PERSPECTIVE
"Understanding Checkpoints" Followed by Self-Evaluation to Assess Students' Scientific Literacy and Mastery of Content and Techniques

Jennifer E. Schaefer
Biology Department, College of St. Benedict/St. John's University, Collegeville, MN 56321.

An Understanding Checkpoint (UC) presents students with previously-unseen figures from published scientific studies accompanied by questions about the study methods, results, and implications. The UCs incorporate content, concepts, and techniques previously discussed in class, although the figures and study from which they are taken are new to students. They are in-class, open-note, time-limited assessments that simultaneously assess course learning goals related to: neurobiology principles and content; process of scientific investigation, including neurobiological research tools and data interpretation; and reading and analyzing primary research literature.

After students submit their work, they are provided the full publication and are asked to grade their own work, providing rationale for their evaluation. The self-evaluative portion of the assignment incentivizes students to identify and remediate ongoing weaknesses. It also provides spaced retrieval practice to enhance learning. The final grade for the UC incorporates the student's original answers and the accuracy of the self-assessment rationale.

Student and instructor feedback indicates that the self -evaluative requirement develops a deeper understanding of the course material and enhances metacognitive effort, in addition to providing an opportunity to improve the UC grade. This strategy was originally presented as a teaching demonstration at the 2017 Faculty for Undergraduate Neuroscience Workshop.

Key words: neuroscience education; STEM education; formative assessment; summative assessment; undergraduate

Best practices in STEM education indicate that we should incorporate meaningful and applied assessments into our courses and that tests are opportunities for learning (McKeachie and Svinicki, 2006). This may be particularly important in a field with rapidly advancing techniques and understanding, such as neuroscience. Neuroscience students should be equipped with the skills to incorporate new findings and technical advances into their knowledge base after they leave our courses. As such, neuroscience educators need strategies to assess students' scientific literacy, experimental design ability, and data interpretation, in addition to content knowledge. We also need strategies that increase retention and transfer of knowledge. Ideally, assessments can function as learning experiences that increase student understanding of course content and concepts and that help students to improve their own metacognitive processes (Ambrose et al., 2010).

Although instructors may desire to incorporate assessments that are applied and meaningful, such assessments can be difficult to design. Particular challenges include finding applied contexts, decreased objectivity when grading "open ended" assessment artifacts, and, sometimes, student resistance to a perceived higher level of challenge in applied assessments. Perhaps the most immediate impediment to developing and adopting meaningful assessments is the time required to prepare and grade these assignments.

The following describes an assessment tool called an "Understanding Checkpoint" (UC). UCs assess student mastery of course content, experimental design, data interpretation, and scientific literacy in a format that is reasonable to prepare and grade. The UCs incorporate current published research and are, therefore, meaningful and applied. Importantly, UCs provide an opportunity for students to evaluate their own work and revise their answers, which increases student buy-in, decreases perceived stakes on the assessment, provides spaced retrieval practice, and transforms a summative assessment into a learning opportunity.

COURSE CONTEXT

Neurobiology is a discussion- and writing-based, upper-division Biology course at a private undergraduate liberal arts institution. The course is enrolled primarily with junior and senior Biology and Psychology majors as well as pre-Physical Therapy students, but is open to any student who has completed two semesters of introductory Biology with grades of C or better. Course enrollment is approximately 20 students. For further course information see Schaefer (2015).

Course learning goals:

1. Students will demonstrate a working understanding of foundational neurobiology principles including neuron and circuit structure and function.
2. Students will become fluent in the process of scientific investigation, the application of common neurobiological research tools, and the interpretation of data produced by these tools.
3. Students will practice communicating science through the process of critically reading,

interpreting, and presenting results from primary research literature in oral and written form.

ASSESSMENT TOOL DESIGN

Students are presented with figures from a previously-unseen, published scientific study alongside modified figure legends and questions about study methods, results, and implications; see example UC in Supplementary Material. The questions in the UCs are designed to assess previously-discussed concepts (learning goal 1) and techniques (learning goal 2) but the figures are new to students and therefore require transfer of knowledge to new data in order to successfully answer the questions. UCs are in-class, open-note, time-limited assessments. Students must provide thorough, clear, and concise written responses (learning goal 3).

