**9D1 (18587)**

**Post-Graduate Adolescent Interviewing Skills: A Reflection of the Sustainability of Structured Formal Undergraduate Training**

*Genna Bourget*, University College Dublin, Dublin, Ireland  
Nadim Joukhadar, Dalhousie University, Halifax, Canada  
Sarah Manos, Dalhousie University, Halifax, Canada  
Karen Mann, Dalhousie University, Halifax, Canada  
Kim Blake, Dalhousie University, Halifax, Canada

**Introduction:** In adolescence, individuals begin to assume responsibility for their own health care. Hence, effective adolescent interviewing and communication by physicians are critical. Education in this area occurs usually in undergraduate medical education (UME); whether these skills are maintained in residency is unclear. Our objective was to determine if entry level residents (PGY1) who had received formal adolescent interview training (with standardized patients and structured feedback) in UME achieved higher scores on a validated rating scale, the Structured Communication Adolescent Guide (SCAG), than those who had not received formal adolescent interview training.

**Methods:** PGY1s, including international medical graduates, were recruited. Each participant conducted an adolescent interview with a standardized adolescent patient and mother pair (SPs). The themes of the patient case focused on sensitive subjects to adolescents, specifically sexual orientation and bullying. The SPs separately scored residents on their interview using the 29 item SCAG comprised of four sections (Getting started, Gathering Information, Teen alone, and Wrap up) each with a total item score and global score. Unpaired t-tests were conducted to compare the total item and global SCAG scores of the 'no formal training' group against the 'formal training' group, using the SP Daughter score and Mother score separately. Unpaired t-tests were also conducted to determine if there were statistical differences between daughter and mother scoring.

**Results:** PGY1's who had received previous formal training (n = 23) had statistically significantly higher scores than those without (n= 29) on both the Total Item scores (maximum 58) and the Global scores (maximum 40). No Formal Training  
SP Adolescent: Total-Item Score Mean (SD) 32.41 (10.12)  
Global Score Mean (SD) 26.10 (6.22)  
Formal Training  
SP Adolescent: Total-Item Score Mean (SD) 40.78 (7.04)  
Global Score Mean (SD) 30.83 (3.33)  
Significance: Total-Item Score P<0.0013  
No Formal Training  
SP Mother: Total-Item Score Mean (SD) 40.48 (7.90)  
Global Score Mean (SD) 29.61 (5.92)  
Significance: Total-Item Score P<0.0347

**Discussion:** Our results suggest that PGY1s who had formal adolescent training in UME had retained the knowledge base and communication skills to interview an adolescent as they entered residency. These skills can be further built on in residency. Agreement between the mother and daughters’ scores provides inter-rater reliability to the scale, further validating the tool.

**Conclusion:** We demonstrated, using an established scale, that there was a sustainable effect of structured training in adolescent interviewing into postgraduate performance. Our findings support the need for formal adolescent training in UME. We suggest that PGY1s with no formal adolescent training should receive SP interviewing with adolescent feedback.

**References:**


---

**9D2 (18902)**

**Checking the checklist: A comparison of case-specific checklists developed by different expert panels**

*Agatha M. Hettinga*, Radboud University Medical Center, Nijmegen, Netherlands  
Cornelis T. Postma, Radboud University Medical Center, Nijmegen, Netherlands

**Introduction:** Clinical competences are commonly assessed by objective structured clinical examinations (OSCEs). A key aspect of the OSCE is the development of a checklist or rating scale. Often expert panels play an important role in the evolution of case-specific checklists. Our research question is: To what extent do case-specific checklists developed by different expert panels agree?

**Methods:** An OSCE scenario was presented to four different expert panels from different faculties. Effort was made to gather four experts for each expert panel. Each panelist was asked to prepare the meeting. Well before the expert meeting the experts were sent the case, they were asked to carefully read the case, and to send a checklist with items for history...
taking, and physical examination. A panel leader ordered the different items. During the meeting the panelists were presented the list of items of their panel and asked to discuss the various checklist items and to develop a final checklist with 10-15 items for history taking and 10-15 items for physical examination. The four checklists for one scenario were compared. This procedure was performed for three common clinical scenarios.

Results: Results are presented in tables, showing different categories of checklist-items, namely ‘items included by all panels’, ‘items included by 3 panels’, ‘items included by 2 panels’ and ‘items included by 1 panel’, for all three scenarios. For each checklist the percentage items on that list that was included by all panels was calculated. The lowest percentage found was 14%, and for the checklist with the highest percentage of items that were included by all panels the percentage was 55%, with an average of 40%. The average percentage of items included by all panels or all but 1 panel was 70%.

