Using Performance Assessments to Bridge the Gap between Accountability and Learning

Dennis Christopher, MS, Assessment Specialist
Les Sewall, Ed.D., Assessment Specialist
Questar Assessment Inc.

> Tests have a bad reputation in education circles these days: They take time, the critics say, put students under pressure, and, in the case of standardized testing, crowd out other educational priorities. But the truth is that, used properly, testing as part of an educational routine provides an important tool not just to measure learning, but to promote it.
>
> We need to change the way we think about testing. It shouldn't be a white-knuckle finale to a semester's work, but the means by which student's progress from the start of a semester to its finish, locking in learning along the way and redirecting their effort to areas of weakness where more work is needed to achieve proficiency.
>
> H.L. Roediger
> *"How Tests Make Us Smarter"*
> New York Times, July 18, 2014

A lot of pressure is being placed on students and educational entities to meet standards of proficiency. Students must pass high-stakes tests to graduate from high school. Schools must meet a level of proficiency to keep their standing within the community and state, as well as maintain steady levels of funding. The state must show that their students, through their schools, are meeting proficiency standards to qualify for funding from the Federal government.

All of this pressure falls on the teachers and the students. Yet students and teachers are the ultimate clients for schools, school districts, state departments of education, the United States Department of Education, education providers and taxpayers. With that in mind, how can we help students become successful? Part of learning is knowing how to use educational strategies to help students retrieve information to which they have been exposed. Those strategies include thought-provoking educational materials (e.g., performance assessments) combined with testing.

Testing

Along with good teachers, proper technology and other educational materials, students need to be tested. One could certainly debate the amount of testing students need; however, testing is a necessary part of education. Not only does testing show what the student knows and needs to know, it also teaches the student how to learn.

"Various kinds of testing when used appropriately encourage students to practice the valuable skill of retrieving and using knowledge." (Roediger, 18, July 2014) Testing arranges the student's understanding of the content in the brain and makes it easier to remember, or retrieve, the content. Part of the learning process is the ability to retrieve information.

Information retrieval can be practiced using assessments. In the classroom, exams are most likely formative assessments, which are a major key to student learning. As Eric Shepherd and Janet Godwin wrote in "Assessments through the Learning Process" (2010):

> Formative assessments provide feedback to individuals and their counselors during the learning process by providing search and retrieval practice. When people must provide answers to questions about material they've learned, their brains must search their memories and retrieve the information. These memory processes help solidify the learners' knowledge and help maintain that information in an accessible state for later recall.
>
> Not only did [students who took formative assessment] discover their level of competence, but they also inadvertently reduced their forgetting curve by experiencing some search and retrieval practice.

Formative assessments are used to define what students currently know and what they should learn next. As students are introduced to information, it is stored as chemical energy in the brain. The information is attached to other pieces of like information. These connections are the key to finding and retrieving that information. It is the practice of retrieving the information that leads to learning and higher scores on exams. In the process of taking formative exams students make connections to where pieces of information are stored in their brain, making retrieval of the information that much easier, and increasing the likelihood of doing well when taking standardized exams later. Roediger (2014) made these key points in his article about the importance of being able to retrieve information:

- Many studies reveal that much of what we learn is quickly forgotten; a central challenge to learning is finding a way to stem forgetting.
- Material tested right after [information was digested] was remembered better later.
- Various kinds of testing, when used appropriately, encourage students to practice the valuable skill of retrieving and using knowledge.
- The fact of improved retention after a quiz makes the learning stronger and embeds it more securely in memory.
- Testing [allows for] new learning within the context of regular classes and study routines.
- The benefit of quizzing remained in a follow-up test eight months later.
- This isn't just a matter of teaching students to be better test takers. As learners encounter increasingly complex ideas, a regimen of retrieval practice helps them to form more sophisticated mental structures that can be applied later in different circumstances.

At the end of the unit, term or year, summative assessments are used. These exams are designed to evaluate the learning system's capacity to help students learn. Teachers need to know how the connection between formative and summative assessments help students learn and do well on standardized exams.

Some recommended methods exist to facilitate retrieval of information embedded in the brain. Roediger (2014) wrote:

Retrieving knowledge from memory is more beneficial when practice sessions are spaced out so that some forgetting occurs before you try to retrieve again. The added effort required to recall the information makes learning stronger.

