



## Twelve tips for embedding assessment *for* and *as* learning practices in a programmatic assessment system

Aubrie Swan Sein , Hanin Rashid , Jennifer Meka , Jonathan Amiel & William Pluta

To cite this article: Aubrie Swan Sein , Hanin Rashid , Jennifer Meka , Jonathan Amiel & William Pluta (2020): Twelve tips for embedding assessment *for* and *as* learning practices in a programmatic assessment system, Medical Teacher, DOI: [10.1080/0142159X.2020.1789081](https://doi.org/10.1080/0142159X.2020.1789081)

To link to this article: <https://doi.org/10.1080/0142159X.2020.1789081>



Published online: 13 Jul 2020.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

## Twelve tips for embedding assessment *for* and *as* learning practices in a programmatic assessment system

Aubrie Swan Sein<sup>a</sup>, Hanin Rashid<sup>b</sup>, Jennifer Meka<sup>c</sup>, Jonathan Amiel<sup>a</sup> and William Pluta<sup>d</sup>

<sup>a</sup>Columbia Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, USA; <sup>b</sup>Rutgers Robert Wood Johnson Medical School, Piscataway, NJ, USA; <sup>c</sup>Jacobs School of Medicine and Biomedical Sciences, State University of New York (SUNY) at Buffalo, Buffalo, NY, USA; <sup>d</sup>Pereiman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

### ABSTRACT

Programmatic assessment supports the evolution from assessment *of* learning to fostering assessment *for* learning and *as* learning practices. A well-designed programmatic assessment system aligns educational objectives, learning opportunities, and assessments with the goals of supporting student learning, making decisions about student competence and promotion decisions, and supporting curriculum evaluation. We present evidence-based guidance for implementing assessment *for* and *as* learning practices in the pre-clinical knowledge assessment system to help students learn, synthesize, master and retain content for the long-term so that they can apply knowledge to patient care. Practical tips are in the domains of culture and motivation of assessment, including how an honour code and competency-based grading system can support an assessment system to develop student self-regulated learning and professional identity, curricular assessment structure, such as how and when to utilize low-stakes and cumulative assessment to drive learning, exam and question structure, including what authentic question and exam types can best facilitate learning, and assessment follow-up and review considerations, such exam retake processes to support learning, and academic success structures. A culture change is likely necessary for administrators, faculty members, and students to embrace assessment as most importantly a learning tool for students and programs.

### KEYWORDS

Feedback; progress testing; self-assessment; written; planning

### Introduction

Assessments can measure student performance, competence, learning deficits, readiness to progress, and entrustment of professional activities. Assessments can also drive or shape learning behaviours to encourage students to direct their own learning (Epstein 2007; Schuwirth and van der Vleuten 2011; van der Vleuten 2016). Assessment systems can also help prepare students for the practice of medicine and to develop important professional values such as trustworthiness. As pre-clerkship students take on responsibility for patient care, pre-clerkship assessment systems that motivate students to manage their long-term learning and to apply knowledge to practice can enhance learning and professional identity formation. Medical schools can shift from emphasizing the importance of assessment as a measure *of* learning to using assessment intentionally *for* facilitating learning and *as* learning to foster student self-regulation (Heeneman et al. 2015; Eva et al. 2016).

A programmatic assessment system is one in which each individual assessment should provide learners with feedback, and each assessment should be viewed as part of a larger system in which the assessments are viewed holistically, to make high-stakes competence/promotion decisions (Schuwirth and van der Vleuten 2011; Heeneman et al. 2015). According to the 2018 Consensus Framework

for Good Assessment, the purposes of Programmatic Assessment are to:

1. Optimize the impact of assessments on learning, decisions regarding individual students, and curriculum quality.
2. Identify and provide feedback on individual student's areas of strength and weakness.
3. Provide students with a good basis for making self-assessments and judging learning needs.
4. Motivate students to remediate areas of weakness.
5. Provide information on instructional effectiveness to guide improvement (p. 1107). (Norcini et al. 2018)

Developing a learning-focused programmatic assessment system that optimizes a school's culture, student population, resources, and mission is a substantial project, and investment in faculty and staff development is needed (Schuwirth and Ash 2013; Schuwirth and van der Vleuten 2019).

We conducted a literature review of programmatic assessment and assessment *for* learning strategies related to cognitive knowledge-focused preclinical exams and reflected on our experiences working across eight medical schools to inform this work. In this article, we present evidence-based assessment tips that schools can use to evolve from an assessment *of* learning structure toward an assessment *for* and *as* learning system and to develop a programmatic assessment structure. While programmatic assessment structures also focus on non-medical knowledge competencies, such as communication and system-based practice, we focus on presenting preclinical medical

knowledge assessment strategies in this paper. These tips can also be used to enhance traditional systems to promote learning through assessment.

