Student Engagement: A Framework for On-demand Performance Assessment Tasks

Prepared by Catherine Taylor,
Kari Kokka, Linda Darling-Hammond, Jack Dieckmann,
Vivian Santana Pacheco, Susan Sandler, and Soung Bae

SCALE
Stanford Center for
Assessment, Learning,
& Equity

scope
Stanford Center for
Opportunity Policy in Education

Acknowledgement:

Student Engagement: A Framework for On-demand Performance Assessment Tasks was prepared with the generous support of the Sandler Foundation. We gratefully acknowledge their support. The research was conducted independently by Stanford University and does not represent the views of the sponsors.

About the authors:

Catherine Taylor, Senior Vice President Measurement Services at Measured Progress

Kari Kokka is a Math Research Associate at the Stanford Center for Assessment, Learning, and Equity

Linda Darling-Hammond is the Charles E. Ducommun professor of Education, emeritus, at Stanford University, and faculty director, Stanford Center for Opportunity Policy in Education (SCOPE)

Jack Dieckmann is the Associate Director of Curriculum at the Stanford Center for Assessment, Learning, and Equity

Vivian Santana Pacheco is a Research Analyst at the Sandler Foundation

Susan Sandler is a Trustee of the Sandler Foundation

Soung Bae is a Senior Research and Policy Analyst at the Stanford Center for Opportunity Policy in Education

Suggested Citation:

Taylor, C., Kokka, K., Darling-Hammond, L., Dieckman, J., Santana Pacheco, V., Sandler, S., & Bae, S. (2016). Student Engagement: A Framework for On-demand Performance Assessment Tasks. Stanford, CA: Stanford Center for Opportunity Policy in Education.

Stanford Center for Opportunity Policy Education
505 Lasuen Mall
Stanford, CA 94305
Email: scope@stanford.edu
http://edpolicy.stanford.edu

Stanford Center for Assessment, Learning and Equity
1705 El Camino Real
Palo Alto, CA 94306
https://scale.stanford.edu/

Introduction

Assessment consortia partnerships designing assessments aligned with the Common Core State Standards have promised innovative approaches that better reflect the full range of standards — including higher-order thinking and performance skills — and are fairer and more accessible to the full range of students in our diverse nation.

Engaging students in meaningful applications of their knowledge is a key aspect of both addressing the standards and providing greater access. Not only do the standards emphasize the importance of meaningful engagement in real-world tasks, but evidence shows that engagement is strongly related to student performance on assessment tasks, especially for students who have been typically less advantaged in school settings. In the traditional assessment paradigm, however, engagement has not been a goal of testing and concerns about equity have focused on issues of bias and accessibility. A common tactic to avoid bias has been to create highly decontextualized items. Unfortunately, this has come at the cost of decreasing students' opportunities to *create meaning* in the task as well as their motivation to *cognitively invest* in the task, thereby undermining students' opportunities to adequately demonstrate their knowledge and skills.

Innovative assessment designs such as performance tasks offer a ripe opportunity to develop tasks that engage all students of diverse backgrounds. Performance tasks provide opportunities for "leveling the playing field," allowing students to demonstrate their evaluation, synthesis, analysis, and application skills in more open-ended ways than multiple-choice items offer.

The goal of this paper is to introduce the ways in which dimensions of engagement may be meaningfully incorporated into assessment tasks so that all students are more fully motivated to complete the tasks and perform them well. In what follows, we begin with an overview of research on student engagement. We then present a set of design challenges, recommendations, and guiding questions for item writers' consideration (see Table 1), followed by design considerations for each of the following elements: Classroom Activities, Design of the Tasks, and Scoring of the Tasks. Summaries and resources of student engagement literature are also presented in Appendix A.

Student Engagement: A Brief Review of Research

Student engagement in classroom activities and assessments is acknowledged to be "a highly desirable goal with positive outcomes for all parties" (Bryson & Hand, 2007, p. 354). It is a complex of internal and external behaviors that are necessary for effective "[mental] interaction with content" (Moore, 1989). School engagement describes students' feelings, behaviors, and thoughts about their school experiences. It is an important predictor of academic outcomes such as achievement, standardized test scores, and high school completion. Bodovski and Farkas (2007) followed students from K through grade 3 and found that student engagement was a stronger predictor of achievement test scores than either initial test scores or time spent studying.

Engagement has been shown to be malleable and responsive to variations in the learning environment (Fredricks, Blumenfeld, & Paris, 2004). For example, engagement can be improved through changes in teachers' relationships with students, instructional strategies, and the nature of tasks and assessments (Dotterer & Lowe, 2011). Therefore, the dimensions of engagement that are important to attend to for developing engaging student assessments include **relevance, autonomy, collaboration**, and **authenticity**, and they are discussed below.

Relevance

Relevance refers to the process by which the learner perceives that the task will satisfy important personal needs, motives, or values (Keller, 1983) and fuels the student's motivation to learn (Brophy, 1986). More specifically, relevance answers the question, why does the educational content *matter* to the student? and provides the student with a reason for doing a task. Relevance may be fostered by making a connection to students' lived experiences, interests, or prior knowledge. These connections create a "need to know" for students and gives them a reason for doing the task.

