Performance Assessments: A Review of Definitions, Quality Characteristics, and Outcomes Associated with Their Use in K-12 Schools

Lynne M. Bland

Chesterfield County Public Schools

Christopher R. Gareis

The College of William and Mary

Abstract

After nearly two decades of federal and state accountability requirements relying on conventional standardized assessments, Virginia and several other states are moving to create more balanced approaches to statewide assessment systems that include the use of performance assessments. But Palm (2008) states, "Performance assessment can mean almost anything" (p. 3). This review of extant literature explores varying ways performance assessments are defined, characteristics of quality performance assessments, and educational outcomes associated with their use in K-12 schools. A rudimentary definition of performance assessment is established at the outset of this article to provide a foundation for undertaking the review, which includes sources from empirical, theoretical, and anecdotal literature. Drawing from the exploration of quality characteristics and evident educational outcomes, a refined definition of performance assessments is offered by way of conclusion to the article with the intent of providing some direction to current and future educators seeking to innovate assessment and accountability practices.

During the past two decades, assessment practices in K-12 schools have shied away from

performance-based and constructed-response formats in favor of objectively scored assessments

such as multiple-choice, matching, true/false, and formulaic writing prompts. Gaps in

performance assessment research for approximately the last 20 years coincide with the onset of

the 2001 *No Child Left Behind Act* (NCLB), which advanced the use of standardized testing as a

means of assessing school children in America (H.R. 1 – No Child Left Behind Act of 2001,

2001; Stecher, 2010). However, this trend is shifting. NCLB has been replaced by the *Every Student Succeeds Act* of 2015 (ESSA), which includes provisions promoting the development of state accountability measures that assess higher-order thinking skills, may include multiple measures rather than single end-of-year measures, and may include measures of academic growth. At the same time, the Commonwealth of Virginia amended the state code in 2014 eliminating the administration of five Standards of Learning (SOL) assessments previously overseen by the Virginia Department of Education, each to be replaced by an "alternative assessment" to be developed by each local school division (*Code of Virginia* §22.1-253.13:3). The intent of this significant policy change in Virginia's accountability system was to reduce the use of standardized, fixed-choice type assessments and to promote the use of more authentic, performance-based assessments.

Set within the context of a move toward performance assessments, this literature review draws from 12 years of performance assessment research available in peer-reviewed journals. Examined are the variety of ways performance assessments are defined; the quality features of performance assessments; and outcomes associated with their use. Key terms such as (*K-12) performance assessment*, *authentic assessment*, *performance-based assessment*, *alternative assessment*, and *portfolio assessment* were used to search for articles that either directly or indirectly focus on the topic of this review. Twenty journal articles from the years 2006 to 2016 were located, which included empirical, theoretical, and anecdotal studies. A scant five articles from the years 1998 to 2005 appear that indirectly support the purpose of the research. Of these 25 articles, only 15 clearly fit the parameters of the search and, therefore, appear in this literature review. The lack of current research suggests that many teachers and administrators today may not be particularly well-prepared for creating and using performance assessments (having relied

for some time fixed-choice and commercially produced assessments) and, therefore, may benefit

from direction, professional development opportunities, and support on the topic (Frey &

Schmitt, 2010). Of course, any such shift in K-12 policy and practice necessitates the attention

of those responsible for preparing new teachers, as teacher educators respond to meet the needs

of the field.

Definitions of Performance Assessment

There are a variety of ways to label and define the work students do to demonstrate

knowledge and skills. The terms *performance assessment* and *performance-based assessment*

(PBA) are used interchangeably throughout much of assessment literature. A rudimentary

definition of performance assessment offered by projectappleseed.org (2018), a nationally

recognized organization known for its advocacy of performance assessments, states that a

performance assessment should allow students to demonstrate knowledge and skills, including

the process by which they solve problems. Turning to sources of extant literature, variances on

this definition are evident. Frey and Schmitt (2010) state that "*performance tests* measure skill

or ability; are sometimes called *alternative assessments* or *authentic assessments;* and scoring

often requires subjective judgment" (p. 109).