### Tip 1

#### ***Use assessment for learning principles to build a curricular and assessment structure to enhance learning***

Curriculum design processes begin with planning the curricular structure that aligns with the intended educational program learning outcomes. Designing an assessment system to enhance learning comes next. Assessment blueprinting can be utilized to develop and communicate how each assessment fits into a broader program of assessment and how progress decisions will be made based on performance across aggregate assessments (Wilkinson and Tweed 2018).

In making an institutional commitment to using assessment for learning and improvement, there is need for clear communication of the purposes of assessment and the use of assessment outcomes. In one study of programmatic assessment implementation, students' perceptions of this structure depended on whether they truly saw exams as learning opportunities, but the programmatic assessment structure did help to spur and direct learning (Heeneman et al. 2015). Intermediate-stakes assessments should provide learners with diagnostic, therapeutic, and prognostic information (van der Vleuten et al. 2015). Ideally designed assessments for learning are those that

- 1) help the learners define where they are in meeting the objectives of a course; 2) identify what they need to do further, 3) prepare them to transfer their knowledge and skills to novel situations; 4) enable them to gain a deeper understanding of the material, and 5) provide them an opportunity to personalize their learning. (p. 9) (Kulasegaram and Rangachari 2018).

Once the learning-focused curricular and assessment systems are designed, appropriate teaching and learning opportunities can be developed.

### Tip 2

#### ***Promote learning by adopting a competency-based education and evaluation structure and a culture of valuing learning over performance***

An ideal learning-focused curriculum and assessment system would promote a growth mindset orientation (Dweck 2006) among learners and teachers. A competency-based education and evaluation structure defines criteria for progression by a learners' attainment of pre-defined thresholds rather than their performance relative to peers. Within the medical education literature, competency-based grading has been often framed in terms of moving to a pass/fail, or pass/not yet, grading system. The adoption of pass/fail grading can reduce competition and foster collaboration among members of the class – core professional values (Reed et al. 2011; Spring et al. 2011), and can promote intrinsic motivation (White and Fantone 2010). It can also help mitigate bias conferred by rewarding learners whose enriched academic backgrounds have taught them to be

quick studiers and punishing those learners who need more time to reach the same degree of competence, potentially improving the cohesion of diverse classes.

Creating an assessment system that promotes self-regulation should intentionally avoid rewarding short-term learning or rote factual memorization using inefficient learning strategies (e.g. cramming). Cramming, as opposed to studying consistently across time and focusing on understanding, leads to poorer academic performance (Bickerdike et al. 2016). Reducing incentives to memorize details to answer every test question correctly in order to earn an 'honours' grade can give students more time to prioritize learning the most important concepts and to recognize errors made. Preparing for long-term board exams and high-stakes assessments can also be framed as a long-term and clinical learning opportunity (Swan Sein et al.).

### Tip 3

#### ***Create a culture of assessment that develops student professional identity as physicians***

We believe the pre-clerkship assessment system and overall culture of assessment should be congruent with the professional norms of medicine by supporting learners' professional identity formation - their core values, moral principles, and self-awareness. However, conducting low or intermediate-stakes examinations under high-security settings can convey a sense of distrust of learners. This is incongruent with the fact that students will soon (or already have) access to sensitive patient information and have responsibility for patient care. An honour code can establish expectations and entrust medical students with core professional values related to honesty and integrity (Irby and Hamstra 2016). By utilizing an honour code to set clear and specific standards for students, and expectation-setting discussions with students, students learn to behave professionally, and to hold peers accountable to these standards (Ross et al. 2019).

New exam administration technologies allow for a shift from high-security, proctored examinations to un-proctored, off-site, and time-flexible examinations. The application of professional standards to assessment also opens up the possibility of adopting these exam administration strategies. Un-proctored and time-flexible exams can help to foster professional identity formation and professionalism (Ross et al. 2019). Time-flexible exams can induce students to make decisions about when they are prepared to demonstrate competence within a knowledge domain.

### Tip 4

#### ***Utilise coaches to help students progress and develop academic success systems to promote learning and growth***

Providing coaches or mentors for students to review performance, reflections, learning opportunities, and next steps is also important to facilitating student learning (van der Vleuten et al. 2015). Coaches can monitor student assessment performance over time to identify students who may benefit from academic support or enrichment. Viewing remediation as an academic support opportunity

and establishing *proactive and preventative* support systems for students can guide learning and improvement, prevent future exam failures, and can orient a school's culture toward learning (Guerrasio 2013; Chou et al. 2014; Kalet and Chou 2014).