When students connect to task scenarios that are relevant to their own lives, engagement and performance improves (Meier, 2008). For example, Walkington (2013) found that context personalization increased student performance on algebraic word problems. She studied 145 ninth graders of three Algebra classes where teachers utilized Cognitive Tutor Algebra, a computer based tutoring system that individualizes instruction through adaptive problem selection, hints, and feedback. Students in the experimental group solved algebraic word problems matched to their self-reported interests (e.g., sports, music, art, games) in prior student surveys and interviews. Connecting the math problems to students' interests increased student performance,

especially with more cognitively challenging problems and for struggling students. In addition, problems *relevant* to students' lives (e.g., paying a cell phone bill) were easier for students to solve than those not connected to their experience, even when they were contextualized to an expressed interest (Walkington & Sherman, 2012).

Relevance of assessment items is of heightened importance for engagement considerations of students of traditionally underserved groups. Students of privileged backgrounds have been found to be more able to compartmentalize and carry out decontextualized tasks that do not have immediate relevance for them (Darling-Hammond, et al., 2008) than traditionally underserved students. Historically privileged students also tend to be more "test wise" or savvy with the unspoken skills and strategies needed to successfully tackle traditional tests (Arbuthnot, 2011). These findings reiterate the importance of making tasks more relevant by making connections to students' lived experiences, interests, or prior knowledge.

Autonomy

Student engagement may also be improved through contexts that support *autonomy* (Connell & Wellborn, 1991). Autonomy derives from self-determination theory and refers to the extent to which one is able to choose or self-initiate an action or experience an action for which one is responsible (Deci & Ryan, 1987). Autonomy may be supported by providing students with latitude and decision-making opportunities (Skinner & Belmont, 1993) and, in particular, affording students opportunities to make cognitive choices as well as organizational and procedural choices regarding their work (Stefanou, Perencevich, DiCintio, & Tural., 2004). Research has shown that intrinsic motivation is enhanced when students are given a choice about what to work on and the amount of time they spend on each task (Zuckerman, Porac, Lathin, Smith, & Dal., 1978). Moreover, when teachers provide autonomy support to students, students demonstrate a preference for more difficult work (Harter, 1978) and strive for conceptual understanding (Grolnick & Ryan, 1989).

In an assessment context, providing students with opportunities to choose or self-initiate an action may be challenging, at least from a psychometric standpoint. For example, Wainer and Thissen (1994) argue that building examinee choice into a test results in unfair tests because the different forms of the tests that are built as a consequence of examinee choices may not be of equal difficulty and cannot be statistically equated, which then renders the scores comparing individuals not comparable. In addition, their review of studies that examined whether two choice questions were of equal difficulty on the College Board's Advanced Placement tests showed that "all examinees do not choose items that will show their proficiency to best advantage… examinee choice is not likely to yield credible estimates of θ_{Max}"[characterization of proficiency that would be obtained if the examinees choose the item that would give them the highest score] (Wainer & Thissen, 1994, p. 170).

Moreover, Powers, Fowles, Farnum, and Gerral. (1992) found that when examinees were given a choice of topics on a test of basic writing skills and asked to choose a topic based on their preference for the topic, preferences had little, if any, relationship to essay scores. In fact, they found a negative, though not highly significant relationship between the mean preference ratings and the mean scores awarded. When controlling for students' mean SAT verbal scores, the correlation between preference ratings and performance dropped and was not statistically significant. However, Powers and colleagues assert that "even if allowing a choice of topics necessitates sacrificing strict task equivalence, this may be a good trade off in terms of equity" (p. 13). Allowing examinees to choose the essay topic may help ensure that the examinee does not write on a topic that is outside his experience, minimizing the chances that the examinee's performance on the exam is inappropriate or difficult for the wrong reasons.

Collaboration

Student collaboration has been found to greatly increase student engagement (Johnson & Johnson, 1987; Slavin, 1990). Collaboration refers to students working with each other in pairs or small groups to share ideas, ask questions, and build on each other's ideas. Collaboration builds upon the cognitive theories of Piaget's social arbitrary knowledge where interactions with others are key to learning and Vygotsky's zone of proximal development (Slavin, 1980). In addition, student talk has been found to improve student learning (Yackel, Cobb, & Wood, 1991), and collaboration provides opportunities for students to engage in such talk. Collaboration can increase student motivation, where a group reward system relies on individuals attaining their goals through the success of the group, facilitating students' encouragement of each other's effort and learning (Slavin, 1990). Collaboration has also been found to foster norms that promote equity with attention to status issues in student groups (Boaler, 2008).

Student collaboration has also been found to increase student performance in testing settings. Skidmore and Aagaard (2004) studied the performance of undergraduate students seeking entry to a teacher education program at a Mid-South state university. The 141 participants were students in four sections of the same course with the same instructor. Skidmore and Aagaard analyzed differences in participant performance of five multiple choice exams. Students worked independently on the first and second exams. For the third exam, students were permitted to bring an 8 ½ by 11 inch "cheat sheet" of notes to the exam. For the fourth exam, students were allowed to discuss the exam questions in groups that were assigned heterogeneously, based on student performance on the first two exams. Students discussed the exam questions in the hallway, but did not take notes or write on their exams during these group discussions. They then returned to the classroom to take the exam independently. For the fifth exam, students were again allowed to bring a "cheat sheet" and

were also engaged in discussions with homogeneously assigned groups in the same fashion as with the heterogeneous groups.