- It also helps when retrieval practice is mixed up — whether you're practicing hitting different kinds of baseball pitches or solving different solid geometry problems in a random sequence, you are better able later to discriminate what kind of pitch or geometry problem you're facing and find the correct solution.
- Students in classes with a regimen of regular low- or no-stakes quizzing carry their learning forward through the term, like compounded interest, and they come to embrace the regimen, even if they are skeptical at first.
- Notably, Mary Pat Wenderoth, a biology professor at the University of Washington, has found that this benefit holds for women and underrepresented minorities, two groups that sometimes experience a high washout rate in fields like the sciences.

For students in identified groups that experience lower retention rates, these testing techniques help students learn; retrieve information for exams; apply information to other situations; and lower test anxiety.

Performance Assessments

The Definition and Purpose of Performance Assessments

The United States Department of Education defines a performance assessment as "... a form of testing that requires students to perform a task rather than select an answer from a ready-made list." (Zimmermann, 1993) This definition implies that a performance assessment leads to a student-generated product or extended response.

The performance assessment approach varies greatly from the common selected response format which has often been criticized for focusing too much on the lower cognitive levels as described by Bloom's Taxonomy or Webb's Depth of Knowledge Levels.

> Although facts and concepts are fundamental in any undergraduate STEM (Science, Technology, Engineering and Mathematics) course, knowledge of methods, procedures and analysis skills that provide context are equally important. Student growth in these latter facets prove somewhat difficult to evaluate, particularly with conventional

> multiple-choice examinations. Performance assessments, used in concert with more traditional forms of assessment, are designed to provide a more complete picture of student achievement. (Slater, 2015)

While performance assessments may show that the student is able to think on a higher level, the product of a performance assessment may not tell us what the student actually knows about the topic. As Slater (2015) wrote, "Performance assessments are typically inappropriate for measuring student knowledge of facts." Having a student build or construct a product does not necessarily imply that the student understands the content behind the production.

In "Including Performance Assessments in Accountability Systems: A Review of Scale-up Efforts," Tung (2010) states that:

> In K–12 education, 'performance assessments' set forth expectations for students that require them to:
>
> - Create an original answer or product
> - Use higher order thinking and 21st century skills
> - Demonstrate thinking processes
> - Evaluate real world situations

Performance assessments are generally considered to consist of a stimulus (passage, audio, video, chart, graphic, etc.) followed by a series of individual items all tied to the stimulus, culminating in a product. Although performance assessments may take on different formats and different names depending on the school system or state, performance assessments all include some type of stimulus, a stem to set up the questions that follow, and a series of questions in differing formats. The questions could be multiple choice single answer, multiple choice multiple answer, hot spot, short essay, and many other types of items. This variety of item types allows for the assessment of knowledge at all cognitive levels.

Performance Assessments Often Cover Multiple Standards

Even if performance assessments are written as traditional constructed response items, they are unique in that they can cover several standards at the same time, which is essentially how content is taught in the classroom. It is probably a fallacy to state that one single standard covers only one piece of content. Although a standard may be a discrete unit, the content covered in that standard may rely upon content from several other standards. The Next Generation Science Standards (NGSS) point out connections between standards and recommends that the content is taught in that fashion. In "HS Evidence Statements: Executive Summary of the Front Matter," the authors state that, "Classroom instruction should be focused on helping students build towards several PEs at one time because many concepts and practices are interrelated." (Achieve, Inc., 2015, p. 2)

Would it make sense to have a series of performance assessments with multiple standards assessed inside of each performance assessment? This approach is not that far removed from the way English Language Arts (ELA) constructed response items are formatted. It also is not that far from testing for certifications and licensing. An ELA exam has a reading passage/stimulus followed by a series of questions. These ELA constructed responses could cover multiple standards. A certification exam could have a stimulus in the form of a case study followed by a series of questions that follow a process. Certification exam stimulus items may also cover multiple strands.

Even though performance assessments may take more time to deliver, covering multiple standards at one time may justify the time. Most exams are written with single unrelated items that can provide an idea of the information the student has retained, but cannot demonstrate whether or not the student can use the information on a higher cognitive level. Most exams are not built to evaluate student thinking above the "understanding" level defined in Bloom's Taxonomy. Higher order thinking skills require the student to match and compare information across standards and content to show application, analysis, synthesis, and evaluation. Performance assessments naturally cross standards and/or content allowing for simultaneous diagnosis of understanding at higher cognitive levels.

Performance Assessments and Simulations

This multiple-standard approach is also being used in curriculum settings. The newest and most advanced curriculum products use simulation and game theory, combined with assessment techniques, to teach and evaluate student understanding. This new curriculum approach can cover several standards at one time. Although interactive simulations can be expensive to produce, and therefore out of the price range of many school districts, lower technology simulations using embedded audio and video could be produced at a reasonable cost.