Certain learning strategies can improve student learning, cognition, and affect, as well as performance on assessments, however many students do not engage in these strategies instinctually (Hattie et al. 1996). Focusing on errors can be a powerful driver of learning (Metcalf 2017). Learning specialists can help all students to use long-term learning strategies, including how to learn from assessments and practice questions (Swan Sein et al. 2020a), process feedback and improve on weaknesses, and model study strategies that utilize self-assessment strategies to foster deeper learning, teach students time management skills, and promote self-testing strategies (Swan Sein et al. 2020b).

### Tip 5

#### **Promote assessment as learning to drive student self-regulated learning**

Assessment *as learning* intertwines learning and assessment and places ownership over learning in the role of the student (Dann 2014). The intention is for students to use metacognitive strategies to plan, monitor, and self-regulate learning. Successful medical students are self-regulated learners who plan, evaluate, and monitor their learning strategies (Artino et al. 2011). Assessment *as learning* tools, such as portfolios or reflection exercises, can help students to become metacognitive and self-regulate their learning. Using student portfolios for programmatic assessment can help compile multiple assessments, collect assessment evidence over time, and amass feedback from multiple evaluators. In combination with faculty coaching and reflective writing, portfolios can facilitate assessment *as learning* and can promote critical thinking, problem solving, and self-assessment to drive life-long learning behaviours (Gadbury-Amyot and Overman 2018). Students can then engage in deliberate practice to improve performance with immediate and useful feedback (Ericsson 2004).

### Tip 6

#### **Promote assessment for learning via frequent low-stakes assessment and feedback**

Ensuring that an assessment structure provides students with frequent low-stakes assessment *for learning* and feedback is important. Learning occurs throughout all phases of the assessment cycle, which includes a pre-assessment, true assessment, and post-assessment phase. In the pre-assessment phase, learning about tests via exam blueprints or practice tests directs students about what to study and learn. In the true assessment phase, engaging in retrieval practice enhances the storage and ability to retrieve information at a future time. In the post-assessment phase, students can learn from their mistakes when they receive feedback about their performance and review incorrect answers (Gielen et al. 2003; Schuwirth and Ash 2013).

Research on testing and assessment *for learning* has shown that repeated testing slows down the forgetting process (Larsen et al. 2008; Rowland 2014). Frequent low-stakes assessment is more likely than infrequent high-stakes assessment to promote learning (Vansteenkiste et al. 2004), reduce the achievement gaps among students from different socioeconomic classes (Pennebaker et al. 2013), promote long-term retention by motivating students to engage frequently with content, and discourage the practice of 'gorge and purge' (Kulasegaram and Rangachari 2018).

Weekly problem sets can be designed to promote learning and provide feedback to students. Low-stakes implicit-confidence testing can be used with multiple choice questions to help students to recognize misconceptions (Klymkowsky et al. 2006) and to learn from errors made with high confidence (Metcalf 2017). Problem sets can also interleave content from across topics, to help students to study and test knowledge on a mixture of topics they have learned, instead of focusing on just one topic for a long period, promoting the ability to apply knowledge in novel settings (Dunlosky et al. 2013). Problem sets of multiple choice and open-ended questions can be assigned and worked on early in a week, and students could then complete the problem set from memory at the end of the week to benefit from the testing effect. Providing students with high quality feedback on their performance, even if not frequently, is an important component of a programmatic assessment structure; narrative feedback can be provided on answers to open-ended questions, for example.

### Tip 7

#### **Use cumulative assessment to encourage student retrieval and application of information over time**

On the basis of the principle of spaced practice, an assessment system should directly or indirectly expose students to content cumulatively and repeatedly. Cumulative assessments, which include recent content and content learned throughout a program, promote student learning and long-term retention, in part because students work to keep content available over time if they expect a cumulative exam (Szpunar et al. 2007; Larsen et al. 2008; Wrigley et al. 2012). Notably, '*it is often only when durability of knowledge is tested in the longer run that the student may perceive a problem with their study performance*' (Kalet and Chou 2014, p. 42). Non-cumulative exam feedback can give students an 'illusion of competence' because it is not known how well this understanding will be retained for the long-term (Koriat and Bjork 2005). Therefore, it is important for content to appear on multiple exams over time in a variety of forms.

Progress testing, which provides students with longitudinal assessment opportunities with repeated comprehensive exams over time, such as giving multiple practice board exams during the preclinical years, is associated with improvements in knowledge retention (Johnson et al. 2014). Students study more continuously for progress exams and build a better knowledge base leading to minimized use of test-driven learning strategies (Norman et al. 2010; Schuwirth and van der Vleuten 2012). Schools could consider implementing regular cumulative examinations, or

integrate questions from previous topics into exams throughout the curriculum, to enhance learning and promote self-regulation of learning.

