Do Direct Observation Tools Generate High Quality Narrative Feedback? Lessons from the Psychopharmacotherapy-Structured Observation Tool

AUTHOR(S):
- John Young, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, USA (Presenter)
- Patricia O'Sullivan, UCSF School of Medicine, USA
- Rebekah Sugarman, Zucker School of Medicine at Hofstra/Northwell, USA
- Eric Holmboe, ACGME, USA

ABSTRACT

Background: Performance assessment based on direct observation has become increasingly important. While many direct observation tools include both rating scales and space for narrative comments, research has largely focused on the quantitative scores generated by the rating scales. Much less is known about the quality of the comments and their relationship to the quantitative scores. This presentation explores these questions via analysis of the comments from the Psychopharmacotherapy-Structured Clinical Observation (P-SCO) tool, a direct observation tool that facilitates assessment of a medication management visit.

Summary of Work: 601 P-SCOs were completed over four academic years in an outpatient psychiatry clinic. 25% of the completed P-SCOs from each academic year were randomly selected. To assess quality, comments were coded for valence (reinforcing or corrective), behavioral specificity, and content. To assess the relationship between comments and scores, the authors calculated the frequency with which a high or low checklist score was accompanied by a comment with the same content and valence as well as the frequency with which a reinforcing or corrective comment was accompanied by a high or low checklist score.

Summary of Results: 91% of the comments were behaviorally specific. 60% were reinforcing, 40% corrective. 8 themes were identified, including two constructs not adequately represented by the checklist. 7% of the low and high checklist scores were associated with a comment of the same valence and content. Only 50% of the overall comments were associated with a checklist score of the same valence and content.

Discussion and Conclusions: The P-SCO, as our exemplar, generated written comments that were based on direct observation, behaviorally specific, clinically relevant, with a balance of corrective and reinforcing feedback - all attributes associated with effective written assessment. Narrative comments both expanded on checklist scores and contributed new content.

Take-home Messages: A direct observation tool such as the P-SCO generates high quality written comments. Narrative comments complement the quantitative scores and should be included in direct observation tools. Moreover, thematic coding of comments can improve the content validity of a checklist and identify weakness in the curriculum.
#81 Short Communications - Assessment - Work-Based Assessment

812 (1314)

Date of Presentation: Tuesday, 27 August 2019

Time of Presentation: 1415-1430

Location of Presentation: Room L2, Level 1

An Innovative and Comprehensive Workplace Based Assessment Framework for Australian General Practice

AUTHOR(S):
- Jill Benson, GPEx, Australia (Presenter)
- Lambert Schuwirth, Prideaux Centre for Research in Health Professionals Education, Australia
- Emily Kirkpatrick, GPEx, Australia
- Shaun Prentice, GPEx, Australia
- Taryn Elliott, GPEx, Australia

ABSTRACT

Background: Work-place based assessments (WBAs) are now accepted as useful and acceptable formative, summative and programmatic assessments in most medical education contexts. They aim to assess the peak of Miller’s pyramid, that the trainee ‘does’ what they are expected to do in a safe and competent manner. But do they really achieve this?

Summary of Work: In order to answer this question, the Royal Australian College of General Practitioners funded a national Education Research Grant through GPEx, the South Australian General Practice (GP) training organisation, in collaboration with the Prideaux Centre at Flinders University, and the GP Training Organisations in every state of Australia. The GPEx team developed a multi-faceted research project consisting of: a 3-cycle hermeneutic literature review with 4 coders; focus groups and interviews with supervisors, registrars and medical educators in every state, on their opinions and experience of WBAs; data collection about how WBAs are performed in each state with analysis using a relational data-base management tool; and 3 nested research projects using quantitative and qualitative data collection on the use of Entrustable Professional Activities, Direct Observation Visits by clinical teachers, and the process of ‘flagging’ registrars with subsequent remediation, and whether this impacts on exam results.

Summary of Results: A WBA framework was developed from the research project, recommending the tools and activities that are more likely to be successful. This includes use in different General Practice contexts (eg rural and remote) and for a variety of purposes such as formative, summative, programmatic and remediation.

Discussion and Conclusions: Recommendations for implementation include: what, who, how, when, where of WBA tools; the training of the assessors and trainees; the support from the training organisation; the relationships between the assessor, the registrar, the tool and the training organisation; and the processes and systems that need to be in place.