Oberg (2010) describes performance-based assessment generally as "one or more

approaches for measuring student progress, skills, and achievement" and that performance

assessments are "the ultimate form of linking instruction with assessment" (p. 5). Oberg (2010)

adds that performance-based assessment should be considered an *alternative assessment* when

the term *authentic* is added. An authentic performance assessment has students demonstrate

understanding and skills in a real-life context "rather than contrived problems for the classroom

setting" (Oberg, 2010, p. 5). Koh, Tan, and Ng (2011) assert that performance assessments are

intended to measure students' knowledge and skills at deeper levels than traditional assessments and tend to prompt students to solve authentic or real-world problems. Newmann et al. (1996) claim that *authentic intellectual work* engages students in much more than routine practices of gathering facts and applying procedures. Higher-order thinking and real-world problem-solving are two important elements that make up the substance of authentic assessment (Koh, Tan, & Ng, 2011, p. 139).

Frey, Schmitt, and Allen (2012) posit that a meaningful, real-world task or problem is characteristic of performance-based assessment at some level. Referencing a book-length critique of standardized tests by Archbald and Newman from 1988, they suggest that for an assessment to be considered authentic, it must have value beyond the actual score or grade, indicating that the assessment task itself should be meaningful (Frey, Schmitt, & Allen, 2012). This suggests that assessments such as multiple-choice standardized tests are not authentic. Along with Archbald and Newman, Frey, Schmitt, and Allen (2012) cite Wiggins as early advocates for authentic assessments:

> The views of Wiggins…support a definition of authentic assessment as assessment that poses an intellectually interesting and personally meaningful problem or task. These types of assessments are realistic because the questions, tasks, or problems have value and interest beyond the classroom into the 'real world' of students' values, abilities, and motivations. (p. 13)

Less critical of conventional assessments, Kan and Bulut (2014) consider a performance assessment simply an alternative method to measure what a student knows and can do. Expanding upon possible forms of performance assessments, Adeyemi's (2015) empirical study connects authentic and portfolio assessments to performance assessments because students'

selection of artifacts to include in a portfolio provide a real-world opportunity to show evidence of mastery of a set of skills, applied knowledge, and attitudes. As such, a portfolio also provides a "means for reflection and critiquing one's own work and evaluating the effectiveness of lessons of interpersonal interactions with students or peers" (p. 127).

Table 1 (next page) shows a side-by-side comparison by skill/ability, process, and value/attitude of various ways to consider performance assessment. The common language throughout suggests the synonymous nature and purpose of performance, authentic, and portfolio assessments. *Alternative assessment* is a term used periodically in the literature to denote assessments not associated with the traditional multiple-choice and fill-in-the blank types of assessment; therefore, the term appears in Table 1 as an umbrella for all others that characterize performance assessment.

While it would be apt to conclude here with an enhanced definition of performance assessment based upon the varying definitions described, the revised definition for purposes of this review is dependent upon the further consideration of quality characteristics of performance assessments, as well as evident educational outcomes of their use.

Quality Characteristics of Performance Assessments

A variety of features of high-quality performance assessments and their use are described in the selected literature. Corcoran, Dershimer, and Tichenor (2004) state that common characteristics of performance assessments should include asking students to perform, create, or produce something; tapping higher-level thinking and problem-solving skills; using tasks that represent meaningful instructional activities; involving real-world applications; and using human judgment in scoring. These characteristics typify performance assessments generally, while other characteristics are associated with specific forms of performance assessments.

