FEATURES

Performance-Based Assessment

The Road to Authentic Learning for the Gifted

Joyce VanTassel-Baska, EdD[1]

Abstract: Performance-based assessment clearly represents an indispensable approach for assessing gifted student learning. Challenging performance tasks allow gifted learners to reveal their considerable intellectual capacity and energy. Through performance tasks, teachers gain insights into a gifted student's true level of capability in a domain of knowledge. As the majority of programs for the gifted employ a project-based approach to curriculum, there is a real need to use a matching assessment model. Performance-based assessment, which includes product assessment, provides just such a match. Moreover, the new assessments for the Common Core State Standards use performance-based assessments as a main format for items as a way to judge the acquisition of higher level skills like developing argument. Thus, the incorporation of performance-based assessment in core content areas would appear a necessary part of designing effective programs for gifted learners and assessing them appropriately.

Keywords: assessment, problem solving, gifted education

Performance-based assessment clearly represents an indispensable approach for assessing gifted student learning. Challenging performance tasks allow gifted learners to reveal their considerable intellectual capacity and energy. Through performance tasks, teachers gain insights into a gifted student's true level of capability in a domain of knowledge. As the majority of programs for the gifted employ a project-based approach to curriculum, there is a real need to use a matching assessment model. Performance-based assessment, which includes product assessment, provides just such a match. Moreover, the new assessments for the Common Core State Standards (CCSS) use performance-based assessments as a main format for items as a way to judge the acquisition of higher level skills like developing argument. Thus, the incorporation of performance-based assessment in core content areas would appear a necessary part of designing effective programs for gifted learners and assessing them appropriately.

> "Through performance tasks, teachers gain insights into a gifted student's true level of capability in a domain of knowledge."

Rationale for the Use of Performance Tasks to Assess the Learning Levels of Gifted Students

The criteria for the creation of good performance assessment items parallel several criteria for the development of sound curriculum for gifted learners. Such criteria call for being open-ended, focusing on higher level thinking and problem solving, and stressing articulation of the thinking processes employed (i.e., metacognition). These features then, incorporated into an assessment protocol, should provide evidence of the level of performance in gifted program classrooms as these same features are cornerstones of most curriculum development efforts, regardless of type of program approach. Thus, a high score on performance assessment items should represent well high-level classroom performance in a gifted program focused in a given domain of learning.

Performance-based assessment provides an alternative way of looking at student ability via contextual performance. The item prototypes developed not only represent the scope of the domain under study, they also represent the major higher level modes of thinking in that domain, a primary issue of interest to gifted educators. Thus, in the new CCSS English language arts assessment prototypes, students are required to demonstrate

DOI:10.1177/1076217513509618. From [1]College of William and Mary. Address correspondence to: Joyce VanTassel-Baska, College of William and Mary, 427 Scotland Street, Williamsburg, VA 23185, USA; email: jlvant@wm.edu.

competency in providing evidence to support ideas, competency in analyzing difficult text, and competency in evaluating comparative text selections.

Critical reading and writing behaviors are a central part of what is being assessed by the new CCSS. For example, Grade 6 students are asked to do the following writing task demand, based on reading two short passages supplied.

> In the Demosthenes biography and the Icarus and Daedalus myth, the main characters exhibit determination in pursuit of their goals. Did determination help both main characters reach their goals or did it lead them to tragedy? Write an argument for whether you believe determination helped or hurt the two main characters.

This task demand directs students back to the two texts to provide details on how determination affected the outcome in each passage and to identify similarities and differences in the way it applied. This assessment prototype is used at most grade levels to assess these higher level skills just as persuasive writing has been a major part of programs for the gifted to enhance the same set of skills.

Finally, research evidence suggests that economically disadvantaged and minority learners perform better on tasks that emphasize fluid over crystallized intelligence (Mills & Tissot, 1995), spatial reasoning over verbal and mathematical (Naglieri, 1999). By employing an assessment approach that contains a strong spatial component, disparities by socioeconomic status (SES) levels or ethnic group may be reduced (Bracken, 2000).

Considerations in Developing Performance-Based Assessment Tools

Performance-based assessment, as an authentic tool for judging learning, offers many advantages for enhancing instruction. These include (a) the use of results as a diagnostic to determine what curriculum needs to be taught and at what level, (b) the use of results for flexible grouping within subjects, and (c) the use of results for instructional emphases or even reteaching of core concepts. However, constructing good performance-based assessments requires attention to important details in the design process. Several considerations important in developing and implementing more authentic and performance-based assessment systems with high-ability learners are addressed below.

