation as a procedure to determine whether a curriculum is up to standard. A more central question is, however, how any accredited curriculum could improve further. Therefore, in our view it is essential that accreditation panels consist of a combination of content experts and educational experts.

Research

Medical education research is becoming more and more important. In the “Research Assessment Exercises” in the UK and Australia for example, medical education research is taken very seriously and compared to biomedical research. Medical education research is also a rapidly expanding discipline (more than 8 international peer-reviewed impact factor journals) that is professionalizing the field. However, local evaluation studies are very hard to get published currently, and this will become even harder in the near future. Anyone who wants to conduct and publish research in medical education, will have to concentrate on research that has a clear relevance outside his or her own institutions.

IT support

IT support is well established in all medical schools. This is important, because not only are our current students better equipped to use computers (RUSSELL &

HANEY, 2000), there are also possibilities in computer-based testing that can improve the quality of our examinations (CLAUSER & SCHUWIRTH, 2002). Thus IT-platforms would not be optimally used if these possibilities were not explored in medical schools.

Repeat Examinations

Many disadvantages of repeat examinations have been reported. They capitalize on false-positive results (students passing purely by chance despite unsatisfactory knowledge) (McMANUS, 1992), and induce unwanted study behaviour. They also increase attrition rates (COHEN-SCHOTANUS, 1999).

In all the answers to the questions have been highly informative aspects. On the one hand it was impressive to see how much progress had been made in the improvement of quality of the curricula as well as the assessment and on the other hand they provide ample handles for future improvements.

5 References

- Bloch, R.** (1998). Assessment and Evaluation in a Modern Curriculum – Monitoring and Improving. *Zeitschrift für Hochschuldidaktik*, 1998. **22**(4): p. 25-36.
- Clauser, B.E. & Schuwirth, L.W.T.** (2002). The use of computers in assessment, in *International handbook of research in medical education*, G.R. Norman, C.P.M. van der Vleuten, and D.I. Newble, Editors. 2002, Kluwer Academic publishers: Dordrecht.
- Cohen-Schotanus, J.** (1999). Student assessment and examination rules. *Medical Teacher*, 1999. **21**(3): p. 318-21.
- Cusimano, M.D.** (1996). Standard setting in medical education. *Academic Medicine*, 1996. **71**(10 Suppl): p. S112-20.
- Engel, C.E.** (1999). Some Thoughts on Formative Assessment. *Zeitschrift für Hochschuldidaktik*, 1999. **23**(2): p. 16-29.
- Hofstee, W.K.B.** (1983). The Case for Compromise in Educational Selection and Grading, in *On educational testing*, S.B. Anderson and J.S. Helmick, Editors. 1983, Josey-Bass: San-Francisco.
- Maatsch, J. & Huang, R.** (1986). An evaluation of the construct validity of four alternative theories of clinical competence. in *Proceedings of the 25th annual RIME conference*. 1986. Chicago: AAMC.
- McManus, I.C.** (1992). Does performance improve when candidates resit a postgraduate examination? *Medical Education*, 1992. **26**(2): p. 157-62.
- Norman, G.R., Smith, E.K., Powles, A.C., Rooney, P.J., Henry, N.L. & Dodd, P.E.** (1987). Factors underlying performance on written tests of knowledge. *Medical Education*, 1987. **21**(4): p. 297-304.
- Russell, M. & Haney, W.** (2000). Bridging the Gap between Testing and Technology in Schools. *Education Policy Analysis Archives*, **8**(19): p. 1-9.
- Schuwirth, L.W.T.** (1996). Quality Control: Assessment and Examinations. *Zeitschrift für Hochschuldidaktik*, 1996. **20**(1-2): p. 66-76.

Schuwirth, L.W., Verheggen, M.M., van der Vleuten, C.P., Boshuizen, H.P. & Dinant, G.J. (2001). Do short cases elicit different thinking processes than factual knowledge questions do? *Medical Education*, 2001. **35**(4): p. 348-56.

Ward, W.C. (1982). A comparison of free-response and multiple-choice forms of verbal aptitude tests. *Applied Psychological Measurement*, 1982. **6**(1): p. 1-11.