Table 1

Summary of Common Characteristics of Alternative, Performance, Authentic, and Portfolio Assessments Recognized by Various Researchers

Alternative Assessment			
	Performance Assessment	**Authentic Assessment**	**Portfolio Assessment**
Skill/Ability	Measures skill or ability (Frey & Schmitt, 2010), clearly articulates what teachers should be teaching and what students should be learning, and allows for evaluation of both process and product (Lai, 2011)	Demonstrates knowledge, skills, or behavior in a real-life context rather than contrived problems for the classroom setting (Oberg, 2010)	Evaluates the effectiveness of lessons of interpersonal interactions with students or peers (Adeyemi, 2015)
Process	Demonstrates knowledge and skills, including the process by which they solve problems (projectappleseed.org) and provides insights to students' cognitive processes and their attitudes toward content (Corcoran, Dershimer, & Tichenor, 2004)	Poses an intellectually interesting and personally meaningful problem or task and engages higher-order thinking (Frey, Schmitt, & Allen, 2012; Koh, Tan, & Ng, 2011)	Shows evidence of mastery of a set of skills, applied knowledge, and attitudes (Adeyemi, 2015)
Value/Attitude	Has inherent value and prompts the interest and motivation of students beyond the classroom into the real-world (Frey, Schmitt, & Allen, 2012; Lai, 2011)	Realistic because the questions, tasks, or problems have value and interest beyond the classroom into the real-world of students' values, abilities, and motivations (Frey, Schmitt, & Allen, 2012; Koh, Tan, & Ng, 2011)	Provides students a real-world opportunity to select artifacts and entries for their portfolio (Adeyemi, 2015)

According to Frey and Schmitt (2010) in their empirical study, writing assignments (such as essays) are noted by some teachers as one of the most common forms of performance assessments used in classrooms. Written assessments can be considered performance-based when their purpose is to measure skill or ability and when they are subjectively scored (Frey & Schmitt, 2010). It is important to note that *subjective scoring* requires subject-specific knowledge on the part of the teacher to evaluate student performance. Corcoran, Dershimer, and Tichenor (2004) suggest journal entries as a meaningful performance assessment because "if it is

tied to teaching and learning objectives, the teacher can gain insight into students' cognitive progress and reveal their attitudes toward content" (p. 214). As another example, historical writing assessments evaluate the skills needed to support historical reasoning, which include analytical and evaluative thinking. Such higher-level thinking skills involve being able to analyze evidence, weigh conflicting accounts, consider biases, construct arguments grounded in evidence, understand the complexities of our social world, evaluate information responsibly, and ask difficult questions (Monte-Sano, 2008). Educators who use historical writing as a performance assessment provide students another avenue by which to develop their literacy skills and to apply them in an authentic way. Monte-Sano (2008) maintains, however, that many teachers do not have the prerequisite skills necessary to teach performance-based tasks such as evidence-based historical thinking and writing, and that professional development is needed.

Beyond the creation and use of performance assessments at an individual classroom level, the country of Singapore offers an example of a national effort to move to more authentic assessment practices. In 1997, the Singapore Ministry of Education articulated a vision of "Thinking Schools, Learning Nation" (TSLN) as a systemic effort to develop creative and critical thinking skills, as well as a lifelong passion for learning and nationalistic commitment, within its entire population of school children (Koh, Tan, & Ng, 2011). Singapore's TSLN involves teachers moving away from traditional approaches to teaching and learning, such as rote memorization, to more authentic learning experiences and assessment practices. Inherent to this initiative has been widespread engagement in professional development in teachers' and school leaders' understanding and use of performance assessments (Koh, Tan, & Ng., 2011). Koh, Tan, and Ng (2011) point out that changes in classroom assessment practices are necessary to promoting thinking schools. The mode of day-to-day classroom assessment practices must foster

engaged learning and enhance students' mastery of 21ˢᵗ century competencies (Koh, Tan, & Ng, 2011). Done well, these assessments reflect more intellectually challenging learning goals and include more authentic, open-ended assessments tasks such as "sustained written prose where students are asked to elaborate on their understanding, explanations, arguments, and/or conclusions" (Koh, Tan, & Ng, 2011, p. 140).