Target High-Level Skills

Given the depth and complexity of gifted learners' cognitive abilities, tests for this population should emphasize high-level thinking and processing skills. That is, the test should go beyond simple recollection of knowledge or facts and require students to operate at higher levels of app1ication, analysis, synthesis, and evaluation. Task demands for gifted learners can make use of thinking processes, often identified as central to differentiation, such as, comparing, classifying, induction, deduction, constructing support, abstracting, decision making, investigation, problem solving, and invention (Marzano, Pickering, & McTighe, 1993). By the same token, expectations for students' performance conveyed, for instance, through scoring rubrics should reflect the same high standards for complexity and sophistication to bring out the best products that gifted learners are capable of generating.

Use Multiple Approaches

To monitor student performance and inform instruction, a teacher needs to collect student performance data all the way through a learning module or unit, using formative and summative assessments. Formative assessments are used to monitor student progress during instruction, while summative assessments are given at the end of instruction for the purpose of certifying mastery or assigning grades (Gronlund, 1998). While some approaches are more suitable for one type of assessment (e.g., portfolios may be used for formative, rather than summative assessment), some approaches can be used for both. To examine a student's performance from various perspectives and under different conditions, it is desirable for teachers to employ multiple assessment approaches in oral and written forms. A combination of approaches generally works to both the teacher's advantage and the student's advantage because different approaches can supplement one another to provide a more comprehensive picture of a student's performance.

Clarify Purpose

An emphasis on performance-based tasks does not replace standardized tests when the latter may function effectively. For instance, although a performance task can allow students to demonstrate their actual writing ability, students may also construct their own sentences in such a way as to bypass their weak areas in sentence structure. If language mechanics are the purpose for an assessment, then a standardized test can better cover a large number of grammar and language points in a relatively short time. It is a more efficient tool for examining students' mastery in key areas. The appropriate assessment approach should always be based on the purpose of the assessment. Generally, if content mastery is being assessed, a paper-and-pencil test with close-ended items may be preferable. If higher order thinking and problem solving are being assessed, a more performance-based approach would be appropriate.

Think Through How to Use Assessment Results

Differentiation for gifted learners typically calls for the use of advanced content, deep processing, and quality products. Where differentiation is occurring, gifted students tend to get harder books to read and more challenging projects to complete than their regular classmates. How do teachers assess their learning outcomes in such a way that these students feel

properly rewarded for their extra labor? How can we encourage gifted students to strive for a higher level when they always compare favorably with their peers in the classroom? And, in attempting a challenging project, how should teachers appropriately weigh the emphasis on their efforts and final results? A disturbing finding that emerged from two district-wide evaluations of gifted programs was that gifted students are not evaluated regularly for their learning in programs (VanTassel-Baska, 2006). Not only is it difficult to provide challenges for the gifted, often we do not document how they handled those challenges to know more about restructuring curriculum the next time. Teachers must consider how to document results and use them for future planning and for evidence of student growth.

Creating Performance-Based Assessment Task Demands

The process for constructing performance-based measures can be lengthy, yet shortcuts are possible if we deliberately apply techniques used in earlier efforts (see VanTassel-Baska, Johnson, & Avery, 2002). These techniques provide ways to construct meaningful tasks that align with curriculum standards and meet technical adequacy considerations.

Selection of Prototypes

To find appropriate prototypes that encompass verbal, math, and spatial spheres, educators need to review several sources. The CCSS standards guidebooks, developed by the Standards Committee for the National Association for Gifted Children, represent an important source of examples of differentiated curriculum and assessments in math and language arts (Hughes, Kettler, Shaughnessy, & VanTassel-Baska, 2013; Johnsen, Riser, & Assouline, 2014; Johnsen & Sheffield, 2012; VanTassel-Baska, 2013) A new guidebook has also been developed for use in constructing differentiated curriculum task demands and performance-based assessments in science (C. Adams, Cotabish, & Ricci, 2014)

Criteria Used for Task Development

A core set of criteria are essential building blocks in creating strong performance-based tasks to assess the learning of gifted students. One criterion is an emphasis on thinking and problem solving to tap fluid rather than crystallized abilities in a domain. A second criterion is to develop off-level tasks, ones that would be challenging to high-ability learners. A third criterion involves the use of an open-ended format to encourage more creative responses and ways of thinking. A fourth criterion deals with the use of manipulatives, a strategy found useful in aiding students in "figuring out" hard problems and especially recommended for use with at-risk students (Ford, 1996; VanTassel-Baska, 1992). Lastly, the criterion of "thinking made visible" should be applied to each task to encourage students to reflect on their problem-solving approaches and self-correct as needed.