### Tip 8

#### ***Use a mix of assessment question types to challenge students to employ a variety of studying strategies***

It is well-established that retrieval practice causes a testing effect and can lead to high quality learning (Larsen et al. 2008, 2013; Larsen and Butler 2013; Green et al. 2018). Engaging in answering questions can serve as a 'desirable difficulty' as questions challenge students to engage in retrieval practice and apply their learning (Marsh and Butler 2013). Multiple choice questions can also promote learning and conceptual understanding. Students who carefully evaluate and generate reasons for why each answer choice is correct or incorrect will engage in the type of deep processing that promotes retention of both tested and untested information (Little and Bjork 2015).

Students learn well when an exam helps them to engage in retrieval practice via answering deep questions and causes them not only to study and learn facts but also to engage in deeper reasoning and elaboration. Questions that promote elaboration, including generation and explanation questions, enhance understanding and long-term retention (Larsen and Butler 2013; van der Vleuten and Driessen 2014). Questions that challenge students to answer *why*, *how*, and *what-if* promote deeper comprehension and levels of learning (Craig et al. 2006). Non-MCQ questions are not trivial to grade, but can be developed and incorporated over time into exams (Bierer et al. 2018). New exam delivery software can facilitate rapid grading of open-ended questions on graded assessments. Students can be provided with worked examples of question answers for learning purposes.

### Tip 9

#### ***Use complex questions to integrate across domains of knowledge***

Students who view basic science as important for their current and future goals will demonstrate greater persistence, employ strategies for deeper learning, and may show increased achievement (Artino et al. 2010). Assessment questions can aid in developing students' perceived value of different domains, such as basic science, social and behavioural sciences, health systems sciences, and humanities, and should be relevant and motivate students to thoughtfully answer them. Growing education evidence highlights the utility of helping learners to integrate basic and clinical science domains of knowledge in their minds (Baghdady et al. 2013; Bandiera et al. 2018), by helping them make conceptual connections across domains. Providing students with clinically oriented practice questions in question banks can help anchor learning and to motivate students to learn and apply their knowledge through clinical reasoning.

A number of assessment question types can be used to assess integrated learning. These include reflection questions that elicit connections between related information.

Multiple choice self-assessment questions and concept appraisal questions can ask for the mechanisms behind a clinical scenarios' findings. Clinical reasoning exercises where students write a paragraph describing a patient's problem and associated mechanisms can also help integrate learning. A diagnostic justification exercise where students suggest a differential diagnosis and the rationale for a patient's problems can also be used (Brauer and Ferguson 2015). Educators can begin by creating low-stakes assessments with these types of questions to guide student learning and comfort with this type of question/assessment activity.

### Tip 10

#### ***Leverage authentic learning practices within the cognitive knowledge assessment system***

Physicians need to be able and willing to work and learn as a team, develop communication skills and peer-assessment skills, and reference and apply new knowledge (Hodges 2013; ten Cate and Chen 2016). Incorporating authentic social and cognitive practices in the assessment system promotes these skills, such as via the social practice of collaborative examinations. Collaborative learning can lead to improved academic performance, interpersonal interactions, perceptions of social support, self-esteem, and retention in academic programs, compared to individual work, as well as compared to competitive learning paradigms (Hmelo-Silver et al. 2013). Low-stakes group assessments, such as Team-Based Learning 'Group Readiness Assurance Tests' or anatomy group-based lab exams can introduce collaborative practice and stimulate an active learning process (Pluta et al. 2013).

An authentic cognitive practice in assessment can be taking open-book examinations. Research shows benefits from both open-book and closed-book exams (Durning et al. 2016). Completing closed-book quizzes on conceptual questions can lead to less forgetting over time than open-book quizzes because students practice recalling information and use the quiz as a retrieval practice exercise (Agarwal et al. 2008). Preparing for an open-book test may encourage students to focus on understanding the important concepts and content, rather than memorizing details. Creating a 'crib-sheet' to prepare for an exam may help students organize knowledge, identify which knowledge is most important, identify which knowledge should be memorized, and which knowledge can/should be looked-up. In the future, students may be provided with additional resources during board exams, such as biological or metabolic pathways or immunization schedules, to assess their ability to apply or synthesize learning, and not just memorize details.

### Tip 11

#### ***Create non-punitive exam retake processes that promote learning***

In some programmatic assessment structures, exam retakes are not utilised because multiple assessments over time are used to make high stakes decisions about student progression, making retakes unnecessary from a student evaluation



standpoint. Schools may choose to use exam retakes so that students identify and learn from errors. By having retakes be a normal and non-punitive part of the assessment system, institutions can begin to change the learning culture so students are accountable for mastering the knowledge, skills, and attitudes necessary to achieve competence. A fair re-assessment should provide students with appropriate remediation and the opportunity to demonstrate successful learning in previously identified areas of deficiency.