The heterogeneously and homogeneously assigned group treatments led to the greatest score gains, with the heterogeneously designed groups resulting in the largest effect. Skidmore and Aagaard offer that collaboration serves as a scaffold for managing the mechanics of testing, such as considering all responses prior to selecting a response. Further, collaboration may improve test performance by fostering motivation and offering opportunities for social interaction and student talk.

Authenticity

Authenticity also matters to student engagement and performance. Authenticity refers to the extent to which the task requires students to solve real-world problems and has value beyond the school (Lombardi, 2007). Real-world problems reflect tasks that are encountered in real professions and everyday settings and are often complex and require sustained effort to solve. The criterion of value beyond school means that the creation of products and performances has personal, utilitarian, or social significance aside from assessing the knowledge and skills of the student (Newmann, Marks, & Gamoran, 1996). That is, the products created by students (e.g., the science lab or research report) have an audience or purpose beyond "getting the grade." Research by Newmann and his colleagues found that when students in elementary and middle school classrooms engage in authentic work, the quality of their academic performance increases (Newmann & Associates, 1996; Newmann, Bryk, & Nagaoka, 2001).

Authenticity may be fostered by creating tasks that promote realistic problem-solving processes (Smith, 1987) and bear significant resemblance to the activities conducted by real practitioners (Brown, Collins, & Duguid, 1989). Exposure to such activities, rather than to disjointed abstract concepts and skills, allows students to "tease out the way a mathematician or historian looks at the world and solves emergent problems" (Brown, Collins, & Duguid, 1989, p. 34). Through authentic tasks and activities, students do not simply learn science by memorizing facts or by following scientific procedures, but rather they engage in scientific discourse or the use of science in ways that a scientist would.

Other strategies for fostering authenticity in learning tasks include the use of problem-based learning instructional approach and the use of contextualized problems in math. Problem-based learning is defined as an "instructional method characterized by the use of 'real-world' problems as a context for students to learn critical thinking and problem-solving skills" (Duch, 1995, paragraph 1). Finkelstein and colleagues (2010) conducted a randomized controlled trial study to assess the impacts of a problem-based learning approach to teaching high school economics on student learning. Sixty-four teachers who volunteered for the study were randomly

assigned to the intervention or control condition (35 and 29 teachers, respectively). The intervention teachers received professional development on a 40-hour, problem-based economics course held over five days in the summer in order to implement the Problem Based Economics curriculum developed by the Buck Institute. The control teachers implemented a textbook-driven economics curriculum, attended their regular annual professional development activities during the school year, and continued their usual instructional practices. The researchers found that students whose teachers taught economics through the problem-based learning approach significantly outperformed their control group peers on the Test of Economic Literacy by an average of 2.6 test items that amounted to an effect size of .32. In addition, a statistically significant difference was found on student measures of problem-solving skills and application to real-world economic dilemmas that favored the intervention group over the control group (point estimate of .54, which corresponded to an effect size of .27).

As another example, Bottge (1999) conducted a quasi-experimental study that examined the effects of contextualized problem-solving instruction on middle school students' math performance. The study showed that low- and average-achieving students who received instruction with contextualized problems (i.e., video-based math problems with "embedded data design," wherein pertinent information needed to solve the problem are not explicitly stated or well formulated as they are in standard word problems) outperformed comparison students who received word problem instruction on the contextualized problem posttest as well as a transfer task that was administered 10 days after the posttests. The author concluded that the students who received instruction in contextualized problems were better skilled at noticing critical features of problems and, thus, were better able to generalize their mathematics problem-solving skills to novel problems.

As good teachers know, these ways of engaging students can trade off against each other. If students need to study something that is far from their experience, they can be engaged by the format of the tasks or the way they are drawn into the work or the opportunity they have to collaborate with others.

Engagement Considerations for Traditionally Underserved Students

Additional engagement considerations for students of historically underserved student groups include concerns of *stereotype threat*. Students of some subgroups often underachieve when faced with stereotype threat; that is, when students' attempt to disprove a negative stereotype of their subgroup results in performance below their actual ability (Steele, 2003), or when presented with abstract tasks irrelevant to their own lives. Steele and his colleagues have tested specific ways of framing and presenting tasks and tests that have demonstrated significantly higher performance from groups that experience stereotype threat (girls in math, African Americans on various kinds of tests, etc.) These approaches should be considered as classroom activities are developed and as test instructions are framed.

Interpretation of text may be of particular concern for students of different socio-economic and cultural backgrounds as well as English Language Learners. Student interpretation of test items is often mediated by socioeconomic and cultural factors, reducing the validity of the assessment. For example, Solano-Flores and Trumbull (2003) analyzed the different ways students interpreted 1996 fourth grade NAEP mathematics items. The "lunch money" item read as follows:

> Sam can purchase his lunch at school. Each day he wants to have juice that costs 50¢, a sandwich that costs 90¢, and fruit that costs 35¢. His mother has only $1.00 bills. What is the least number of $1.00 bills that his mother should give him so that he will have enough money to buy lunch for 5 days?

This item is intended to measure proficiency with addition, multiplication, and rounding. However, interviews of students' interpretation of the question of three student subgroups (high SES suburban white, low SES inner city African American, and low SES rural Native American) revealed great variation. Eighty-four percent of white students read the question as intended, whereas only 56% and 52%, respectively, of Native American and African American students read the sentence as intended. Solano-Flores and Trumbull also found that 10% and 18%, respectively, of the Native American and African American students interpreted the word only as restricting the number of dollars (i.e., "His mother has only one dollar"); however, this interpretation was not observed with white students in the study.