Off level/advanced. Because the population of interest is high-ability learners, the criterion of developing "off-level" tasks is crucial. The power of the tasks ultimately lay in the ability to challenge the learner at an authentic level. By using an advanced and open-ended task, students are not in danger of bumping up against an artificial ceiling, a common problem with traditional in-grade achievement tests for these learners. In many advanced tasks, students have the opportunity to demonstrate sophistication in their thinking through their writing—a common approach for assessing reasoning ability (Paul, 1992). Moreover, they are encouraged to be fluent in expressing and elaborating their ideas.

Open-ended format. Many performance-based tasks should be open-ended, either because multiple answers are possible or because different approaches to answers are possible. When the former case is operant, students are instructed to write as many solutions as they can find. It prompts students to find multiple solutions at three levels of complexity. Students are given a fairly wide framework within which they can show how well they can see patterns. Multiple responses are judged equally effective as long as basic parameters of the problem are honored. Elaboration of response is also encouraged and rewarded.

Emphasis on articulation of thinking processes. In performance assessments, students are expected to provide some evidence of the thinking processes used in obtaining a solution for verbal, mathematical, and spatial-visual tasks. In some tasks, students are asked to show in words, pictures, or symbols how they reached their solutions. In other tasks, articulating the solution to a nonverbal problem by writing it out is an important part of the task.

Development of Rubrics and Exemplars

The rubric development process also involves a careful delineation of a range of responses obtained on a 0 to 4 scale from high-level response (4) to a low-level response (0). Rubric scores are used to discriminate among student performances. Once pilot test data are obtained, a set of exemplars can be developed for each point total value to aid in understanding and scoring the tasks. Answer sets for each task can be constructed and used as a basis for each rubric score.

Using Existing Models for Performance-Based Assessment

While creating new assessments can be accomplished by using the steps outlined, many educators may wisely opt to use or adapt existing performance-based assessments that already meet technical adequacy and have a history of successful use.

Pretest Response	Posttest Response
I don't think earthworms like light, because most of them live underground unless it rains or something and they get washed out of the dirt. I could always do an experiment to make sure, thow. For an experiment, I might taken an earthworm, with some kind of light, an dirt, and see if it stays out in the light, or trys to get away from the light by going under the dirt. Score: 5	Title: "Are bees attracted to diet cola?" Hypothesis: I don't think bees are attracted to diets just to regular. For example: coke, sprite, Dr. Pepper Materials: Bee, diet cola, container Description of what I would do: Take one can of diet cola and pour about 1 cup of it into a dish, bowl, etc. Then release a bee about a foot away and see if it moves toward the diet cola. If it does- you know bees like diet cola, but if it moves away from the diet cola, or doesn't respond to it you know bees don't like diet cola. When you are done with your experiment carefully release you bee, pour out your soda, and put back the way you found them. What will you record: If the bees are attracted to the diet cola or if they are attracted to the none diet liquids. Data Table: Trys: 1 2 3 4 5 6 Reactions: Score: 12

Figure 1. Sample fifth-grade response.

An Elementary Level Performance-Based Science Assessment: The Diet Cola Test

The Diet Cola Test was developed by Cain (1990) to assess students' understanding of experiments. It is an open-ended test that requires students to design an experiment to determine whether bumble bees are attracted to diet cola. A parallel form to the Diet Cola Test, the Earthworm Test, asks students to design an experiment to find out whether earthworms are attracted to light (C. M. Adams & Callahan, 1995). Both instruments were adopted for use on a pretest and posttest basis for their adequate reflection of the unit objectives to develop student experimental research skills, the similar age range targeted, and their sufficiently high ceilings (VanTassel-Baska, Bass, Ries, Poland, & Avery, 1998).