In addition to an evident connection between teaching and assessment practices, performance assessments may also serve formative purposes at the classroom level. Oberg (2010) advocates the use of performance assessments as a pre-assessment alternative to the traditional paper-pencil methods to support teachers in their instructional planning. Pre-assessments administered in an authentic format can provide teachers an important glimpse not only to students' prior experiences and knowledge but also to their current skills and interests relative to the intended learning outcomes of instruction (Oberg, 2010). Oberg (2010) suggests that teachers consider the following guidelines when developing performance-based pre-assessments:

1. What is important about the lesson and student learning that the teacher wants to know if the student knows;

2. How students can demonstrate knowledge in a unique or non-standard way;

3. What the criteria are for competence and mastery;

4. How to judge student competence;

5. How to provide feedback in a constructive manner;

6. How to include the student within the review process;

7. How the results will be used to guide instruction and differentiation.

These guidelines are intended to provide support when creating a performance-based pre-assessment. For instance, consider a mock store in the classroom to observe how well students can count money when purchasing items, as well as when receiving or giving change: "Teachers observe adding, subtraction, multiplication skills as well as problem solving skills, language skills, and social interactions" (Oberg, 2010, p. 6). Through the use of this performance and a teacher's observation checklist, a pre-assessment of authentically applied understandings and skill related to the intended learning of rudimentary financial literacy can be obtained.

Performance assessments can also be used for the purpose of monitoring student growth. VanTassel-Baska (2013) contends that performance assessments themselves serve as a basis for continued, authentic learning for students, including gifted students. Adeyemi (2015) demonstrates in a study of middle school students that authentic and portfolio assessments contribute to the teaching and learning process. Each of these assessment types involves students in the process, thus giving them a more meaningful role in improving achievement. Learners' achievement and their feelings of responsibility for monitoring their own progress provide an intrinsic motivation of interest in a task. Not only did the students in this study relate the different aspects of the information to one another, they also related them to their previous learning and personal experiences (Adeyemi, 2015).

As Stecher (2010) suggests, performance assessments can take a broad array of forms. For instance, Barber, King, and Buchanan (2015) report on a measure of performance assessment known as Digital Moments. Although this study was set in a course at the undergraduate level, it is indicative of similar technology enhanced innovations in K-12 schools. In this university example, students enrolled in a course entitled "Psychological Foundations and Digital Technology" and engaged in Digital Moments activities over a 12-week period. Students

watched three hours of podcasts per week and submitted reflections to a common online learning management system (i.e., Blackboard). Students were encouraged to structure their reflections in ways meaningful to them, which included words, phrases, pictures, colors, musical links, and more (Barber, King, & Buchanan, 2015). Students viewed their classmates' work and gained new ideas with immediate relevance to them regarding ways to complete their reflections and assignments. Students also collaboratively developed criteria and tools for assessing their own work and the work of others, as well as for providing feedback to each other. As such, Digital Moments functions as a means of assessing and learning, involving characteristics of authenticity, relevance, collaboration, and technology (Barber, King, & Buchanan, 2015).

In general, high-quality performance assessments should focus on important intended learning outcomes; engage students' higher-order thinking skills; integrate authentic tasks and problems; foster both independent and collaborative work; integrate assessment within the learning process; leverage technology when appropriate; and have meaningful success criteria. In addition, performance assessments may be used for each and any of the classic uses of classroom assessment, namely as pre-assessments, formative assessments, or summative assessments. Performance-based assessments are also developmentally appropriate for grades K-12, while also being appropriate in post-secondary education. They can be designed and used in face-to-face settings, as well as online learning environments. Finally, as alluded to at the start of this review of literature, they may be used as classroom assessments, district benchmark assessments, statewide accountability measures, and even national indicators of educational achievement.