For English Language Learners, inclusion of unnecessary *linguistic complexity* can also provide inaccurate measures of student understanding. However, Performance

Assessments that use clear language embedded in realistic contexts have been found to provide ELL students greater opportunities to demonstrate understanding than traditional discrete-item assessments (Abedi, 2010). In addition, tasks can be linguistically modified to make the content more accessible. Linguistic modifications include reducing the number of words in the task by eliminating superfluous language, using active voice and excluding conditional clauses, and using more familiar words. These modifications can make the task easier to read while not altering the knowledge and skills being measured.

Scoring rubrics, while not tied explicitly to engagement, are of further consideration for students of traditionally underserved backgrounds. For example, Moschkovich (2008) found multiple meanings behind students' reasoning expressions when examining eighth grade English Language Learners' problem-solving discourse, highlighting the importance of evaluators' interpretation of students' mathematical reasoning. Rubrics can help ensure that scorers attend to students' meaning rather than being overly influenced by specific modes of expression.

We address concerns highlighted in this literature review for assessment design in the following section.

Assessment Tensions and Design Recommendations

Several tensions present challenges to assessment design. For example, efforts to personalize assessments are met with questions about standardization. Similarly student collaboration in the assessment context can raise concerns about measuring individual performance. Below, we list tensions in assessment design between concerns for engagement and concerns associated with the ways in which fairness has often been sought in traditional test design through standardization and decontextualization. We also present design recommendations that address these tensions in Table 1. Each design recommendation is also accompanied with questions for item writers' consideration. Following Table 1, we offer a bulleted list of design recommendations by each of the following on-demand performance task components: classroom activity, design of the task, and scoring of task.

Among the tensions that emerge as assessment designers seek to balance engagement with a range of other goals for assessment, are the following:

1. How can assessments offer relevant themes, when students have often had different experiences, interests, and exposure to background knowledge?

2. How can assessments be personalized without threatening key aspects of standardization?

3. How can assessments include opportunities for collaboration when the outcome is intended to describe individual performance?

4. How can assessments balance the demands of authentic intellectual work, which results in open-ended responses that must be interpreted, with the desire for rapid, low-cost, accurate scoring?

5. How can assessments balance the rich context needed for authentic performance tasks (i.e. a context that is sufficiently detailed to describe a realistic situation) with concerns for minimizing linguistic complexity?

6. How can assessments stimulate engagement and encourage students to understand and complete the task when test objectivity has traditionally been defined as precluding engagement with peers or adults?

7. How can assessments be scored in ways that allow diversity of responses without sacrificing reliability, accuracy, and ease of scoring?

These tensions are discussed further, with suggestions for resolution in Table 1. We note that there are ways to balance different pathways to engagement in tasks, such that task developers can aim to purposefully engage students on one or more levels, by:

- connecting to young people's interests and experiences,
- provoking their curiosity with lively stimuli,
- communicating through technologically current media, or
- drawing them into a dramatic scenario in which they take personal autonomy to complete a purposeful task.

For an example that illustrates how a performance task can be modified to increase engagement by enhancing relevance and studency autonomy, please refer to Appendix B.

Table 1. Design Recommendations to Increase Student Engagement

Tension	The Traditional Testing Approach	Recommended Approach (in Task and through Classroom Interaction Activity)	Questions for task developers to consider
1. Familiarity and relevance matter to student engagement and performance, yet students have had different experiences. Different contexts are more familiar and relevant to some students than others.	Design tasks for a "mainstream" student audience. Avoid contexts or examples that may be unfamiliar to subgroups of students, such as sports, hobbies, activities, or social/political issues.	Connect the task/topic/context to students' lived experiences, interests, or prior knowledge. The contexts should be **familiar** and **relevant** to many students at the age group being tested and should be sensitive to the experiences of non-dominant and traditionally underserved students of diverse backgrounds. Include a mix of items and tasks that draw on the experiences and cultural contexts of both dominant and non-dominant groups. As a group, performance tasks should represent a variety of diverse backgrounds. In ELA, for example, tasks might present students with texts discussing the Harlem Renaissance, Chinese Exclusion Act, Delano Grape Strike, and texts from a variety of authors (e.g., Chinua Achebe, Toni Morrison, Gabriel Garcia Marquez, Jhumpa Lahiri, and James Baldwin). Use sports, hobbies, and other contexts as interesting activities, but choose widely known sports or activities and avoid those that are known primarily to privileged groups. Identify the prior knowledge, familiarity, or experience that is expected, implied, assumed, and/or required of the task. Every item or task will have a context that is differentially known to different students by geography, interests, activities, etc. As much as possible, background knowledge associated with the context should be made explicit to the students and incorporated into the classroom activity or the task itself. It should also not require extensive amounts of background knowledge such that the explanations in the classroom activity and in the task itself are sufficient for students with no prior exposure to the context. Be sure to provide sufficient background knowledge about activities that are part of a task context for students to be able to perform the task. For example, do not assume that all students know how a particular sport is scored or played. Use classroom interaction activities to familiarize students with the context of the task in an engaging way that reduces anxiety about non-familiarity.	What prior knowledge, familiarity, or experience is expected, implied, assumed, and/or required of the task (e.g., how basketball games are scored; how the stock market works; or how sleds, surfboards, or boats operate)? Why would a student care about this topic? Consider a diversity of students in addressing this question from different backgrounds (such as, but not limited to, different socioeconomic, cultural, age, geographic, physical and/or mental wellness, sexual orientation, religious, and language backgrounds). If the context is largely known primarily to upper-income students (e.g., traveling to Europe on vacation or sailing in a regatta), its appropriateness for the purpose of this assessment should be questioned.

