**DOCS**

**September 2019 Journal Watch**

Articles included in review:

*Academic Medicine -* Volume 94(8) pgs. 1061- 1255 August 2019

*Teaching and Learning in Medicine -* Latest Articles

*Medical Teacher ­*– Volume 41 (8)

*Medical Education­* *­*– Volume 53(8)

### **Can Change to Clerkship Assessment Practices Create a More Equitable Clerkship Grading Process?**

### Authors: Teherani, Arianne PhD; Hauer, Karen E. MD, PhD; Lucey, Catherine MD

Academic Medicine DOI: 10.1097/ACM.0000000000002824 Accession: 00001888-201909000-00006

**Abstract:**

We implemented changes to our clerkship student assessment process aimed at creating equity for our underrepresented in medicine (UIM) medical students. The concept of "differential attainment" addresses how educational outcomes diverge for different demographic groups assessed in the same way. We recently reported differences in scores on assessments at our institution, the University of California, San Francisco, School of Medicine, a public, research-intensive health sciences university, that consistently favored not-underrepresented in medicine (not-UIM) students.1 These assessment differences narrowed throughout medical school. These small differences nonetheless produced what we called the "amplification cascade"-small differences in assessed performance that yield larger differences in grades and selection for awards.1 The amplification cascade creates inequities with serious consequences, such as diminished opportunities to compete for selective residency programs and enter academic careers. The amplification cascade raises questions about clerkship grading policies at our institution that also concern medical schools nationwide.

To address these inequities, leaders at our institution brainstormed countermeasures to differential attainment to prevent the amplification cascade. We focused on a coordinated and integrated system of change in our clerkship assessment methods starting in 2016. First, we introduced grading committees into all core clerkships to mitigate unchecked bias by single supervisors or clerkship directors. Second, we eliminated a threshold examination score to qualify for clerkship honors, a change subsequently endorsed in the literature. Third, we liberalized the honors cap per clerkship from 25% to 45% to provide more opportunities for all students to attain honors.

To determine whether there was greater equity in the number of honors received by UIM students, we examined 3 years preclerkship (N = 501) and 2 years postclerkship (N = 307) assessment countermeasure changes. Post changes, both UIM and not-UIM students received more honors. In the 3 years preceding clerkship assessment changes, UIM students earned on average 17% and not-UIM students earned 34% honors grades. In the 2 years since implementing our clerkship assessment changes, UIM students received on average approximately 32% and not-UIM students 46% honors grades (P = .001-.05). Although differences in attaining honors narrowed slightly postclerkship after assessment changes, the results were similar and smaller than expected. We continue to consider ongoing contributors to this discrepancy and strategies to optimize equity in medical student assessment.

## Implicit Bias and the Feedback Paradox: Exploring How Health Professionals Engage With Feedback While Questioning Its Credibility

Authors: Sukhera, Javeed MD, PhD, DABPN, FRCPC; Wodzinski, Michael; Milne, Alexandra RN; Teunissen, Pim W. MD, PhD; Lingard, Lorelei PhD; Watling, Chris MD, PhD, FRCPC

Academic Medicine. Volume 94(8), August 2019, p 1204-1210

DOI: 10.1097/ACM.0000000000002782 Accession: 00001888-201908000-00038

|  |
| --- |
| **Abstract:** |
| Purpose: Learners and practicing health professionals may dismiss emotionally charged feedback related to self, yet little research has examined how to address feedback that threatens an individual's identity. The implicit association test (IAT) provides feedback to individuals regarding their implicit biases. Anticipating feedback about implicit bias might be emotionally charged for mental health professionals, this study explored their experience of taking the IAT and receiving their results, to better understand the challenges of identity-threatening feedback.  Method: The researchers sampled 32 psychiatry nurses, psychiatrists, and psychiatric residents at Western University in Ontario, Canada, after they completed the mental illness IAT and received their results. Using constructivist grounded theory, semistructured interviews were conducted from April to October 2017 regarding participants' experience of taking the IAT. Using constant comparative analysis, transcripts were iteratively coded and analyzed for results.  Results: While most participants critiqued the IAT and questioned its credibility, many also described the experience of receiving feedback about their implicit biases as positive or neutral. Most justified their implicit biases while acknowledging the need to better manage them.  Conclusions: These findings highlight a feedback paradox, calling into question assumptions regarding self-related feedback. Participants' reactions to the IAT suggest that potentially threatening self-related feedback may still be useful to participants who question its credibility. Further exploration of how the feedback conversation influences engagement with self-related feedback is needed. | |