Outcomes Associated with Performance Assessments

The outcomes, results, or effects associated with the use of performance assessments generally reveal that the depth of knowledge and types of skills developed by students are greatly influenced by the nature and format of the assessments themselves (Darling-Hammond & Adamson, 2010). Returning to the example from an undergraduate course, Barber, King, and Buchanan (2015) note outcomes of the use of an authentic assessment in a study of digital pedagogy as the following: "Students began to exhibit greater competence and confidence in using digital open resources, needed less direction from the instructor, and enjoyed taking the reins of their own learning. Also, learners developed autonomy, engagement, and motivation; self and peer assessments grew to be more meaningful; and there was a shared development of collective knowledge" (p. 65). As a point particularly relevant to the use of performance-based assessments in the PK-16 continuum of elementary, secondary, and post-secondary education, Barber, King, and Buchanan (2015) suggest that performance assessments have future implications for students regardless of level, building a foundation of problem-solving, self-directed learning, and constructive collaboration for future learning.

Other recent studies situated explicitly in K-12 educational settings garner other evident learning outcomes. Kim, VanTassel-Baska, Bracken, Feng, and Stambaugh, (2014) found indicators of science reasoning and conceptual understanding in the primary grades using standardized and performance-based assessments. In a study of the efficacy of authentic and portfolio-based assessment in learning social studies, Adeyemi (2015) concludes that both modes of assessment are correlated to academic success in learners and also that these types of assessments contribute to the formative use of assessment results by both teachers and students. Several studies from the height of implementation of NCLB accountability testing also provide

some evident outcomes of the use of performance assessments. In an article about young adult literature and alternative assessment measures, Comer (2011) asserts that performance assessments can be constructed and used to engage students metacognitively. Additionally, Comer, (2011) states that there is more opportunity for students to pull from knowledge they already have to answer questions, thus making active application of their knowledge. Hallam, Grisham-Brown, Gao, and Brookshire (2007) empirically studied the effects of outcomes-driven authentic assessment on classroom quality. Their study suggests that an authentic assessment approach, which includes performance assessments embedded within the curriculum, may have a positive impact on the language and literacy environment within classrooms and as outcomes for students. In *A Teacher's Guide to Alternative Assessment: Taking the First Steps*, Corcoran, Dershimer, and Tichenor (2004) write about students being more engaged and more willing to assess their own learning. They found that creative student projects reflect a range of intelligences, resulting from having students perform, create, produce, or do something involving higher-level thinking and real-world applications.

Although the past decade-and-a-half of high-stakes accountability testing has seen a precipitous drop in the study of performance assessment, the studies that have been published within this time provide some evidence of a number of desirable learning outcomes associated with the use of such alternative assessments. Table 2 (next page) summarizes selected examples of these studies, including their specific type of performance assessment, key characteristics, and observed learning outcomes.

Table 2

Samples of PBA Types and Outcomes

Type of PBA	Source	Characteristics of PBA	Evident Learning Outcomes
Essay	Frey & Schmitt (2012)	Essay type assessments can be considered performance-based when their purpose is to measure skill or ability and are scored by subject experts.	Increased learning in the classroom, as well as increased test scores
Journal	Corcoran, Dershimer, & Tichenor (2004)	Writing tied to the learning objectives and use human judgment to do the scoring	Insight into cognitive processes of students and attitudes toward the content
Historical Writing	Monte-Sano (2008)	Writing tied to analysis of evidence, weighing conflicting accounts, determining bias, constructing arguments, and asking difficult questions	Development of literacy skills and application in an authentic way
Pre-assessment	Oberg (2010)	A mock store in the classroom reflecting an authentic scenario to observe how well students can count money when purchasing items, as well as when receiving or giving change	Teachers observe adding, subtraction, and multiplication skills, as well as problem solving skills, language skills, and social interactions in an authentic format to help develop high-quality and effective instruction and curriculum.
Portfolio	Adeyemi (2015)	Students involved in the process of monitoring their own learning and communicating their learning and previous experiences to others	Intrinsic student motivation of interest in a task and feelings of responsibility for monitoring their own progress, which gives students a meaningful role in improving achievement
Digital Moments	Barber, King, & Buchanan (2015)	Students electronically complete tasks of their choosing, give and receive feedback to self and others, and immediately apply knowledge to move their project forward.	Effective use of feedback develops valuable online collaboration and communication skills and embeds assessment within the learning process.

