

ABSTRACT NUMBER: OA1

Quantitative Assessment: Compromising Student And Patient Outcomes In Chiropractic Teaching Clinics

Adrian Repka

RMIT University, Australia

Background

This study was undertaken to determine the extent to which quantitative methods of assessment in the RMIT Chiropractic Teaching Clinics compromise qualitative outcomes for both patients and students.

Methods

A sample of final year students in the RMIT chiropractic program each reviewed and summarised health care records of 10 randomly selected patients attending the University student teaching clinics during the preceding 6 months, resulting in data on the management of 400 patients.

Results

Patient records were reviewed to determine the manner in which students diagnosed and managed their patients. Of the 400 patient diagnoses reviewed, 355 (88.75%) were regarded as spinal in nature and warranted chiropractic intervention and management. Of the 355, only 13 (3.7%) were considered for management beyond the symptomatic presentation of spinal pain.

Conclusion

Students diagnosed and managed patients attending the clinic in only the most minimalist manner, rather than with a preferred holistic approach. Despite being provided with an extensive variety of diagnostic and management tools that enable them to assess and manage patients' health issues beyond just mechanical joint pain, the greater majority of students within the RMIT teaching clinics do not do so. Holistic patient management requires the student to spend more time on each case presentation. Students are assessed on the number of presentations seen throughout their clinical internship, not necessarily on the manner in which they manage those patients. As a result, they may not see it as being in their best interest to manage beyond the simplistic. The author suggests that the problem may be addressed through the development and implementation of a qualitative method of assessment, where student learning outcomes are assessed primarily against patient management.

ABSTRACT NUMBER: OA2

The Discriminative Value of Blooms' Level Of Abstraction In Evaluating OSPE Scores Among Final Year Medical Students

K Y Loh¹, Sivalingam N², Yushak A W³, Zainur Rashid Z²

¹Department of Family Medicine, International Medical University Malaysia, ²Department of O&G, International Medical University Malaysia, ³Department of Surgery, International Medical University Malaysia

Introduction

Objective Structured Practical Examination (OSPE) is a better assessment technique over the traditional clinical examination method in testing a wide range of medical knowledge and practical skills. In OSPE the variables and complexity of the examination are more easily controlled and the marking is objective. In 1956 Benjamin Bloom created a taxonomy for categorizing levels of abstraction of questions that commonly occur in educational settings into 6 levels, i.e: (1) knowledge (recall), (2) comprehension, (3) application, (4) analysis, (5) synthesis and (6) evaluation. As the medical students proceed to higher semesters, we expect them to be more competent and be able to perform better at answering questions at the higher level of Bloom's hierarchy.

Objective

- (i) to review the final year OSPE questions and categorize the level of abstraction according to Bloom's taxonomy.
- (ii) to compare the difference in the student's performance score at different levels of Blooms' taxonomy.

Material and Methods

This is a retrospective descriptive survey. A total of 18 OSPE questions were reviewed by two authors to determine the level of abstraction using criteria stated in Blooms' taxonomy. The OSPE stations comprise questions requiring students to interpret laboratory results, interpret radiological imaging, identify medical instruments and address evidence-based commentaries. The mean score for each station was determined from the students' scores. The domains tested in each station was determined by the two senior authors and compared with the levels of Bloom's taxonomy.

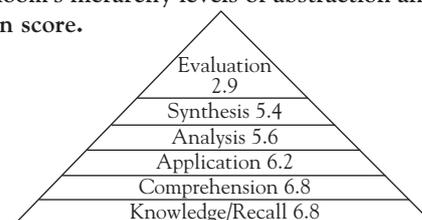
Results

Eight out of 18 stations (44.4%) belong to category levels 1, 2, 3. Only 2 stations tested at higher levels (4, 5,6). Only 1 question tested at the highest level (level 6). Figure 1 shows the mean scores of the OSPE according to the categorized level of abstraction. As the Blooms' level of abstraction goes higher, the mean score of students declined. (Figure 1) There is a statistically significant difference between the score for level 1 (knowledge and recall) compared to level 6 (evaluation) (p=0.001).