Discussion and Conclusions: This study has shown that different expert panels develop different case-specific checklists for OSCE scenarios. One important question is whether the ranking of students changes using different checklists. Further research is needed to better understand the meaning of these differences.


9D3 (18661)
Validation of an Instrument for Measuring the Quality of Teamwork in Teaching Teams (Team Q)

Irene A. Slootweg*, University of Maastricht/Academic Medical Centre, Department of Educational Development and Research/Professional Performance Research Group, Maastricht/Amsterdam, Netherlands
Kiki Lombarts, Academic Medical Center, Professional Performance Research Group, Amsterdam, Netherlands
Benjamin Boerebach, Academic Medical Centre, Professional Performance Center, Amsterdam, Netherlands
Albert Scherbier, University of Maastricht, Faculty of Health, Medicine and Life Sciences, Maastricht, Netherlands
Cees Vleuten, van der, University of Maastricht, Department of Educational Development and Research, Maastricht, Netherlands

Introduction: Tackling the issue of teamwork between clinical teachers is one of the challenges in reforming high-quality postgraduate medical training. Clinical teachers should ideally share a mutual vision and arrive at specific agreements to supervise residents. Information that provides insight in these mutual visions and agreements is an essential first step in improving teamwork. Few evaluation instruments are available for measuring teamwork in health care settings, however none of them was considered suitable, in terms of aim and use of specific language, for measuring teamwork in a teaching team[1]. The aim of this study is to investigate the validity and generalizability of a tool for measuring teamwork in teaching teams (the TeamQ instrument). The research questions are: 1) what are the psychometric properties of the TeamQ instrument and 2) how many evaluations are needed per teaching team to generate a reliable teamwork measure.

Methods: The initial TeamQ instrument consisted of teamwork items that were selected based on a focus group and Delphi study. [2] From January 2012 till December 2013 we tested these items in teaching teams of multiple specialties in multiple organizations. In total, 1446 clinical teachers from 116 teaching teams were invited to complete the web-based TeamQ instrument. First, we conducted a principal component analysis to extract the number of factors (composite scales) underlying the TeamQ items. Second, the internal consistency reliability coefficient Cronbach’s alpha was calculated for each composite scale. Third, we calculated the number of clinician teacher evaluations needed to obtain reliable measures of the TeamQ scales (generalizability analysis).

Results: Data of 114 teaching teams were included in our study. The median response rate was 7 evaluations per team. The sample contained 48 small groups (<10), 54 medium sized groups (10-20) and 14 large groups (>20). The sample included surgical (n=45) and non-surgical specialties (n=53) as well as auxiliary disciplines (n=18). The PCA revealed a 8-factor structure of the TeamQ questionnaire. The factors were labelled as Task-expertise; Team-expertise; Decision-making; Leadership; Feedback-culture; Team-results; Residents'engagement; Residents' empowerment. The scales counted 3 to 11 items per scale. The reliability of 7 TeamQ scales ranged from 0.75 for Decision-making to 0.93 for Leadership. The scale Residents'empowerment had a lower reliability coefficient of 0.66. The generalizability analysis revealed that 5 to 7 evaluations were needed to obtain reliability coefficients of 0.70 for all scales and 6 to 8 evaluations were needed to obtain a reliability of 0.80.

Discussion and Conclusions: This study provides a first indication of the validity of a new tool for measuring teamwork in teaching teams. The high response rates and the low number of evaluations needed for reliably measuring teamwork indicate the feasibility of the TeamQ instrument in the evaluation of teamwork in teaching teams in practice. For high quality residency training teams of clinical teachers may use the TeamQ results to reflect on and potentially improve the quality of the different aspects of teamwork. Future research could focus on further exploring the use of the TeamQ instrument for teaching teams of different sizes.
References:

9D4 (18856)
Validity: One word with a plurality of meanings

Christina St-Onge*, Université de Sherbrooke, University of British Columbia, Sherbrooke, Canada
Meredith Young, McGill University, Montreal, Canada
Kevin Eva, University of Toronto, Vancouver, Canada
Brian Hodges, Université de Sherbrooke, Toronto, Canada

Introduction: Compared to recommended standards, several works have claimed “less than optimal validation practices” are present in medical education research. Whether such disconnects between recommended practice and published work reflect genuine disagreements between practitioners and theoreticians, a lag in the uptake of the recommended standards, or implicit differences in different groups’ fundamental conception of validity, remains to be determined. However, since validity as a concept is expected to inform assessment development, implementation, and interpretation, it is critical to investigate the mechanism whereby such disconnects arise. The primary purpose of this study was to identify ways in which the term validity is used within medical education research to examine the potential for variable, implicit and sometimes unavowed conceptions of validity.