Take-home Messages: This evidence-based, innovative and comprehensive WBA framework is not only relevant in Australian General Practice but has the flexibility to be applied across multiple contexts.
How students and specialists appreciate the Mini-Clinical Evaluation Exercise (Mini-CEX) in Indonesian clerkships

AUTHOR(S):

- Yoyo Suhoyo, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Indonesia (Presenter)

ABSTRACT

Background: Cultural differences might challenge the acceptance of the implementation of assessment formats that are developed in other countries. Acceptance of assessment formats is essential for its effectiveness; therefore, we explored the views of students and specialists on the practicability and impact on learning of these formats. This study was conducted to explore Indonesian students' and specialists' appreciation of the implementation of the Mini-Clinical Evaluation Exercise (Mini-CEX) in Indonesian clerkships.

Summary of Work: This study was conducted at the Universitas Gadjah Mada, Indonesia. Participants were 52 students and 21 specialists in neurology and 78 students and 50 specialists in internal medicine. They were asked to complete a 19-item questionnaire that covered the characteristics of the mini-CEX such as observation and feedback, its practicability, and the impact on learning and professional development. We used a Mann-Whitney U test to analyse the data.

Summary of Results: In total, 124 students (46 from neurology and 78 from internal medicine) and 38 specialists (13 from neurology and 25 from internal medicine) participated in this study. Students and specialists were positive about the practicability of the mini-CEX and the impact of this assessment format on learning and on professional development. The Mann-Whitney U test showed that there were no significant differences between students' and specialists' opinions on the mini-CEX, except for 2 items: specialists' appreciation of direct observation (mean rank=93.16) was statistically significantly higher than students' appreciation of it (mean rank=77.93; z=2.065; p<0.05), but students' appreciation of the item that students' past mini-CEX results affected their recent mini-CEX outcomes (mean rank=85.29) was significantly higher than specialists' appreciation of it (mean rank=69.12; z=2.140; p<0.05).

Discussion and Conclusions: Students and specialists were positive about the mini-CEX in Indonesian clerkships, although it was developed and validated in another culture. We found only small differences between their appreciations, which could be explained by the patterns of specialist-student interaction in Indonesian culture as large power distance and low individualism country.

Take-home Messages: Culture is not necessarily an obstacle when implementing educational principles stemming from countries with a different culture. We just need to take into account the characteristics of the local culture.
How do postgraduate surgeons-in-training learn through the use of workplace-based assessment?

AUTHOR(S):

- Arpan Tahim, UCL, UK (Presenter)
- Deborah Gill, UCL, UK
- Jeff Bezemer, UCL, UK

ABSTRACT

Background: Existing research on workplace-based assessment (WBA) in surgical education has focused on their outcomes (for example, how well a learner performs) and user perception (such as, how participants feel about it). There is limited understanding of what actually happens during WBAs and therefore no empirical basis for explaining either outcomes, users’ perceptions, or their impact on learning. Recently it has been proposed that re-imaging WBA practices as intricate, real-time, social phenomena may be a more appropriate way to interpret these complicated interactions. Taking this socio-cultural view may help us understand the context, content and process of the WBA and help explore what use learners make of them in their education.

Summary of Work: This study explores the nature of workplace-based assessment in UK postgraduate surgical training, through a socio-cultural lens, in relation to the following questions; i) What does a WBA look like in situ, in the context of surgical education, ii) What does a surgeon-in-training identify has having learnt from a particular WBA, and iii) How do these learners respond to it in their future practice? Data was generated from a cohort of surgeons-in-training using video-recording of in situ WBA events, the learner’s formal WBA reports, and through reflexive interviews conducted with the learner.

Summary of Results: Video-ethnographic and -reflexive techniques were used to explore the these workplace learning phenomena through a socio-cultural lens in order to examine the unique context, content, process and impact of these events.

Discussion and Conclusions: Viewing WBAs through a socio-cultural lens provides new understandings of their function as a surgical educational tool, which could improve their future use. More broadly, the complex learning needs that surgery shares with other professions should not go unrecognised. In fact, surgical education could represent an important ‘telling case’ to explore workplace learning and assessment in greater detail. Insights gained could improve education practice and shed light on the potential of assessment in the wider workplace.