1. **The Hidden Curriculum of Compassionate Care: Can Assessment Drive Compassion?**

Wright, Sarah R. MBA, PhD; Boyd, Victoria A. MPC; Ginsburg, Shiphra MD, PhD, FRCPC

Academic Medicine. Volume 94(8), August 2019, p 1164-1169.

DOI: 10.1097/ACM.0000000000002773

Accession: 00001888-201908000-00033

**Abstract:**

Purpose: Medical schools are expected to promote compassionate care among learners. Assessment is a key way to communicate values to learners but can create a hidden curriculum. Assessing compassionate care is challenging; however, not assessing it can communicate to students that such care is not valued. The purpose of this study was to explore how current assessment strategies promote or suppress the idea that caring behaviors are valued learning objectives.

Method: Data sources were third-year course documents; interviews of 9 faculty, conducted between December 2015 and February 2016; and focus groups with 13 third-year medical students and an interview with 1 third-year medical student, conducted between February and June 2016. The stated intentions of third-year assessments were compared with the behaviors rewarded through the assessment process and the messages students received about what is valued in medical school.

Results: Syllabi did not include caring as a learning outcome. Participants recognized assessment as a key influence on student focus. Faculty perspectives varied on the role of medical schools in assessing students' caring and compassion. Students prioritized studying for assessments but described learning about caring and compassion from interactions such as meaningful patient encounters and both positive and negative role models that were not captured in assessments.

Conclusions: Faculty members expressed concern about not assessing caring and compassion but acknowledged the difficulty in doing so. While students admitted that assessments influenced their studying, their reported experiences revealed that the idea that "assessment drives learning" did not capture the complexity of their learning.

1. **The Emotional and Behavioral Impact of Delivering Bad News to Virtual versus Real Standardized Patients: A Pilot Study**

Authors: Sarah R. O’Rourke, Kelly R. Branford, Taylor L. Brooks, Lindsay T. Ives, Arjun Nagendran & Scott N. Compton

Teaching and Learning in Medicine. Published online: 22 Aug 2019DOI: 10.1080/10401334.2019.1652180

## Abstract:

Phenomenon: Virtual standardized patients (vSPs) are becoming increasingly common in medical education, though one limitation of vSPs is the artificiality of computer-based simulators. Past research on the use of vSPs has not clearly established whether learners have different emotional responses to real SPs (rSPs) compared with vSPs; however, understanding learners’ emotional responses to vSPs is important in providing realistic learning experiences and establishing the validity of this teaching and assessment tool. This study compared the emotional experiences of individuals who interacted with rSPs and vSPs. Approach: Sixty medical students at a medical school in the southeastern United States participated in the study. Participants were randomly assigned to deliver bad news to an rSP or vSP. The vSP for this study used a hybrid intelligence model that allowed a person to “inhabit” the vSP. Salivary cortisol and a self-report measure of mood—the Profile of Mood States, Second Edition (POMS 2)—were gathered before and after delivering the bad news. The SP and 2 independent evaluators rated the behavioral performance of each participant in real and virtual conditions. Participants also rated the performance of the SP. Findings: Participants in both conditions reported increased negative emotionality on the POMS 2 following the SP interaction. There were no significant between-group differences on the POMS 2 or salivary cortisol concentration following the SP interaction. Ratings by the SP and independent evaluators indicated that participants performed similarly on most interpersonal dimensions, except tone of voice. Participants perceived the vSP as less realistic than the rSP. Insights: These results suggest that medical students may have similar emotional and behavioral responses when delivering bad news to a vSP when compared to an rSP. These findings provide support for the continued use of vSPs in training learners to deliver bad news and other communication-based skills and to assess their performance on these tasks.