The review of extant literature suggests the promise of performance assessments to result in real and substantive learning outcomes not only in the cognitive domain of learning but also in the socio-emotional domain. Notably, more than one study tied student performance on conventional assessments to their engagement with performance assessments. However, it is also important to note the reliance of performance assessments on the use of use of rubrics to

gauge student outcomes. The creation and use of rubrics is clearly integral to determining evident outcomes of learning and, therefore, merits scrutiny. For instance, an empirical study conducted by Kan and Bulut (2014) examined the effects of teacher experience and rubric use in performance assessments. In this study, eighth-grade students were given a performance task, and 17 teachers with and without a rubric graded their responses. The performance assessment first graded without the use of a rubric resulted in an inconsistency of scores among the teachers. When the assessments were graded several weeks later using a rubric, the consistency among the scores given by the teachers increased, thus supporting the use of rubrics when grading performance assessments. An additional noted outcome was that teachers with more years of teaching experience tend to score performance tasks more leniently than teachers who do not have long years of teaching experience. Additional research is needed to determine why this phenomenon occurs. However, another clear implication is the need for teacher educators to prepare pre-service teachers in the technical knowledge and skills necessary to create and use valid and reliable rubrics in order to achieve the promise of performance assessments.

Conclusion

The purpose of this literature review was to explore a sampling of research regarding definitions, quality characteristics, and evidence of educational outcomes of performance assessments. In this review, the concept of performance assessment was interpreted in different ways depending upon the context in which it was used. A basic definition from the educational advocacy group projectappleseed.org offers that *performance assessments require students to demonstrate knowledge and skills, including the process by which they solve problems*. This definition coincides with the extant sampling of research; however, other words and concepts began to appear in the literature that offered more depth. Table 3 (next page) identifies the

frequency of common words and concepts associated with performance assessments by various researchers that appeared in Tables 1 and 2.

Table 3

Content Analysis of Common Words and Concepts Associated with Performance Assessments by Various Researchers from Tables 1 and 2

Words and Concepts	Number of Occurrences in Tables 1 and 2
Skills/Abilities	12
Authentic or Real-world Contexts	11
Tasks	9
Value/Attitudes	9
Knowledge/Understanding	7
Process	6
Social/Emotional	5
Subjective Scoring	2

Based upon the frequency of these words and concepts projected in the review of the literature, an enhanced version of the rudimentary definition developed into the following: *Performance assessments, which can develop as a task or product, necessitate subjective judgment to measure students' abilities to authentically demonstrate knowledge, skills, and processes in a way that provides value, interest, and motivation to students beyond the actual score or grade.* This definition is not exclusive, nor do we contend that it is complete. Rather, it expands upon an oft-accessed definition from a popular online resource and provides additional depth and breadth to the definition based upon a sampling of extant research.

Additional research in the area of performance assessment will be beneficial as the United States moves from the standardized testing phase of the *No Child Left Behind Act* of 2001

to opportunity to innovate with more authentic assessments under the *Every Student Succeeds Act* of 2015. While not all standardized tests will be abolished under ESSA, there will be a move away from sole dependence upon standardized, fixed-choice assessments and more flexibility given to local school divisions to assess students in a variety of ways. In this respect, Virginia is a leader among a handful of other states (such as New Hampshire, New York, and Colorado) moving toward the innovative use of locally-developed performance-based assessments. This important and strategic move toward the more authentic assessment of real-world, transferable understandings and higher-order thinking skills will necessitate increasingly clear and commonly shared understandings of the definition, characteristics, and intended learning outcomes associated with high-quality performance assessments for in-service and pre-service teachers alike. Furthermore, as the experience of Singaporean educational system suggests, this move will necessitate the continued professional development of in-service teachers to create and use performance assessments, as well as the revision of teacher preparation coursework and field experiences to prepare novice teachers for the challenge and opportunity of this change (Koh, Tan, & Ng, 2011).