Tension	The Traditional Testing Approach	Recommended Approach (in Task and through Classroom Interaction Activity)	Questions for task developers to consider
2. Student autonomy increases engagement. Standardization has often been interpreted as precluding choice.	No choices are offered in an effort to preserve standardization and fairness.	To engage students, tasks, while standardized, should present opportunities for student decision-making. Tasks can promote autonomy by providing choices about how to approach the problem and what decisions to make. Offer opportunities for student autonomy that are age and developmentally appropriate. For example, elementary aged students may engage in a task to write a letter to a parent or family member to advocate and argue for something they desire and care about while a high school aged student may prepare a speech to student government (although not all schools have student government programs), principal, or teachers in their school to suggest changes in school policies. Where possible, allow choice of tasks and prompts as is common in IB, AP, and other countries' assessments. This allows students to choose tasks that connect with student interest, lived experiences, and/or background knowledge.	• How are students invited to enact autonomy and/or choice through the task? • How is the task structured to provide a range of possible solutions?
3. Student collaboration improves engagement. A focus on standardization and individual performance has typically eliminated opportunities for collaboration.	Proctor reads a script with minimal instructions when administering the exam.	Include collaboration with peers in the classroom activity that is standardized in terms of the activities it includes. Teachers and students have no knowledge of the specific details of the performance task to follow, alleviating concerns of so-called "cheating" or measuring students' individual performance. Classroom activity should be facilitated by students' regular classroom teacher to increase familiarity, reduce anxiety, and enhance students' sense of belonging.	• Are students afforded opportunities to engage with peers and ask questions of their teacher such as through a classroom activity that builds background knowledge? • Does the classroom activity provide a standard set of stimuli with teacher facilitation information/guidelines that are clear and common, but not a lockstep script? • How will the teacher facilitator's role contribute to increasing engagement? • How will the classroom activity include a variety of modalities (e.g., partner/group work, listening to audio, watching video, etc.)? • What modes can be used to increase engagement (e.g., partner or group work, brief debate or student presentation, etc.)?

Tension	The Traditional Testing Approach	Recommended Approach (in Task and through Classroom Interaction Activity)	Questions for task developers to consider
4. Students invest in authentic and purposeful intellectual work that requires reasoning. Such work is more complex and less easily scored by machine.	Problems are more procedural in nature and do not require students to apply their knowledge to novel problems. No consideration is given to the student's role in tasks; items are posed in the third person (e.g., Mary went to the store). There is an over reliance on items that are multiple-choice, discrete, and computer scorable.	Invite students to engage with **authentic intellectual** tasks that involve construction of knowledge, disciplined inquiry, and value beyond the classroom. Emphasize real-world connections. Provide opportunities for students to demonstrate original applications of knowledge and skills used in the real world. Incorporate a variety of information sources and stimuli (e.g., original text, realistic data figures, or tables and charts) that are representative of artifacts used in the world beyond the classroom. Use simulations and plausible scenarios to require students to communicate their knowledge and skills to an audience beyond the teacher, classroom, and school.	• How does the task emphasize real-world connections? • To what extent does the task address the central ideas and modes of inquiry in the discipline? • How is the task structured so that students demonstrate novel applications of knowledge and skills? • To what degree does the task involve students in manipulating information and ideas to arrive at conclusions that solve an open-ended problem? • How is the task structured to provide a range of acceptable right answers that can be analytically scored? • What is the student's role in the task? Is this role purposeful, meaningful, appropriate, and accessible? • Is the rubric built to value the construction of knowledge and creative approaches?
5. Tasks with rich contexts invite engagement. Linguistically dense items may disadvantage English Language Learners or struggling readers.	Create unnecessary text complexity (that is not construct-relevant). Alternatively, reduce text and reduce context in an effort to offer manageable text. (However, this can result in decontextualized, inauthentic items not relevant to students' lives.)	Assessment items should present **manageable text** (both for English Language Learners and for assessments where reading comprehension is not being measured) without stripping the task of its **rich context.** Use clear text, use familiar words, and remove construct-irrelevant complex language without removing construct-relevant language. Use the classroom activity to be sure that key concepts and vocabulary have been introduced and to familiarize students with the context. Present stimuli in more than one mode and format. These may include graphic, tabular, audio-visual, and pictorial modes as well as text-based modes (though not all at once in order to avoid overwhelming students). Allow students ready access to glossaries to decode text. Invite students to demonstrate their understanding through multiple modes (e.g., using graphing tools or creation of tables, equations, and figures to communicate solution strategies in mathematics).	• Does the performance task offer a variety of information sources (e.g., text, visuals, figures, tables, charts, etc.)? • Should the text be linguistically modified to be readily accessible? • What additional formats may be used to present stimuli that may engage more students? • In what multiple ways are students invited to demonstrate their understanding? • Do student response capabilities also allow for multiple solution pathways and modes of representation? • Is text comprehension and understanding of the context/ scenario supported through the classroom activity?