Re-assessment can be accomplished via a number of different formats, including an assessment customized to a student's particular areas of difficulty (Hauer et al. 2009; Hawthorne et al. 2014). One practical retake approach is to have students retake the same exam that they have not mastered, but to also require them to provide explanations of the answers to each question, to demonstrate that they have moved beyond memorizing the correct answer, and to help them to learn important concepts. During re-grading, students can be given credit only if the question explanation is also correct.

The best retake timing likely depends on the quantity and types of errors made. Retakes can happen close to a failure, after engagement in a subsequent learning activity, and/or when a student feels prepared. Advantages of quick retakes include learning material needed for the next block of the curriculum and not falling behind. Remediation during another course might be distracting to learning new content. Students with major conceptual or fund of knowledge deficits may benefit from dedicated study time and additional support for learning.

## Tip 12

### *Utilise exam feedback data for student learning and ultimately for high stakes decisions*

Giving students exam feedback can have more learning impact than many instructional strategies (Hattie and Timperley 2007). Learning and long-term retention are also enhanced by appropriately timed feedback (Black and Wiliam 2006; Harks et al. 2014; van de Ridder et al. 2015). Rapid feedback on high-stakes assessments may have a greater impact on learning since students are often not able to take the test again (Hattie and Timperley 2007); wrong answers may be learned and reinforced if immediate feedback is not given and utilized by students (Butler and Roediger 2008). With the advent of electronic testing systems, exams can easily be reviewed by students, with a variety of security options available, such as during post-exam review sessions. Feedback can be facilitated by scheduling post-exam reviews sessions either immediately or shortly following exams.

Delayed feedback in low-stakes assessments can help students can spend time working through the problem until they find the correct answer (Butler et al. 2007; Mullet et al. 2014). There is also evidence that narrative and explanation feedback is superior to simply correct answer feedback (Black and Wiliam 2006). Students can also be provided with electronic question explanations, or with worked exemplar responses once they work to answer questions independently. Feedback that reveals patterns in students' weaknesses, that focuses on specific areas of

improvement and that details how good performance is achieved will promote self-regulated learning (Nicol and Macfarlane-Dick 2006).

Leveraging emerging assessment technologies such as tagging questions from intermediate-stakes assessments with topic keywords can allow schools to develop dashboards that track student progress towards specific benchmarks, and students can identify potential content gaps and use these gaps to inform their self-directed learning. Student coaches or learning specialists can meet with students to review and make use of these data.

As programmatic assessment systems rely on using multiple data points from assessments collected over time for high stakes decisions, using assessment technologies to compile data for student progression committee reviews is also important. Assessment data can and should also be used for curricular and program evaluation and improvement purposes (van der Vleuten et al. 2015).

## Conclusion

A well-designed assessment program can have many benefits for learning (Roediger et al. 2011), but developing and implementing evidence-based assessment systems is an area in need of further consideration and study (Norcini et al. 2018; Schuwirth and van der Vleuten 2019). These tips demonstrate that assessment systems can optimize learning when: (1) learners enter into and promote an environment of trust, (2) schools set thresholds for competency and allow learners the opportunity to retake exams until they have reached competency, (3) schools promote self-regulated learning with ample feedback and resources, (4) learners and schools collaboratively focus on learning and reducing the stigma of 'failure' – some content requires more time or effort, depending on background and learning strategies, (5) assessments focus on clinical reasoning, not only recall of facts, and (6) schools offer cumulative assessments so students do not forget what they learned before.

We have provided tips in order to 'disrupt thinking' and help schools to develop a program of assessment to generate positive learning behaviours so students are better equipped to transfer and apply knowledge during patient care activities. A culture change is likely necessary for all administrators, faculty members, and students to embrace assessment as most importantly a learning tool for students and programs. Educational settings are complex; ultimately, schools need to balance what is best for long-term student learning, what is best for student well-being, what is feasible, and what works with a school's learning culture.

## Acknowledgements

The authors would like to thank the members of the Society of the Directors of Research in Medical Education, the International Association of Medical Science Educators, the Northeast Group on Education Affairs, the Medical Education Learning Specialists, the American Dental Education Association groups, and the article reviewers for guidance in the development of these tips and advice.

## Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

## Notes on contributors

**Aubrie Swan Sein**, PhD, EdM, is Director of the Center for Education Research and Evaluation and Assistant Professor of Education Assessment in Pediatrics and Dental Medicine at Columbia Vagelos College of Physicians and Surgeons, New York, NY.