Students' responses are scored according to a checklist of science process skills, with points assigned for addressing each skill and additional points for skills addressed in greater detail: plans for *safety*, stating the *problem* or *question, giving a hypothesis* describing three *steps* or more, arranging steps in a *sequential order, listing materials* needed, plans to *repeat testing, defining terms*, plans for *observation*, plans for *measurement*, plans for *data collection*, plans for *interpreting data*, plans to make *conclusions based on data*, and plans to *control variables.*

To illustrate students' increased understanding in experimental design and data collection after their exposure to the units, a sample response from a fifth grader is presented in Figure 1. The Figure 1 example demonstrates the growth in student understanding of experimental design from a raw score of 5 to 12. He has been able to structure a design in an acceptable format, describe elements of the process, and set up a rudimentary data table. Because of their emphasis on advanced level work, the assessment approaches employed in International Baccalaureate (IB) and Advanced Placement (AP) programs are illustrative of assessments commonly used with academically oriented learners.

The IB assessment model measures the performance of students against the main objectives of the program by using a combination of external and internal assessment methods in written and oral modes. External assessments are provided and scored by the International Baccalaureate Organization (IBO). Internal assessments, which are also provided by the IBO, are scored by classroom teachers who are required to send representative scores of high, low, and average levels to the IBO for verification of their having correctly used the scoring rubric. The purpose of this is to ensure that students are assessed fairly according to international standards. The IB Language A1 externally assessed exam includes the components of commentary and essay papers on seen and unseen texts and two written assignments of comparative and imaginative/ creative nature. The external assessments account for 70% of the overall Language A1 assessment. The internally assessed component consists of two compulsory oral activities, one commentary on a teacher-selected reading, and one oral presentation on a student-selected topic. The oral component accounts for 30% of the total assessment (IBO, 1999). Scoring rubrics for the written work typically contain six levels to differentiate the degrees of *none, little, some, adequate, good*, and *excellent* demonstration of required ability, skills, or presentation. These assessments demand such abilities as appreciation, interpretation, comparison, critique, analysis, evaluation, and creativity.

The AP exam for each of 38 courses provides another example of a secondary level performance-based approach, seeking carefully constructed and scored responses that require depth of knowledge and thought. The exams generally contain two question types: multiple-choice and free response. The multiple-choice section emphasizes the breadth of the student's knowledge and understanding of the content. The free-response section emphasizes the application of these core principles in greater depth in solving more extended problems, or analyzing complex issues and texts (e.g., College Board Advanced Placement Program, 1999a, 1999b, 1999c, 1999d). For example, a student taking an exam in English Language and Composition might be asked to analyze the rhetoric of a given passage; a student taking English Literature and Composition might be asked to use examples from literature selections he or she has read to support a generalization about character or theme in the literature. Students taking a science or statistics exam may be given a situation and asked to design an experiment to answer a question of interest. In general, the free-response questions are designed so that different students are able to draw on the different experiences and texts they have encountered in their courses to respond to the question, thus allowing choice for teacher and student while still maintaining a common course framework.

The free-response section is scored against carefully developed guidelines that are drafted by individual item developers, reviewed and revised collectively by a committee, and modified based on student responses. Scorers of the free-response section are trained to apply the guidelines using exemplary student responses. Sample free-response questions for all exams, demonstrating the emphasis on higher level thinking required of students, are available through the College Board at http://www.collegeboard.com.

Along with demonstrating emphasis on higher level thinking and problem-solving skills, the AP and IB exams also illustrate the proper use of different test formats to serve different purposes of assessment. Moreover, these exams are exemplary for high-stakes testing in terms of their careful construction with consideration of the technical concepts of validity, reliability, and ceiling effect.

Although the resources available to the College Board and IB for developing their assessments far exceed those available to the average classroom teacher or district curriculum developer, the procedures used by these organizations are useful for developing even small-scale classroom assessments. The emphases on determining key principles, concepts, and content for assessment; using multiple formats for question development; encouraging review by a group of educators and content experts; and revising careful scoring guidelines based on the test framework and student response are important considerations that educators may use as foundations for adapting their own assessments.

Conclusion

In this age of CCSS, gifted educators need to provide performance-based assessment protocols in all subject areas that meet the criteria outlined in this article for advanced, higher level thinking and problem-solving, and open-ended task demands that truly challenge gifted learners and provide demonstrable evidence of their learning at elementary and secondary levels.