After students submit their work at the end of the class period, the instructor provides the full publication. Students grade their own work, providing rationale for their evaluation, and submit the evaluation to the instructor. Students are instructed to evaluate their answers' correctness, clarity, and completeness in the rationales. The self-evaluative portion of the assignment incentivizes students to identify and remediate ongoing weaknesses, addresses learning goals 1-3, and provides an opportunity to improve the UC grade. It also builds in another contact with the course material and figures in the paper. thereby forcing spaced repetition with the material.

The instructor grades the original submission as well as the self-evaluation and rationale. The final grade for the UC incorporates both components (the student's original answers to the UC and the accuracy of the self-assessment rationale). I grade each answer on a "plus" (100%), "check-plus" (85%), "check" (60%), "zero" (0%) scale. The final assignment score uses the same scale, and is the most common/average of the individual answer scores. This grading scale is efficient and, because students have already assessed their own work using the same scale, does not require extensive commentary from the instructor.

RESULTS

Students view UCs as authentic, motivating, and meaningful--largely because the material is drawn from current scientific studies (anecdotal and course-reviews). For example, one student comment on an open ended course survey question about beneficial aspects of the course stated:

> "I also liked the understanding checkpoints. They were framed in a way that I found myself learning as I moved through the questions on them. It was a way of kind of guiding us to apply principles we had learned in a real research example. I also liked that we were able to correct our mistakes afterwards and earn points that way as well, because that helped the learning process. Traditional tests would not have been as effective for this class, so the checkpoints were a nice spin."

Students who prefer standardized assessment of content seem largely satisfied with the UCs, presumably because of the opportunity for self-assessment and, thereby, grade improvement.

As the instructor, I find that the self-evaluation component helps student to develop a deeper understanding of the course material and enhances student metacognitive effort. Many students submit very detailed rationales in the self-evaluations, which helps me, as the instructor, to be confident in estimating their mastery level. I have observed a decrease in student questions about grading because students actively evaluate their responses before they see my feedback. My grading of the original submissions more often than not agrees with the students' self-assessments (anecdotal). I receive almost no complaints or arguments from students about grading and my impression is that this is because of the prior self-evaluation.

Finally, I find that the UCs are enjoyable to develop and grade. I apply the template in the Supplementary Material to a new, current paper for each UC. Often, questions can be repeated between UCs because many are generalizable about experimental design, controls, interpretation, etc. I enjoy reading the students' self-evaluations because they are, by and large, insightful and honest, and often delightful and funny.

An instructor at a different institution (C. Favero, Ursinus College) piloted UCs in her Developmental Neurobiology course using the Supplementary Material and the grading scheme described in "Assessment Tool Design" as a template. The course is upper-division, writing-intensive, and typically enrolls 9-12 students. Feedback from Favero indicated that the UCs were reasonable in time and energy required to prepare and grade (personal communication). She typically included two figures in each UC and sometimes a third, more challenging figure, as a bonus question. Her students were receptive to the UCs as an assessment and were often very detailed in their self-evaluation rationales, similar to my observations above. Favero noted as additional benefits that questions could be reused on subsequent UCs (for example: identifying treatments as gain or loss of function), and that the UCs brought rigor to the course. A drawback noted by Favero was that UCs should be reused with caution if they include the self-evaluation component because students in subsequent semesters could have access to the source paper and other materials from former students.