**Hanin Rashid**, PhD, EdM, is Associate Director for the Office for Advancing Learning, Teaching, and Assessment, a Learning Specialist in the Cognitive Skills Program, and Assistant Professor of Psychiatry at Rutgers Robert Wood Johnson Medical School, Piscataway, NJ.

**Jennifer Meka**, PhD, EdM, is Director of the Medical Education and Educational Research Institute, Assistant Dean for Medical Education, and Assistant Professor of Medicine at The Jacobs School of Medicine and Biomedical Sciences at The State University of New York (SUNY) at Buffalo, Buffalo, NY.

**Jonathan Amiel**, MD, is co-Interim Vice Dean for Education and Senior Associate Dean for Curricular Affairs, and Associate Professor of Psychiatry, Columbia Vagelos College of Physicians and Surgeons, New York, NY.

**William Pluta**, PhD, EdM, is Director of UME Evaluation at the Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA.

## References

- Agarwal PK, Karpicke JD, Kang SHK, Roediger HL, McDermott KB. 2008. Examining the testing effect with open- and closed-book tests. *Appl Cognit Psychol*. 22(7):861–876.
- Artino A, Hemmer P, Durning S. 2011. Using self-regulated learning theory to understand the beliefs, emotions, and behaviors of struggling medical students. *Acad Med*. 86(10 Suppl):S35–S38.
- Artino A, La Rochelle J, Durning S. 2010. Second-year medical students' motivational beliefs, emotions, and achievement. *Med Educ*. 44(12):1203–1212.
- Baghdady MT, Carnahan H, Lam EW, Woods NN. 2013. Integration of basic sciences and clinical sciences in oral radiology education for dental students. *J Dent Educ*. 77(6):757–763.
- Bandiera G, Kuper A, Mylopoulos M, Whitehead C, Ruetalo M, Kulasegaram K, Woods NN. 2018. Back from basics: integration of science and practice in medical education. *Med Educ*. 52(1):78–85.
- Bickerdike A, O'Deasmhunaigh C, O'Flynn S, O'Tuathigh C. 2016. Learning strategies, study habits and social networking activity of undergraduate medical students. *Int J Med Educ*. 7:230–236.
- Bierer SB, Colbert CY, Foshee CM, French JC, Pien LC. 2018. Tool for diagnosing gaps within a competency-based assessment system. *Acad Med*. 93(3):512.
- Black P, Wiliam D. 2006. Assessment and classroom learning. *assessment in education: principles. Policy & Practice*. 5(1):7–74.
- Brauer DG, Ferguson KJ. 2015. The integrated curriculum in medical education: AMEE Guide No. 96. *Med Teach*. 37(4):312–322.
- Butler A, Karpicke J, Roediger H. 2007. The effect of type and timing of feedback on learning from multiple-choice tests. *J Exp Psychol Appl*. 13(4):273–281.
- Butler A, Roediger H. 2008. Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Mem Cognit*. 36(3):604–616.
- Chou CL, Kalet A, Hauer KE. 2014. A research agenda for remediation in medical education. In: *Remediation in medical education*. New York (NY): Springer Books; p. 339–348.
- Craig SD, Sullins J, Witherspoon A, Gholson B. 2006. The deep-level-reasoning-question effect: the role of dialogue and deep-level-reasoning questions during vicarious learning. *Cogn Instr*. 24(4):565–591.
- Dann R. 2014. Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice. *Assess Educ Princ Policy Pract*. 21(2):149–166.
- Dunlosky J, Rawson KA, Marsh EJ, Nathan MJ, Willingham DT. 2013. Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol Sci Public Interest*. 14(1):4–58.
- Durning SJ, Dong T, Ratcliffe T, Schuwirth L, Artino AR Jr., Boulet JR, Eva K. 2016. Comparing open-book and closed-book examinations: a systematic review. *Acad Med*. 91(4):583–599.