Conflict of Interest

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

References

Adams, C., Cotabish, A., & Ricci, M. K. (2014). *Using the next generation science standards with gifted and advanced learners*. Waco, TX: Prufrock Press

Adams, C. M., & Callahan, C. M. (1995). The reliability and validity of a performance task for evaluating science process skills. *Gifted Child Quarterly*, *39*, 14-20.

Bracken, B. (2000, April). *An approach for identifying underrepresented populations for G/T programs: The UNIT test*. Presentation at the College of William and Mary, Williamsburg, VA.

Cain, M. F. (1990). The diet cola test. *Science Scope*, *13*(4), 32-34.

College Board Advanced Placement Program. (1999a). *5-years of free-response questions 1995-1999: English*. New York, NY: College Entrance Examination Board and Educational Testing Service.

College Board Advanced Placement Program. (1999b). *Released exam 1997: AP statistics*. New York, NY: College Entrance Examination Board and Educational Testing Service.

College Board Advanced Placement Program. (1999c). *Released exam 1998: AP environmental science*. New York, NY: College Entrance Examination Board and Educational Testing Service.

College Board Advanced Placement Program. (1999d). *Released exams: 1998 AP Physics B and Physics C*. New York, NY: College Entrance Examination Board and Educational Testing Service.

Ford, D. Y. (1996). *Reversing underachievement among gifted Black students: Promising programs and practices*. New York, NY: Teachers College Press.

Gronlund, N. E. (1998). *Assessment of student achievement* (6th ed.). Boston, MA: Allyn & Bacon.

Hughes, C., Kettler, T., Shaughnessy, E., & VanTassel-Baska, J. (2013). *A guide to Differentiation of the CCSS ELA Standards for Advanced Learners* (Vol. II). Waco, TX: Prufrock Press.

International Baccalaureate Organization. (1999). *International Baccalaureate Language A1 guide*. Geneva, Switzerland: Author.

Johnsen, S., Riser, G., & Assouline, S. (2014). *A teachers' guide to using the common core state standards with mathematically gifted and advanced learners*. Waco, TX: Prufrock Press.

Johnsen, S., & Sheffield, L. (2012). *A guide to mathematics differentiation of the CCSS for advanced and gifted learners* (Vol. I). Waco, TX: Prufrock Press.

Marzano, R. S., Pickering, D., & McTighe, S. (1993). *Assessing student outcomes: Performance assessment using the dimensions of learning model*. Alexandria, VA: Association for Supervision and Curriculum Development.

Mills, C., & Tissot, S. (1995). Identifying academic potential in students from underrepresented populations: Is using the Ravens Progressive Matrices a good idea? *Gifted Child Quarterly*, *39*, 209-217.

Naglieri, J. A. (1999). *The essentials of CAS assessment*. New York, NY: Wiley.

Paul, R. (1992). *Critical thinking: What every person needs to survive in a rapidly changing world*. Sonoma, CA: Foundation for Critical Thinking.

VanTassel-Baska, J. (1992). *Planning effective curriculum for gifted learners*. Denver, CO: Love.

VanTassel-Baska, J. (2006). A content analysis of evaluation findings across 20 gifted programs: A clarion call for enhanced gifted program development. *Gifted Child Quarterly*, *50*, 199-215.

VanTassel-Baska, J. (Ed.). (2013). *Using the common core state standards for English language arts with gifted and advanced learners*. Waco, TX: Prufrock Press.

VanTassel-Baska, J., Bass, G., Ries, R., Poland, D., & Avery, L. D. (1998). A national study of science curriculum effectiveness with high-ability students. *Gifted Child Quarterly*, *42*, 200-211.

VanTassel-Baska, J., Johnson, D., & Avery, L. D. (2002). Using performance tasks in the identification of economically disadvantaged and minority gifted learners: Findings from Project STAR. *Gifted Child Quarterly*, *46*, 110-123.

Bio

Joyce VanTassel-Baska is the Smith professor emerita at The College of William and Mary in Virginia where she developed a graduate program and a research and development center in gifted education. Formerly, she initiated and directed the Center for Talent Development at Northwestern University. She has also served as the state director of gifted programs for Illinois, as a regional director of a gifted service center in the Chicago area, as coordinator of gifted programs for the Toledo, Ohio public school system, and as a teacher of gifted high school students in English and Latin. She has published widely including 29 books and more than 550 refereed journal articles, book chapters, and scholarly reports. Her major research interests are on the talent development process and effective curricular interventions with the gifted.