Conclusion

The majority of the OSPE stations were at level 1 (knowledge/recall), level 2 (comprehension) and level 3 (application) according to the Blooms' taxonomy. This survey demonstrates that if the nature of the OSPE is to be maintained, it is not suitable to test higher levels of abstraction. Other examination tools utilized at the IMU such as modified essay questions (MEQ) would then be required to compensate for the inability of OSPE to test final year students at higher levels of Bloom's taxonomy.

Diagram 1: The Bloom's hierarchy levels of abstraction and the associated mean score.



ABSTRACT NUMBER: OA3

Assessment Of Simulated Patients: Is The Instrument Used Valid And Reliable In The Malaysian Setting?

Nagarajah Lee, Jennifer Perera, Joachim Perera, Juriyah Abdullah

International Medical University, Kuala Lumpur, Malaysia

Introduction

In developing a sound instrument, the researcher needs to understand the exact nature of the problem and the research objectives and then systematically develop the questionnaire taking into consideration the issues of validity and reliability. According to Churchill (1979), when developing a new instrument the systematic process should define the construct to establishing reliability and validity. If the researcher would like to use a psychometrically well established instrument, checking the reliability and validity is the minimum required.

Purpose and Research Design

The purpose was to establish the reliability and validity of the Maastricht University (MaSP) instrument used to assess simulated patients, developed and validated in a Western country. The expert panel validation and questionnaire survey was used as main tools. The expert panel was conducted to establish face and content validity, while data from the questionnaire survey was used to ascertain reliability and construct validity.

Results

Video recordings of nine simulated patients were assessed on two broad aspects: authenticity and attentiveness during consultation and provision of feedback after consultation. The respondents for this assessment comprised of 63 student interviewers, 762 peers, and 54 tutors. The Cronbach alpha coefficient for the various items are greater than 0.650 indicating acceptable reliability. The convergent validity using the correlation analysis between the individual items and the overall assessment score as proposed by Narver & Slater (1990) provided values ranging from 0.318 to 0.612 which indicates good convergent validity. The concurrent validity using the independent sample t-test, comparing the differences in 'individual item scores' and 'overall evaluation score' shows that there are significant differences between the two groups (favourable overall assessment and unfavourable overall assessment). This is evidence for good concurrent validity.

Conclusion

The instrument to assess the simulated patients consists of two dimensions: authenticity and feedback after consultation. The reliability and various validity checks provided evidence that the instrument used to assess simulated patients is reliable and valid to a good extent in the Malaysian setting.

REFERENCE

1. Churchill, G.A, 1979 A Paradigm for Developing Better Measures of Marketing Constructs', *Journal of Marketing Research*, pp. 64 - 73.
2. Narver, J and Slater, S, 1990, The effect of a market orientation on business profitability' *Journal of marketing*, vol.54, pp. 20-3.

ABSTRACT NUMBER: OA4

Performance At Formative and Summative Assessments In Musculo-Skeletal System: A Comparative Study

Mitra N K, Judson J P, Perera, J

International Medical University, Kuala Lumpur, Malaysia

Introduction

Formative assessments help the learners to improve themselves through self reflection and awareness of any gaps between their desired goals and current knowledge or skills^{1,2}. In contrast, summative assessments are conducted to appraise a student's overall performance³. Formative assessments enhance student learning and improve their performance at summative assessments⁴. At the International Medical University, in-course-assessments (ICA) are considered formative as they help the students to prepare for the end of semester (EOS) summative assessment.

Objective

To study the correlation between the students' performance at the musculo-skeletal (MSK) system at the ICA (formative) and the EOS (summative) assessment.

Materials & Methods

The scores obtained by students of two cohorts at the ICA of the MSK system and the MSK component of the EOS were analysed. As both assessments comprised of problem cases (PC) and Objective Structured Practical examination (OSPE) their scores were analysed separately using SPSS 11.5.