References

Adeyemi, B. (2015). The efficacy of authentic assessment and portfolio assessment in the learning of social studies in junior secondary schools in Osun state, Nigeria. *Ife PsychologyIA, 23*(2), 125-132.

Barber, W., King, S., & Buchanan, S. (2015). Problem based learning and authentic assessment in digital pedagogy: embracing the role of collaborative communities. *The Electronic Journal of e-Learning, 13*(2), 59-67. Retrieved from: http://www.ejel.org

Code of Virginia. The Commonwealth of Virginia. §22.1-253.13:3.

Corcoran, C., Dershimer, E., & Tichenor, M. (2004). A teacher's guide to alternative assessment: Taking the first steps. *The Clearing House 77*(5), 213-216. Retrieved from http://www.jstor.org/stable/30189900

Every Student Succeeds Act (2015) Pub. L. No. 114–95. Retrieved from https://www.congress.gov/bill/114th-congress/senate-bill/1177

Frey, B., & Schmitt, V. (2010). Teachers' classroom assessment practices. Middle Grades Research Journal, 5(3), 107-117.

Frey, B., Schmitt, V., & Allen, J. (2012). Defining authentic classroom assessment. *Practical Assessment Research & Evaluation, 17*(2), 1-18.

Hallam, R., Grisham-Brown, J., Gao, X., & Brookshire, R. R. (2007). The effects of outcomes-driven authentic assessment on classroom quality. *Early Childhood Research & Practice, 9*(2). Retrieved from http://go.galegroup.com/ps/i.do?id=GALE%7CA178716829&v=2.1&u=ccpshigh&it=r&p=GPS&sw=w

No Child Left Behind Act of 2001. (2001). Retrieved from https://www.congress.gov/bill/107th-congress/house-bill/1

Kan, A., & Bulut, O. (2014). Crossed random-effect modeling: Examining the effects of teacher experience and rubric use in performance assessments. *Eurasian Journal of Educational Research*, 57, 1-28.

Kim, H., VanTassel-Baska, J., Bracken, B., Feng, A., & Stambaugh, T. (2014). Assessing science reasoning and conceptual understanding in the primary grades using standardized and performance-based assessments. *Journal of Advanced Academics,*

25(1), 47-66.

Koh, K., Tan, C., & Ng, P. (2011). Creating thinking schools through authentic assessment: The

case in Singapore. *Educational Assessment Evaluation and Accountability 24*, 135-149,

doi: 10.1007/s11092-011-9138-y

Lai, E. (2011). Performance-based assessment: Some new thoughts on an old idea. Retrieved

from: http://researchnetwork.pearson.com/wp-content/uploads/Performance-based-

assessment.pdf

Monte-Sano, C. (2008). Qualities of historical writing instruction: A comparative case study of

two teachers' practices. *American Educational Research Journal, 45*(4), 1045-1079. doi:

10.3102/0002831208319733

Oberg, C. (2010). Guiding classroom instruction through performance assessment. *Journal of

Case Studies in Accreditation and Assessment, 1*, 1-11.

Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the

literature. *Practical Assessment, Research & Evaluation, 13*(4), 1-11.

Project Appleseed The National Campaign for Public School Improvement. (2018). *Testing,

testing* [Brochure]. St. Louis, MO. Retrieved from

http://www.projectappleseed.org/#!assessment/cwvf

Stecher, B. (2010). *Performance assessment in an era of standards-based educational

accountability*. Stanford, CA: Stanford University, Stanford Center for Opportunity

Policy in Education.

VanTassel-Baska, J. (2013). Performance-based assessment the road to authentic learning for the

gifted. *Gifted Child Today*, *37*(1), 41-47. doi: 10.1177/1076217513509618