Take-home Messages: WBAs are unique and individual experiences and may be better suited to interpretation through socio-cultural learning theories.
Practical Strategies for Evaluating and Improving the Content of a Workplace-based Observational Assessment

AUTHOR(S):
- Mark Raymond, National Board of Medical Examiners, USA (Presenter)
- Melissa Margolis, National Board of Medical Examiners, USA
- Brian Clauser, National Board of Medical Examiners, USA
- Marcia Winward, National Board of Medical Examiners, USA
- Alan Schwartz, UIC Department of Medical Education, USA
- Carol Carraccio, American Board of Pediatrics, USA
- Patricia Hicks, University of Pennsylvania, USA

ABSTRACT

Background: Observing and documenting the performance of medical trainees in the clinical workplace allows for collecting important evidence about competence in critical domains. Unfortunately, the resulting data often are subjective and unreliable because the instruments used to collect these behavioral data typically do not go through the robust development and revision process that is common for professionally developed tests. The present research addresses this problem by describing practical approaches that can be used to evaluate (and ultimately improve) the quality of observational assessment instruments.

Summary of Work: Three different observational assessment instruments were used to collect data on 258 first-year Pediatrics residents from 10 United States Pediatrics residency programs. A total of 2,591 instruments were completed by Faculty, Residents, Nurses, and other members of the healthcare team. Statistical (descriptive, correlational, and generalizability) analyses were used to evaluate item and instrument performance.

Summary of Results: Results indicated that (1) observers often did not use the full response scale; (2) some items did not perform well either because they were highly correlated with other items or had near zero correlations with other items; and (3) reliability varied considerably across observer groups (.76 for residents, .66 for faculty).

Discussion and Conclusions: The quality of an assessment instrument is directly related to confidence in the outcomes. This practical approach to instrument evaluation allowed for identification of items that did not perform well and informed revisions (i.e., scale changes, revision or deletion of items that were unrelated or too closely related to other items) that will improve instrument quality and by extension will increase confidence in the resulting data.

Take-home Messages: Data collected via observational assessment in the authentic clinical environment can provide important evidence about the competence of medical trainees. The present research demonstrates that relatively simple statistical procedures can be used to guide evaluation of item and instrument performance. This has the potential to substantially improve the reliability of scores from observational assessment instruments.
#8I Short Communications - Assessment - Work-Based Assessment

816 (2329)

**Date of Presentation:** Tuesday, 27 August 2019  
**Time of Presentation:** 1515-1530  
**Location of Presentation:** Room L2, Level 1

**Who's Watching? Bedside Observation and Feedback Practices in Internal Medicine**

**AUTHOR(S):**
- Michael Wang, McMaster University, Canada (Presenter)
- Kristyne Onizuka, McMaster University, Canada
- Christopher Foster, Western University, Canada
- Ramy Khalil, Queens University, Canada
- Daniel Brandt Vegas, McMaster University, Canada

**ABSTRACT**

**Background:** Opportunities to perform work-based assessments (WBAs) frequently occur during bedside teaching (BST). However, observation and feedback practices at the bedside are not well documented in the literature. Our objective was to assess the current use of bedside teaching, observation, and feedback during internal medicine rotations.

**Summary of Work:** We conducted a longitudinal survey study at five academic-affiliated hospitals associated with McMaster University. Medical students and residents completing inpatient internal medicine ward rotations between October 2017 and May 2018 were included. Survey items were constructed around three themes - bedside teaching, direct observation of bedside skills, and subsequent feedback. The survey encompassed bedside experiences over the past one week. We quantified the frequency of observation and feedback by the type of clinical skill, and utilized a 5-point Likert scale to measure the overall perceived quality of feedback.

**Summary of Results:** The response rate was 63% (n = 192/304). Learners received BST on 4 patients each week, and was often led predominantly by senior residents (43% of respondents). One-fifth of respondents reported receiving no BST, while a quarter reported receiving no direct observation. Bedside assessment of discharge planning and history taking occurred least frequently, with three-quarters of respondents (75% and 78%, respectively) reporting no observations. When observation of history taking and discharge planning occurred, feedback was provided to less than half of learners (44% and 42%, respectively). Attending physicians offered feedback more regularly than senior residents (p = 0.003), though there was no difference in its perceived quality.

**Discussion and Conclusions:** Direct observation at the bedside remains infrequent and inconsistently delivered. Ensuring routine feedback delivery after observation remains a challenge. Observation of history taking was particularly neglected, despite its importance in forming accurate diagnoses. Residents often led bedside teaching, and may provide a similar quality of feedback to learners compared to attending physicians.

**Take-home Messages:** Bedside observation and feedback remain under-utilized tools for learner assessment. Low rates of assessment were particularly noted for several core clinical skills, including history taking. Training programs transitioning towards CBME should focus on ensuring adequate opportunities for direct observation and feedback of these skills.