Tension	The Traditional Testing Approach	Recommended Approach (in Task and through Classroom Interaction Activity)	Questions for task developers to consider
6. Clear and encouraging expectations engage students in the intended task. Testing instructions must be minimal in order to be fair.	Minimal directions are provided to students. Tests assume that students understand the format of the test and the types of responses expected.	Task expectations should be clear, and students encouraged to persevere with assessments. For example, the audience and format expected in responses should be clear (e.g., "write a letter to your school principal with your recommendation"). Communicate that effort and persistence make a difference. Structure assessments in such a way that students feel their success is valued and expected. Minimize messages regarding competitive performance. Minimize stress by providing adequate time, access to resources, and focus on concepts and skills students perceive to be important.	• In what ways do the classroom activity and performance task communicate test-taking strategies and high expectations for students' success? • Are expectations for student responses (including conventions) clear to the student?
7. Performance-based tasks are more engaging for students, but less reliable and valid measures. Multiple choice items are not as engaging for students, but are more reliable and valid measures.	Multiple choice items are scored by the computer.	Scoring rubrics should be analytic, provide specificity, and should articulate the diversity and range of appropriate student responses to ensure fairness. Performance tasks should be hand scored, with capability for students to respond in some languages other than English. Specify weights for each element in the rubric. Include a diversity of possible student responses with opportunities to earn partial credit.	• Is the rubric aligned with the intent of the task? • Does the rubric anticipate the full range of possible student responses? • Is the rubric clear enough for scorers to fully understand how to award points? • Is the rubric specific enough to increase inter-rater reliability and norming of scores? • Is the rubric susceptible to bias (intentional or implicit)? • Does the rubric include specific weights for each element, including opportunities for partial credit?

Design Recommendations by Task Component

Student engagement shows positive outcomes for achievement, standardized test scores, and connection with the academic discipline. Positive effects of engagement range from primary school through college. Although prior achievement is the best predictor of later achievement for high-performing students, engagement is the best predictor of achievement for low-performing students. Design recommendations that improve student engagement are listed below by Task Component.

Classroom Activity

Performance assessments typically involve a classroom activity component where students are offered opportunities to build background knowledge, work collaboratively with peers, and ask clarifying questions. Performance tasks in Smarter Balanced Assessment Consortium assessments include a classroom activity prior to exam administration. The classroom activity may be up to 30 minutes in length. It is intended to ensure students' understanding of the context of the performance task as to not disadvantage those unfamiliar with the context of the performance task (Smarter Balanced Assessment Consortium, 2014). Students benefit from the classroom activity (whether for in-class assessments or for Smarter Balanced assessments) because of the opportunities to collaborate and build the background knowledge required to successfully complete the task.

- The classroom activity should ***activate students' prior knowledge and build background knowledge*** needed to comprehend and complete the task.

- The classroom activity can enhance engagement and access by creating ***familiarity*** with the context and introducing key concepts and vocabulary so that students are ready to engage the substance of the task.

- The classroom activity may provide an opportunity to offer ***personalization, student choice, and autonomy*** through the type of activities it uses. For example, students may choose which side of an argument they wish to represent in a brief class debate. Students might vote on their preferences or share their opinions about a topic.

- Inclusion of video, audio, and use of manipulatives during the classroom activity may further invite engagement through the addi-

tion of ***multiple communication modalities***. For example, students might take their own heartbeat in the classroom activity as presented in the sample performance task in Appendix B.

- The classroom activity can help create a supportive learning environment focused on all students' success, through ***active and collaborative learning*** (such as hands-on learning in small groups where students work together to make sense of concepts and solve problems).

- The classroom activity should incorporate ***visualization*** of important ideas and concepts. The use of visual models accompanied by discussion to support an understanding of the context is likely to increase engagement with the performance task.

- The classroom activity provides an opportunity to ***decrease stress*** by creating a climate of care, focusing on big ideas rather than discrete bits of knowledge, and communicating directions and expectations for the subsequent performance task. Improving socio-emotional context (a sense of belonging, social support, good relationships with teachers) through the classroom activity may engage more students in the subsequent performance task. Stress can also be minimized by providing adequate time, access to resources, and focus on concepts and skills students perceive to be important.

- Assessment events should be introduced and structured in such a way that students feel their ***success is valued and expected.*** Messages regarding competitive performance should be minimized. Teacher facilitators can be prompted to convey expectations to students such as, "On this assessment, there are many different ways to accomplish the task. This task has several parts and there are several 'right' ways to solve the problem. Look over the entire task to understand the overarching goal. Use what you know; think creatively. If you get stuck, review the questions and information provided to see what you have already answered. If possible, skip to the following prompt and return to the one that stumped you later." Messages can also be constructed to ***minimize stereotype threat,*** such as affirmations about the fact that, while difficult, the task is one at which students are expected to be able to succeed.

Design of the Tasks

- Items and performance tasks should engage students in ***authentic, hands-on work*** that involves inquiry, use of authentic disciplinary thinking and processes, and the construction of new knowledge. Tasks should involve inquiry-based learning (questioning, investigating, drawing conclusions, reasoning from evidence), higher-order reasoning, and "sense-making."

- Tasks should be ***relevant*** to students' own lives and experiences and have value within and beyond school. The work should be anchored in ***real-world contexts*** and tasks placed in academic/theoretical contexts should be work worth doing – work that the student feels is purposeful, interesting, and challenging, yet achievable.

- The focus of assessments should be on ***important, core ideas*** rather than discrete, abstract bits of knowledge and/or skill.