1. **The Facilitator Instrument for Team-Based Learning (FIT)**

Authors: Ruth E. Levine, Peggy Hsieh, P. Adam Kelly, Lisa Carchedi, Jennifer Gibson, Paul Haidet, Paul Koles, Lindsey Pershern, Dawnelle Schatte, Brenda Talley, Dwight Wolf & Britta Thompson

Teaching and Learning in Medicine. Published online: 07 Aug 2019. DOI: 10.1080/10401334.2019.1646653

**Abstract:**

Construct: We sought to evaluate the quality of Team-Based Learning facilitation in both large and small group settings. Background: Team Based Learning (TBL) is an increasingly popular small group instructional strategy in health science education. TBL facilitation skills are unique and differ from those needed to lecture or facilitate other types of small groups. Measuring facilitation skills and providing feedback to TBL instructors is important, yet to date no valid instrument has been developed and published for this purpose. Approach: We created an 11-item instrument (ratings of each item on a 7-point scale) designed to assess TBL facilitation skills, considering major sources of validity. Twelve experts in TBL facilitation and training developed the content of the FIT. To ensure response processes were valid, we used an immediate retrospective probing technique with 4th year medical students who were not part of the study. The Facilitator Instrument for Team-Based Learning (FIT) was piloted with 2,840 medical students in 7 schools in large (year 1 and 2) and small (year 3) courses. The internal structure of the FIT was analyzed. Results: In total, 1,559 and 1,281 medical students in large and small TBL classes, respectively (response rate 88%) rated 33 TBL facilitators. The composite mean score for the FIT was 6.19 (SD = 1.10). Exploratory factor analysis and Cronbach’s alpha indicated that all items loaded on 1 factor, accounting for 77% of the item variance. Cronbach’s alpha for the 11 items was 0.97. Analysis of facilitator variables and course context indicated that FIT scores were statistically significantly correlated with type of class (pre-clinical or clinical) and size of class as well as the facilitator enjoyment in using TBL as a method. Gender and the amount that facilitators used TBL each year was weakly correlated, with other factors not correlated (years facilitating TBL, confidence in facilitating TBL, and age). Conclusions: Analysis of FIT scores from 2,840 medical students across multiple institutions and teaching settings suggests the utility of the FIT in determining the quality of TBL facilitation across a range of medical education settings. Future research is needed to further analyze course contexts and facilitator variables that may influence FIT scores with additional facilitators. Additionally, FIT scores should be correlated with additional measures of TBL facilitator quality, such as direct observations, especially if these data are used for summative decision-making purposes.

1. **Minimal effects of reduced teaching hours on undergraduate medical student learning outcomes and course evaluations**

Authors: Derek L. Choi-LundbergORCID Icon, Hayder A. Al-AubaidyORCID Icon, John R. Burgess, Christine A. CliffordORCID Icon, William A. CuellarORCID Icon, Judi A. Errey, Amanda J. HarperORCID Icon, Roslyn C. MalleyORCID Icon, Renee M. RossORCID Icon, Anne-Marie M. WilliamsORCID Icon & Richard Hays

Medical Teacher, DOI: 10.1080/0142159X.2019.1652258

Published online: 22 Aug 2019

**Abstract:**

Introduction: Various pressures exist for curricular change, including economic forces, burgeoning knowledge, broadening learning outcomes, and improving quality and outcomes of learning experiences. In an Australian 5-year undergraduate medical course, staff were asked to reduce teaching hours by 20% to alleviate perceived overcrowded preclinical curriculum, achieve operating efficiencies and liberate time for students’ self-directed learning.

Methods: A case study design with mixed methods was used to evaluate outcomes.

Results: Teaching hours were reduced by 198 hours (14%) overall, lectures by 153 hours (19%) and other learning activities by 45 hours (7%). Summative assessment scores did not change significantly after the reductions: 0.4% increase, 1.5% decrease and 1.7% increase in Years 1, 2 and 3, respectively. The percentage of students successfully completing their academic year did not change significantly: 94.4% before and 93.3% after the reductions. Student evaluations from eVALUate surveys changed little, except workload was perceived to be more reasonable.

Conclusions: Teaching hours, particularly lectures, can be moderately reduced with little impact on student learning outcomes or satisfaction with an undergraduate medical course.