Simulation-enhanced learning can potentially increase test scores. Simulations put the student front and center in the learning activity. If the student is engaged, the brain produces endorphins that cause attention span and information uptake to increase dramatically. Neurochemistry researchers call this enhanced brain energy a "flow state." As Steve Kotler (2015) stated in a video titled, "The Neurochemistry of Flow States":

> ...a quick shorthand for how learning works is the more neurochemicals that show up during [the] experience, the better chance that experience has of moving from short term holding into long term storage.... Neurochemicals, among their many other functions,...one of them is to tag experiences. Big neon sign saying 'really important, save for later', because flow is this giant neurochemical dump. It massively amplifies learning.

Many commercial educational products offer simulations that allow the student to conduct research, answer questions, and produce a product. In most cases the simulations simply ask the student to answer questions while manipulating a simulation. In the case of two highly rated programs, the students not only answer questions about the content, they must produce a paper on the topic that would be presented to a specific audience. Engaging the student to produce a performance piece

based on a fun interaction helps to create a flow state as discussed by Kotler, (2015). The student is more likely to remember the information covered.

Most simulations ask the student to use the GRASP model as described by Wiggins and McTighe in "Performance Assessment: GRASP" (2004). GRASP, stands for Goal, Role, Audience, Situation and Products/Performance. Some variations on this model include a final S which stands for Standards. This model lays out the key points for a scenario-based authentic task, or performance assessment.

G = What is the **GOAL** in the scenario?

R = What is the **ROLE**?

A = Who is the **AUDIENCE**?

S = What is your **SITUATION**?

P = What is the **PERFORMANCE** challenge?

S = By what **STANDARDS** will work be judged in the scenario?

With GRASP, students are engaged in a scenario that makes them a part of a story. The student becomes immersed in a situation in which they must produce something that will be inspected or presented to a group of experts. In GRASP, expectations for the product and how it will be evaluated are outlined for the student. The student, in essence, becomes part of a role play that may happen in the real world.

Brett Moulding, Director Partnership for Effective Science Teaching and Learning and an NGSS Writer, stated in a video by the Teaching Channel titled, "NGSS: A Vision for K-12 Science Education" (2015, March 3):

> When we learn things it isn't for memorizing pieces of information. Just reciting science facts and information isn't what we want children to do. We want them to go out into the world and make sense of novel phenomena. So, making sense of things really is engaging in a performance in saying, I need to construct an explanation in why or how this occurs.

Conclusion

Testing is an integral part of the learning process. Testing allows for the continued retrieval of information that the brain is exposed to during learning. The ideas discussed above are arguably curriculum pieces that could be used as formative assessment. The student interacts with the simulation and is given feedback for future learning. Could this idea become part of summative assessment? The Next Generation Science Standards recommend such a direction (DeBarger, Penuel, & Harris, 2013, p.5).

Imagine a student working through one of these simulations as formative assessment. The teacher uses this format and follows the recommendations for helping students learn. At the end of the unit, quarter, or term the student is given a similarly formatted simulation as a standardized exam. Because the student is familiar with this format, has learned how to retrieve information, and is engaged in the assessment, you can imagine the student testing in a lowered stress environment and being able to retrieve key content information, possibly increasing test scores.

References

Achieve, Inc. (2015). Retrieved March 24, 2015 from website: HS-ETS1 Engineering Design | Next Generation Science Standards.

DeBarger, A. H., Penuel, W. R., & Harris, C. J. (2013). Retrieved from K-12 Center at ETS website: Designing NGSS Assessments to Evaluate the Efficacy of Curriculum Interventions.

Kotler, S. (2015, February 27). Retrieved from video file: "The Neurochemistry of Flow States, with Steven Kotler".

Roediger, H. L. (2014, July 18). Retrieved from website: How Tests Make Us Smarter.

Slater, T. F. (2015, March 24). Retrieved from website: CL-1: Field-tested Learning Assessment Guide (FLAG): CATs: Performance Assessment.

Teaching Channel. (2015, March 3). Retrieved from video file: Next Generation Science Standards

Tung, R. (2010, January). Retrieved from a full-text document that includes: Performance Assessments in Accountability Systems: A Review of Scale-up Efforts

Wiggins, G., & McTighe, J. (2011, March). Retrieved from website: Performance Assessment: GRASP.

Zimmermann, J. (1993, September). Retrieved from website: Archived: PERFORMANCE ASSESSMENT.