- Whenever possible, assessment events should provide opportunities for ***autonomy and choice***. Students should be able to take a role in solving a meaningful problem in a fashion that allows them to choose how they will approach the problem and that allows them to make decisions in the process.

- Where possible, assessments should provide ***auto-feedback*** (i.e., clues to the causes of difficulties as well as opportunities for attacking the task in a new, more informed way).

- Where possible, performance tasks should ***integrate subject matter***, demonstrating how concepts and skills relate across disciplines and how students can learn in context.

- Assessment item types should be ***familiar*** and ***relevant*** to students. Contexts should be familiar and relevant to many students at the age group being tested and should be sensitive to the experiences of non-dominant and traditionally underserved students of diverse backgrounds. Sports, hobbies, and other activities may be used as contexts, but developers should choose widely known sports or activities and avoid those that are known primarily to privileged groups. Be sure to provide sufficient background knowledge about activities that are part of a task context for students to perform the task (i.e., don't assume that students know how a game is scored or played).

- Assessment items should present ***manageable text*** (both for English Language Learners and for struggling readers in assessments where reading comprehension is not construct-relevant) without stripping the task of its ***rich context.*** Provide clear text and use familiar words without removing construct-relevant language.Present stimuli in ***multiple modes and formats*** (use graphs, charts, figures, etc.). Where possible, invite students to demonstrate their understanding through multiple modes (e.g., using graphing tools or creation of tables, equations, and figures to communicate solution strategies in mathematics).

- Assessments can offer ***choices of tasks*** for students. Student choice of task completion from a menu of tasks is successfully used in other exams such as the A-Level British Council and International Baccalaureate assessments.

- Assessments might include as an option that students be surveyed about their interest (from a fixed number of choices) prior to presenting the performance task. A question may ask students, "From the following topics, in which do you have the most interest? (a) music, (b) sports, (c) art, or (d) games." The subsequent mathematics performance task may be set using the context chosen by the student. If the following performance task aims to capture student understanding of linear functions and lines of best fit and a student chooses music as their topic of interest (from the limited four choices), data can represent record sales trends over time and ask the student to compare and contrast CD sales and iTunes sales patterns. Alternatively, if a student selects sports, the performance task may present Oakland A's and Los Angeles Angels ticket sales over time, ask for the lines of best fit, and include other questions related to prediction and comparison in order to capture student thinking and facility with using linear functions to make real-world predictions and decisions. These performance tasks can measure the same mathematics content but through the context in which the student may be most interested.

Scoring of the Tasks

- Scoring rubrics should be ***analytically constructed*** and should include specific weights for each element, emphasizing meaning and central understandings and minimizing aspects of communica-

tion that may be associated with culture and language background where these are not the focus of the construct.

- Rubrics or scoring criteria should anticipate the ***diversity and range*** of possible student responses to ensure fairness in evaluation, particularly for students of traditionally underserved backgrounds. Scoring criteria should include the range of plausible anticipated student answers to alert scorers to award credit and partial credit when they see a novel or untraditional way of answering a question that they themselves might not anticipate as a "correct" or "partially correct" response.

- Performance tasks should maintain ***open-ended responses*** that are, at least initially, human scored, in order to assess complex modeling, reasoning, and communication skills.

- Performance tasks should, in some cases, allow students to respond in ***languages other than English.***

Conclusion

Assessments for college and career readiness intend to measure 21st century skills and higher-order thinking. By definition a performance task "is an item type designed to provide students with an opportunity to demonstrate their ability to apply their knowledge and higher-order thinking skills to explore and analyze a complex, real-world scenario" (Smarter Balanced Assessment Consortium, 2014, p. 3). The recommendations provided in this document offer a framework for considerations of engagement for all students of diverse backgrounds (such as, but not limited to, different socioeconomic, cultural, age, geographic, sexual orientation, religious, and language backgrounds).

Engagement considerations are of ongoing concern and importance. We see this as an important first step and look forward to continued dialogue and actions to continually address concerns of student engagement in assessments.

References

Abedi J. (2010). Performance Assessment for English Language Learners. Stanford, CA: Stanford University. Stanford Center for Opportunity Policy in Education.

Arbuthnot, K. (2011). *Filling in the Blanks: Understanding Standardized Testing and the Black-White Achievement Gap*. Charlotte, NC: Information Age Publishing.

Boaler, J. (2008). Promoting 'relational equity' and high mathematics achievement through an innovative mixed-ability approach. *British Educational Research Journal*, *34*(2), 167-194. doi: 10.1080/01411920701532145

Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *Elementary School Journal, 108,* 115-130.

Bottge, B. A. (1999). Effects of contextualized math instruction on problem solving of average and below-average achieving students. *The Journal of Special Education, 33*(2), 81-92.

Brophy, J. (1986). *Socializing student motivation to learn* (Research Series No. 169). East Lansing, MI: The Institute for Research on Teaching. Retrieved on September 25, 2014 from http://education.msu.edu/irt/PDFs/Research-Series/rs169.pdf

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Bryson, C., & Hand, L. (2007). The role of engagement in inspiring teaching and learning. *Innovations in Education and Teaching International 44,* 349–62.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. G. Gunnar & L. A. Sroufe (Eds.), *Self processes and development. The Minnesota symposia on child psychology, Vol. 23*, (pp. 43-77). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Darling-Hammond, L., Barron, B., Pearson, D. P., et al. (2008). P*owerful Learning: What We*

Know About Teaching for Understanding, pp. 74-76. San Francisco: Jossey-Bass.