DISCUSSION

UCs incorporate multiple best practices for assessment. This course does not have traditional exams and these are the closest equivalent assignment to an exam in the course. This assessment shifts from a high stakes exam to a more low stakes endeavor because the in-class first attempt is open-note and because the subsequent self-assessment and rationale are weighted equally in the assessment grade with the first attempt. Low stakes assessments have been shown to increase student learning and decrease student anxiety, particularly for underrepresented student populations, including women (Cotner and Ballen, 2017). Further, scientists and professionals rarely operate in a truly "closed book" environment, so the open-note aspect of the assessment is more realistic for future professionals. The UCs could be administered as closed-book, closed-note assessments but are challenging even as open-book, open-note assignments because they present open-ended questions that require higher order thinking. The subsequent self-evaluation requires entirely higher order thinking in order to provide rationale behind students' assessment of their work.

Adding self-assessment to a UC also ensures that students will interact with the course material at least one additional time. This additional interaction with the material comes at least two days after the first interaction because the original work is not returned to the students until the following class period. The self-assessment, therefore, requires students to space their practice even if they massed practice or "crammed" to study for the UC—which is important for maximizing retention and transfer of concepts and skills (Moulton et al., 2006; Butler, 2010; Dunlosky et al., 2013; Kang, 2016; Weinstein et al., 2018). In this context, the effect of spaced learning is enhanced when combined with the beneficial effects of testing on retention (Delaney et al., 2010; Roedinger and Butler, 2011).

Assessment has been shown to increase student metacognition and, therefore, enhance learning (Butler et al., 2008). Although UCs are summative assessments in the context of the course structure, the self-evaluation with rationale enhances their formative nature. Students have also received extensive formative feedback prior to each UC in the context of graded reading guides for articles discussed in class and class discussions. These formative assignments, along with the self-assessment portion of the UCs, may combine to increase student metacognition and enhance learning. Future work could evaluate the relative effectiveness of primary article reading guides and discussions versus UCs for enhancing student metacognition and learning.

REFERENCES

Ambrose SA, Bridges MW, DiPietro M, Lovett MC, Norman MK (2010) How do students become self-directed learners? In: How learning works. pp188-216. San Francisco, CA: Jossey-Bass.

Butler AC (2010) Repeated testing produces superior transfer of learning relative to repeated studying. J Exp Psychol Learn Mem Cogn 36(5):1118-33.

Butler AC, Karpicke JD, Roedinger HL (2008) Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. J Exp Psychol Learn Mem Cogn 34(4):918-928.

Cotner S, Ballen CJ (2017). Can mixed assessment methods make biology classes more equitable? PLoS One 12(12): e0189610.

Delaney PF, Verkoeijen PPJL, Sprigel A (2010) Spacing and testing effects: a deeply critical, lengthy, and at time discursive review of the literature. Psych Learn Motiv 53:63-417.

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.

Kang SHK (2016) Spaced repetition promotes efficient and effective learning: policy implications for instruction. Pol Insights Behav Brain Sci 3(1):12-19.

McKeachie WJ, Svinicki M (2006) Assessing, testing, and evaluating: grading is not the most important function. In: McKeachie's teaching tips: strategies, research, and theory for college and university teachers. pp74-77. Boston, MA: Houghton Mifflin.

Moulton CE, Dubrowski A, MacRae H, Graham B, Grober E, Reznick R (2006) Teaching surgical skills: what kind of practice makes perfect? Ann Surg 244(3):400-409.

Roedinger HL, Butler AC (2011) The critical role of retrieval practice in long term retention. Trends Cogn Sci 15(1):20-27.

Schaefer J (2015) The BRAIN Initiative provides a unifying context for integrating core STEM competencies into a neurobiology course. J Undergrad Neurosci Educ 14(2):A97-A103.

Weinstein Y, Madan CR, Sumeracki MA (2018) Teaching the science of learning. Cogn Res Princ Implic 3:2. Available at https://doi.org/10.1186/s41235-017-0087-y.

Received August 3, 2018; revised September 24, 2018; accepted September 25, 2018.

The author thanks Dr. C. Favero for feedback after using this exercise in her course.

Address correspondence to: Dr. Jennifer E Schaefer, Biology Department, PO Box 3000, College of St. Benedict/St. John's University, Collegeville, MN 56321. Email: jschaefer@csbsju.edu