Results & Discussion

In the PC at ICA, 70% of students scored marks exceeding 65% whereas at EOS 90% secured marks above 65%. In OSPE at ICA, 79% of students scored marks exceeding 65% whereas at EOS only 54% exceeded the 65% limit. This may be due to students concentrating more in the preparation for the theory component than that of the practical component. Students' inability in applying knowledge in practical situations may be another contributory factor. This is corroborated by the results obtained by analyzing performances in OSPE (in one Anatomy OSPE only 32% secured >65%). Regression analysis showed only 28% and 22% variability in the performance of problem case and OSPE in EOS respectively indicating that the performance at two assessments was not well correlated. Differences in the difficulty levels of questions at two assessments may account for this.

Conclusion

Students' performance in musculoskeletal component of EOS is not well correlated to the performance at in-course-assessment. The overall poor performance at OSPE may be due to the superficial knowledge acquired during the practical component of the course.

REFERENCES

1. Ramaprasad A. On the definition of feedback. *Behavioral Science*, 1983; 28 (1): 4-13.
2. Sadler DR. Formative assessment and the design of instructional systems. *Instructional Science*, 1989; 18 (2): 119-144.
3. McAlpine M. Principles of Assessment. *Bluepaper Number 1* 2002(Glasgow: Robert Clark Centre for Technological Education)
4. Yorke M. Formative Assessment and Its Relevance to Retention. *Higher Education Research and Development* 2001; 20(2): 115-126.

ABSTRACT NUMBER: OA5

The Use Of Objectively Structured Self-Assessment And Peer-Feedback (OSSP) For Teaching Communication Skills; Perceptions Of Students

Jennifer Perera, Satpal Kaur & Galy Mohamadou

International Medical University, Kuala Lumpur, Malaysia

Introduction

Provision of feedback is an important component of effective learning.¹ The usefulness of teacher feedback on improving communication skills (CS) has been well documented. It has been proposed that self-assessment and peer feedback have an equally important role to play in enhancing learning,^{2,3} and may enable students to acquire generic skills on self-assessment and self-regulation of their learning and behaviour. Thus, *objectively structured self-assessment and peer feedback* (OSSP) was used in CS teaching/learning (CSTL) sessions. However, learner acceptability is an important component in any new T/L programme.

Methods

In CSTL sessions it has been the practice for facilitators and simulated patients (SPs) to provide feedback. OSSP was introduced to 96 semester one medical students who learn CS (history taking) in the undergraduate course as an innovation to improve learning. Structured self-assessment and peer-feedback with the aid of a questionnaire consisting of close and open ended questions, designed for the purpose was used to provide OSSP after history taking sessions with SPs. The degree of experience on self-assessment and peer-feedback during pre-university education and perceptions with regard to benefits and acceptability of OSSP in learning, were sought from the 96 students and 84(87.5%) responded. Among the 84 respondents, 59 made free text comments.

Results

With regard to prior experience, 33% of students had never privately assessed either self or peer work in a formal manner and 42% had never assessed either self or peer work publicly during pre-university years. In terms of acceptability, only 10.6% felt shy frequently, 9.6% awkward, 4.8% uncomfortable, 1.2% harassed during OSSP, 70% learnt new aspects with regard to their performance, during both self and peer evaluation, 90% made use of the method during subsequent CSTL sessions conducted in the absence of tutors. Among the 59 free text comments only 3 were negative and team spirit, small group, transparency, accepting constructive criticism, skill of self and peer evaluation were recurring positive themes.

Discussion

It is encouraging to note that in spite of the paltry experience of students with regard to self-assessment and peer-feedback, and when viewed in the cultural context, the learner acceptability of OSSP is very high.

REFERENCES

1. Ende J. Feedback in medical education. JAMA, 1983; 250(6) : 777- 81
2. McDonald B & Boud D The impact of self-assessment on achievement: the effects of self-assessment training on performance in external examinations. Assessment in Higher Education 2003 10 (2), 209 -220
3. Nicol D J & Boyle J T . Peer instruction Vs class wide discussion in large classes: a comparison of two interaction methods in the wired classroom. Studies in Higher Education, 2003 28(4), 457 - 473.