Methods: We used discourse analysis “to dissect, disrupt and render the familiar strange” (Graham, 2005, p.4) through an interrogation of the medical education literature that examined the ways in which the concept of validity is presented. The aim was to document the multiple possible “truths” surrounding the concept of validity in this field. We used a purposeful sample of references identified by a search of PubMed, ERIC, PsychINFO and the authors’ accumulated references. An iterative process was used to discuss and identify emerging discourses until consensus among the research team was achieved.

Results: Three discourses were identified. Validity as a test characteristic is underpinned by the notion that validity is an intrinsic property of an assessment instrument itself. Adopting this discourse makes it possible to label “tools” as “gold standard” and implies that validity claims apply across contexts. Validity as a hypothesis-driven evidentiary-chain emphasizes the importance of supporting the interpretation of assessment results with sound evidence and continuous analysis. In this discourse, validity does not belong to the tool/instrument itself, but reflects the process used for validation. Validity as a social imperative takes into account the downstream effect of assessment (e.g. a ‘fail’ decision), positive or negative, on assessees. Here, validity is called upon to justify the decision process regarding the consequences for assessees and the social implications of assessment.

Discussion and Conclusions: Validity can be called upon to signify that a tool can or should be used because it meets a certain standard. For others it speaks to a continuous quality assurance process put in place to justify use of a tool or to mechanism to ensure that the assessment process maintains social good. Independent of the underlying conceptualization, validity is a commonly-used and highly-loaded term in medical education. The observed discourses may explain – in part – the observed discrepancies between recommended and adopted validation practices. Each may fill very different and sometimes implicit needs within medical education. Validation appears to have several meanings in medical education; further research is needed to better understand who participates in each discourse and how as well as to understand the impacts of adopting different conceptions of validity on the development and monitoring of assessment.


9D5 (18840)
Junior doctor intercultural clinical communication: Lessons for transition to practice medical education

Robyn Woodward-Kron*, University of Melbourne, Melbourne Medical School, Melbourne, Australia
Hashrul Nizam, Goulburn Valley Health, Melbourne, Australia
Jonathon Au, Goulburn Valley Health, Shepparton, Australia

Introduction: In many Western countries, junior doctors including International Medical Graduate (IMG) doctors and medical students are increasingly culturally diverse due to factors such as globalisation of the tertiary education system and health workforce shortages. In migration and refugee destination countries such as Australia, these graduates are delivering healthcare in a context where a significant proportion of the population may not have the English language skills to communicate effectively about their healthcare. Effective communication with patients from different language and cultural backgrounds is acknowledged as a desirable graduate attribute internationally in numerous medical education settings and communication curricula. In a study by junior doctors from the U.K., the U.S., and Europe, the authors point out that one barrier to the transfer of communication skills learned at medical school is the cultural heterogeneity of simulated patients compared to the cultural and social diversity of patients they
encounter in their training environments (1). Medical educators who are concerned with cultural and language aspects of healthcare communication seldom investigate or conceptualise communication in healthcare settings where cultural and linguistic diversity is commonplace amongst doctors, medical students, and patients. The aim of this study was to investigate junior doctor clinical communication in a culturally diverse healthcare setting, including investigating communication behaviours, challenges, and enabling strategies that the junior doctors engage in when delivering healthcare in a setting where patient and staff cultural diversity is commonplace.

Methods: We adopted a qualitative research design including focus group discussions with junior doctors and semi-structured interviews with senior hospital staff employed at one regional hospital. There were 5 focus group interviews with 20 junior doctors in total as well as semi-structured interviews with 10 senior doctors and 4 senior administrative staff. These data were analysed thematically.

Results: The two major themes to emerge in the discussion about intercultural communication with the junior doctors were language as a barrier and cultural influences on healthcare communication. The sub-themes for language as a barrier were fluency, impact, and mediated communication. The sub-themes for cultural influences were uncertainty and cultural comfort. The results for the semi-structured interviews with senior clinical and hospital staff were likewise language as a barrier and cultural influences on healthcare communication. Additional themes were clinical skills, including communication, and patient education and counseling.

Discussion and Conclusions: The findings suggest that while junior doctors have some strategies to address communication challenges such as language barriers and cultural sensitivities, they can struggle with cultural differences in patient expectations of healthcare and expression of symptoms. Senior medical staff reported that there is room for improvement, particularly in explaining the diagnosis and management plan to patients. There appears to be a disconnect between how well the junior doctors think they communicate and how the senior medical and hospital staff evaluate the junior doctors’ communication skills. Cultural and language barriers to effective communication can also be considerable. The findings form a needs analysis for culturally effective communication to inform transition to practice medical education.