- Dweck C. 2006. *Mindset: the new psychology of success*. New York (NY): Random House.
- Epstein RM. 2007. Assessment in medical education. *N Engl J Med*. 356(4):387–396.
- Ericsson KA. 2004. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med*. 79(10 Suppl):S70–S81.
- Eva KW, Bordage G, Campbell C, Galbraith R, Ginsburg S, Holmboe E, Regehr G. 2016. Towards a program of assessment for health professionals: from training into practice. *Adv Health Sci Educ Theory Pract*. 21(4):897–913.
- Gadbury-Amyot CC, Overman PR. 2018. Implementation of portfolios as a programmatic global assessment measure in dental education. *J Dent Educ*. 82(6):557–564.
- Gielen S, Dochy F, Dierick S. 2003. Evaluating the consequential validity of new modes of assessment: The influence of assessment on learning, including pre-, post-, and true assessment effects. In: Segers M, Dochy F, Cascallar E, editors. *Optimising new modes of assessment: in search of qualities and standards*. New York (NY): Kluwer Academic Publishers.
- Green ML, Moeller JJ, Spak JM. 2018. Test-enhanced learning in health professions education: a systematic review: BEME Guide No. 48. *Med Teach*. 40(4):337–350.
- Guerrassio J. 2013. *Remediation of the struggling medical learner*. Irwin (PA): Association for Hospital Medical Education.
- Harks B, Rakoczy K, Hattie J, Besser M, Klieme E. 2014. The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness. *Educ Psychol*. 34(3):269–290.
- Hattie J, Biggs J, Purdie N. 1996. Effects of learning skills interventions on student learning: a meta-analysis. *Rev Educ Res*. 66(2):99–136.
- Hattie J, Timperley H. 2007. The power of feedback. *Rev Educ Res*. 77(1):81–112.
- Hauer KE, Ciccone A, Henzel TR, Katsufraakis P, Miller SH, Norcross WA, Papadakis MA, Irby DM. 2009. Remediation of the deficiencies of physicians across the continuum from medical school to practice: a thematic review of the literature. *Acad Med*. 84(12):1822–1832.
- Hawthorne MR, Chretien KC, Torre D, Chheda SG. 2014. Re-demonstration without remediation—a missed opportunity? A national survey of internal medicine clerkship directors. *Med Educ Online*. 19(1):25991.
- Heeneman S, Oudkerk Pool A, Schuwirth LW, van der Vleuten CP, Driessen EW. 2015. The impact of programmatic assessment on student learning: theory versus practice. *Med Educ*. 49(5):487–498.
- Hmelo-Silver CE, Chinn CA, O'Donnell AM, Chan C. 2013. *The international handbook of collaborative learning*. New York (NY): Routledge.
- Hodges B. 2013. Assessment in the post-psychometric era: learning to love the subjective and collective. *Med Teach*. 35(7):564–568.
- Irby DM, Hamstra SJ. 2016. Parting the clouds: three professionalism frameworks in medical education. *Acad Med*. 91(12):1606–1611.
- Johnson TR, Khalil MK, Pepler RD, Davey DD, Kibble JD. 2014. Use of the NBME Comprehensive Basic Science Examination as a progress test in the preclerkship curriculum of a new medical school. *Adv Physiol Educ*. 38(4):315–320.
- Kalet A, Chou C. 2014. *Remediation in medical education: a mid-course correction*. New York (NY): Springer Books.
- Koriat A, Bjork RA. 2005. Illusions of competence in monitoring one's knowledge during study. *J Exp Psychol Learn Mem Cogn*. 31(2):187–194.
- Kulasegaram K, Rangachari PK. 2018. Beyond “formative”: assessments to enrich student learning. *Adv Physiol Educ*. 42(1):5–14.
- Larsen D, Butler A. 2013. Chapter 38: test-enhanced learning. In: Walsh K, editor. *Oxford textbook of medical education*. Oxford (UK): Oxford University Press; p. 443–452.