ABSTRACT NUMBER: OA6

Evaluation Of The Quality Of Multiple Choice Questions

Cho Min Naing & Mala Maung

International Medical University, Kuala Lumpur, Malaysia

Background

The appropriately constructed multiple choice questions (MCQs) pose objectivity in marking. However, items are prepared by multiple writers from multiple disciplines, thus the variation in the quality of question items (QI) is a concern. The research questions are, thus, raised: (i) what is the quality of QI used? and (ii) if necessary, how can the quality of QI be improved? Therefore, a study was performed with two main objectives: (i) to investigate the quality of QI using standard indicators, and (ii) to provide suggestions for the improvement of QI.

Methodology

A documentary analysis using the sample of all 50 QI of MCQ in the one-best-answer format, administered on 198 semester-2 medical students from M 1/07 cohort, for end-of course assessment at the International Medical University. The items incorporated 5 disciplines. "Item analysis" was performed using 3 basic indices: (i) the difficulty index (DF), (ii) the discrimination index (DC) and (iii) a combined indicator, *power index* (PI). The simplified formulas were applied for the calculation of these indices.

Results

Out of 50 items, 10 (20%) items were with non-functional distractors (NFD), and 19 (38%) with negative phrases. The mean DF and DC were 0.67 (SD = 0.22) and 0.29 (SD = 0.14), respectively. No negative DC was found. Notably, 14 (28%) items had a PI value above 110, which deserved to be retained for future use. The difference of quality in terms of PI among disciplines was not statistically significant ($\chi^2 = 5.77$; $df = 4$; $p = 0.217$).

Conclusions

The findings suggest re-wording of negatively phrased items, and to improve the options of NFD items rather than a total discard. The items that deserve to be retained in future are indicated. The consistency of the quality is assured among disciplines. A yearly comparison was not attempted as it would have been affected by "selection and/or detection bias". Considering results from all indicators can serve as a guide to the faculties for writing better QI.

ABSTRACT NUMBER: OA7

Outcome Analysis As A Validation Tool For Student Assessment

John Paul Evangel Judson, Khin Win, Joachim Perera, Nilesh Kumar Mitra

The International Medical University, Kuala Lumpur, Malaysia

Introduction

Assessment is an integral part of the medical curriculum and is often the driving tool for learning. Although not always possible, assessments need to be scientific, objective, criterion-based and uniform to ensure fairness as best as possible and to instill confidence among the students. In the selective programme in anatomy, group-based poster presentation was used as a tool to assess the learning outcomes in IMU.

Objective

To determine the outcomes of developing and presenting posters as a learning tool in relation to the expected institutional learning outcomes.

Methodology

Four academics independently assessed the posters presented by 44 students (clustered in 11 groups of 4 students each) during a Selective course in Anatomy. Each poster was graded (on a scale of 1-4) for each of the 8 institutional outcomes.

Results

More than 54% of the posters were in agreement with all IMU outcomes except the clinical skills. This was as expected since the selective programme was not focused on learning clinical skills. The highest agreement was found in communication skills (100%), with professionalism, ethics and personal development scoring 91% each. The communication was assessed by evaluating the scores given by the judges at the poster presentation. Least agreement was found in the application of basic science in clinical medicine and disease prevention and health promotion (55%). This may be due to the fact that the poster topics were not focused on these specific outcomes. The inter-rater correlation ranged from adequate to good (Alpha 0.67 to 0.93) ($p < 0.005$).

Conclusion

The study showed that the curricular content of topics were in agreement with majority of the expected curricular outcomes. It also showed that posters can be used as a reliable assessment of learning tool to achieve outcomes such as professionalism, ethics and personal development which are difficult to assess within the conventional curriculum. Further improvements can be achieved by check listing the institutional outcomes against the learning outcomes when selecting and developing topics for posters.