1. **Are rating scales really better than checklists for measuring increasing levels of expertise?**

Authors: Timothy J. WoodORCID Icon & Debra PughORCID Icon

Medical Teacher. Published online: 20 Aug 2019 DOI: 10.1080/0142159X.2019.1652260

**Abstract:**

**Background:** It is a doctrine that OSCE checklists are not sensitive to increasing levels of expertise whereas rating scales are. This claim is based primarily on a study that used two psychiatry stations and it is not clear to what degree the finding generalizes to other clinical contexts. The purpose of our study was to reexamine the relationship between increasing training and scoring instruments within an OSCE.

**Approach:** A 9-station OSCE progress test was administered to Internal Medicine residents in post-graduate years (PGY) 1–4. Residents were scored using checklists and rating scales. Standard scores from three administrations (27 stations) were analyzed.

**Findings:** Only one station produced a result in which checklist scores did not increase as a function of training level, but the rating scales did. For 13 stations, scores increased as a function of PGY equally for both checklists and rating scales.

**Conclusion:** Checklist scores were as sensitive to the level of training as rating scales for most stations, suggesting that checklists can capture increasing levels of expertise. The choice of which measure is used should be based on the purpose of the examination and not on a belief that one measure can better capture increases in expertise.

## Critical thinking, biases and dual processing: The enduring myth of generalisable skills

## Authors: Sandra Monteiro Jonathan Sherbino Matthew Sibbald Geoff Norman

## Medical Education 29 August 2019

**Abstract:**

Context

The myth of generalisable thinking skills in medical education is gaining popularity once again. The implications are significant as medical educators decide on how best to use limited resources to prepare trainees for safe medical practice. This myth‐busting critical review cautions against the proliferation of curricular interventions based on the acquisition of generalisable skills.

Structure

This paper begins by examining the recent history of general thinking skills, as defined by research in cognitive psychology and medical education. We describe three distinct epochs: (a) the Renaissance, which marked the beginning of cognitive psychology as a discipline in the 1960s and 1970s and was paralleled by educational reforms in medical education focused on problem solving and problem‐based learning; (b) the Enlightenment, when an accumulation of evidence in psychology and in medical education cast doubt on the assumption of general reasoning or problem‐solving skill and shifted the focus to consideration of the role of knowledge in expert clinical performance; and (c) the Counter‐Enlightenment, in the current time, when the notion of general thinking skills has reappeared under different guises, but the fundamental problems related to lack of generality of skills and centrality of knowledge remain.

Conclusions

The myth of general thinking skills persists, despite the lack of evidence. Progress in medical education is more likely to arise from devising strategies to improve the breadth and depth of experiential knowledge.

1. **Medical student strategies for actively negotiating hierarchy in the clinical environment**

Authors: Meredith Vanstone Lawrence Grierson

Medical Education. 28 August 2019

Context

Medical learning takes place in an extremely hierarchical environment. Medical students may struggle to understand how to succeed in such a rule‐bound environment that leaves them vulnerable to the influences of social power. This study explores how medical students experience the clinical learning environment from their low‐status positions in the social hierarchy.

Methods

Using constructivist grounded theory, we collected 88 hours of observation and 13 interviews with medical students completing clinical clerkships. Data collection focused on students’ interactions with their supervisors, colleagues and other staff members as they completed the core rotations of their clinical clerkships. Data analysts used a constant comparative approach to remain alert to the different ways in which medical students experienced and responded to social power used by their supervisors and colleagues.

Results

We describe a cyclical theory of how medical students appraised the environment, the needs and preferences of their supervisors and their personal resources in order to select and enact a strategy for interacting. They used these strategies when in the presence of supervisors, but also when supervisors were absent in preparation for the next interaction. The ways in which medical students chose and employed these strategies reflect a significant use of social and cognitive resources.

Conclusions

Power is an important component of the social culture of clinical learning environments; understanding the ways in which medical students experience and react to power can help educators, learners and administrators optimise learning opportunities. Medical education increasingly encourages students to exercise agency in seeking feedback and directing their own learning; this may be particularly challenging for students who cannot interpret social cues well, and those who lack social capital.