- Larsen DP, Butler AC, Roediger HL. 2008. Test-enhanced learning in medical education. *Med Educ.* 42(10):959–966.
- Larsen DP, Butler AC, Roediger HL. 2013. Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Med Educ.* 47(7):674–682.
- Little J, Bjork EL. 2015. Optimizing multiple-choice tests as tools for learning. *Mem Cognit.* 43(1):14–26.
- Marsh EJ, Butler AC. 2013. Memory in educational settings. In: Reisberg D, editor. *The Oxford handbook of cognitive psychology*. New York: Oxford University Press; p. 299–317.
- Metcalfe J. 2017. Learning from errors. *Annu Rev Psychol.* 68:465–489.
- Klymkowsky MW, Taylor LB, Spindler SR, Garvin-Doxas RK. 2006. Two-dimensional, implicit confidence tests as a tool for recognizing student misconceptions. *J Coll Sci Teach.* 36(3):44–48.
- Mullet HG, Butler AC, Verdin B, von Borries R, Marsh EJ. 2014. Delaying feedback promotes transfer of knowledge despite student preferences to receive feedback immediately. *J Appl Res Mem Cogn.* 3(3): 222–229.
- Nicol DJ, Macfarlane-Dick D. 2006. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Stud Higher Educ.* 31(2):199–218.
- Norcini J, Anderson MB, Bollela V, Burch V, Costa MJ, Duvivier R, Hays R, Palacios Mackay MF, Roberts T, Swanson D. 2018. 2018 Consensus framework for good assessment. *Med Teach.* 40(11): 1102–1109.
- Norman G, Neville A, Blake JM, Mueller B. 2010. Assessment steers learning down the right road: impact of progress testing on licensing examination performance. *Med Teach.* 32(6):496–499.
- Pennebaker JW, Gosling SD, Ferrell JD. 2013. Daily online testing in large classes: boosting college performance while reducing achievement gaps. *PLoS One.* 8(11):e79774.
- Pluta WJ, Richards BF, Mutnick A. 2013. PBL and beyond: trends in collaborative learning. *Teach Learn Med.* 25 Suppl 1:S9–S16.
- Reed DA, Shanafelt TD, Satele DW, Power DV, Eacker A, Harper W, Moutier C, Durning S, Massie FS, Jr., Thomas MR. 2011. Relationship of pass/fail grading and curriculum structure with well-being among preclinical medical students: a multi-institutional study. *Acad Med.* 86(11):1367–1373.
- Roediger HL, Putnam AL, Smith MA. 2011. Ten benefits of testing and their applications to educational practice. *Psychol Learn Motiv Cogn Educ.* 55:1–36.
- Ross PT, Keeley MG, Mangrulkar RS, Karani R, Gliatto P, Santen SA. 2019. Developing professionalism and professional identity through unproctored, flexible testing. *Acad Med.* 94(4):490–495.
- Rowland CA. 2014. The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol Bull.* 140(6): 1432–1463.
- Schuwirth L, Ash J. 2013. Chapter 35, Principles of assessment. In: Walsh K, editor. *Oxford textbook of medical education*. Oxford, United Kingdom: Oxford University Press; p. 409–420.
- Schuwirth L, van der Vleuten C. 2011. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach.* 33(6):478–485.
- Schuwirth L, van der Vleuten C. 2012. The use of progress testing. *Perspect Med Educ.* 1(1):24–30.
- Schuwirth L, van der Vleuten C. 2019. How ‘testing’ has become ‘programmatic assessment for learning’. *Health Prof Educ.* 5(3): 177–184.
- Spring L, Robillard D, Gehlbach L, Simas TA. 2011. Impact of pass/fail grading on medical students’ well-being and academic outcomes. *Med Educ.* 45(9):867–877.
- Swan Sein A, Cuffney F, Clinchot D. 2020a. How to help students strategically prepare for the MCAT exam and learn foundational knowledge needed for medical school. *Acad Med.* 95(3):484.
- Swan Sein A, Daniel M, Fleming A, Morrison G, Christner J, Esposito K, Pock A, O’Connor Grochowski C, Dalrymple J, Santen S. 2020b. Identifying and supporting struggling students to prevent failures when the United States Medical Licensing Examination (USMLE) step 1 is moved to after clerkships. *Acad Med.* [accessed 2020 March 3]. <https://doi.org/10.1097/ACM.0000000000003272>
- Szpunar KK, McDermott KB, Roediger HL. 3rd. 2007. Expectation of a final cumulative test enhances long-term retention. *Mem Cognit.* 35(5):1007–1013.
- ten Cate O, Chen H. 2016. The parts, the sum and the whole—evaluating students in teams. *Med Teach.* 38(7):639–641.
- van de Ridder JM, McGaghie WC, Stokking KM, ten Cate OT. 2015. Variables that affect the process and outcome of feedback, relevant for medical training: a meta-review. *Med Educ.* 49(7):658–673.
- van der Vleuten C. 2016. Revisiting ‘Assessing professional competence: from methods to programmes’. *Med Educ.* 50(9):885–888.
- van der Vleuten CP, Driessen EW. 2014. What would happen to education if we take education evidence seriously? *Perspect Med Educ.* 3(3):222–232.
- van der Vleuten C, Schuwirth L, Driessen E, Govaerts M, Heeneman S. 2015. Twelve Tips for programmatic assessment. *Med Teach.* 37(7): 641–646.
- Vansteenkiste M, Simons J, Lens W, Sheldon KM, Deci EL. 2004. Motivating learning, performance, and persistence: the synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *J Pers Soc Psychol.* 87(2):246–260.
- White CB, Fantone JC. 2010. Pass-fail grading: laying the foundation for self-regulated learning. *Adv Health Sci Educ Theory Pract.* 15(4): 469–477.
- Wilkinson TJ, Tweed MJ. 2018. Deconstructing programmatic assessment. *Adv Med Educ Pract.* 9:191–197.
- Wrigley W, van der Vleuten CP, Freeman A, Muijtjens A. 2012. A systemic framework for the progress test: strengths, constraints and issues: AMEE Guide No. 71. *Med Teach.* 34(9):683–697.