Deci, E. L., & Ryan, R. M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology, 53*, 1024-1037.

Duch, B. (1995). What is problem-based learning? *About teaching 47.* A newsletter of the Center for Teaching Effectiveness. Retrieved on November 18, 2014, from http://www.udel.edu/pbl/cte/jan95-what.html

Dotterer, A. M., & Lowe, K. (2011). Classroom context, school engagement, and academic achievement in early adolescence. *Journal of Youth and Adolescence, 40*, 1649-1660.

Finkelstein, N., Hanson, T., Huang, C.-W., Hirschman, B., and Huang, M. (2010). *Effects of Problem Based Economics on high school economics instruction.* (NCEE 2010-4002). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59–109.

Grolnick, W., S., & Ryan, R. M. (1989). Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology, 81*, 143-154.

Harter, S. (1978). Pleasure derived from optimal challenge and the effects of extrinsic rewards on children's difficulty level choices. *Child Development, 49*, 788-799.

Johnson, D. W., & Johnson, R. T. (1987). *Learning Together and Alone: Cooperative, Competitive, and Individualistic Learning* (2nd ed.). Englewood Cliffs, N.J.: Prentice-Hall.

Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status* (pp, 383-434). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lombardi, M. M. (2007). *Authentic learning for the 21st century: An overview*. Educause Learning Initiative. Retrieved on October 15, 2014, from https://net.educause.edu/ir/library/pdf/eli3009.pdf

Meier, T. (2008). *Black Communications and Learning to Read: Building on Children's Linguistic and Cultural Strengths*. New York: Lawrence Erlbaum.

Moore, M. (1989). Three types of interaction. *American Journal of Distance Education* 2: 1–6.

Moschkovich, J. N. (2008). "I went by twos, he went by one": Multiple interpretations of inscriptions as resources for mathematical discussions. *The Journal of the Learning Sciences*, *17*(4), 551-587.

Newmann, F. M., & Associates (1996). *Authentic achievement: Restructuring schools for intellectual quality*. San Francisco: Jossey-Bass.

Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence?* Chicago: Consortium on Chicago School Research.

Newmann, F. M., Marks, H. M., & Gamoran, A. (1996). Authentic pedagogy and student performance. *American Journal of Education, 104*(4), 280-312.

Powers, D. E., Fowles, M. E., Farnum, M., & Gerritz, K. (1992). *Giving a choice of topics on a test of basic writing skills: Does it make any difference?* (RR. 92-19). Princeton, NJ: Educational Testing Service.

Skidmore, R. L., & Aagaard, L. (2004). The Relationship between Testing Condition and Student Test Scores. *Journal of Instructional Psychology*, *31*(4), 304-313.

Skinner, E. A., & Belmont, M. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85*, 571-581.

Slavin, R. (1980). Cooperative learning. *Review of Educational Research*, *50*(2), 315-342.

Slavin, R. E. (1990). *Cooperative Learning: Theory, Research, and Practice*. Englewood Cliffs, NJ: Prentice Hall.

Smarter Balanced Assessment Consortium. (2014). Mathematics Performance Task

Specifications. Retrieved from: http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/05/Mathematics-Performance-Task-Specifications.pdf

Smith, P. E. (1987). Simulating the classroom with media and computers. *Simulation and Games, 18*(3), 395-413.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new

research and practice paradigms in the testing of English-language learners. *Educational Researcher*, *32*(2), 3-13.

Steele, C. (2003). Stereotype threat and African-American student achievement. In T. Perry, C.

Steele & A. G. Hilliard (Eds.), *Young, gifted, and black: Promoting high achievement among African-American students*. Boston: Beacon Press, pp. 109-130.

Stefanou, C. R., Perencevich, K. C., DiCintio, M., & Turner, J. C. (2004). Supporting autonomy in the classroom: Ways teachers encourage student decision making and ownership. *Educational Psychologist, 39*(2), 97-110.

Wainer, H., & Thissen, D. (1994). On examinee choice in educational testing. *Review of Educational Research, 64*(1), 159-195.

Walkington, C. A. (2013). Using adaptive learning technologies to personalize instruction to

student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology*, *105*(4), 932-945.

Walkington, C., & Sherman, M. (2012). Using adaptive learning technologies to personalize instruction: The impact of interest-based scenarios on performance in algebra. In *Proceedings of the 10th International Conference of the Learning Sciences. Sydney,*

Yackel, E., Cobb, P., & Wood, T. L. (1991). Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education*, *22*, 390-408. doi: 10.2307/749187

Zuckerman, M., Porac, J., Lathin, D., Smith, R., & Deci, E. L. (1978). On the importance of self-determination for intrinsically motivated behavior. *Personality and Social Psychology Bulletin, 4*, 443-446.

Appendix A:
Summaries and Resources of Student Engagement Research

Factor	Brief Summary of Research	Sources
Work that is relevant to the lives of students	Student engagement increased from 68% to 91% when students did "…activities that encourage students to draw on their previous knowledge and experiences, engage in critical thinking, and apply what they learn to their own lives." Type of instruction (lecture versus authentic work) explained 15% of the variance in student engagement scores.	Cawthon, S. W., Dawson, K., & Ihorn, S. (2011). Activating student engagement through drama-based instruction. *Journal for Learning through the Arts, 7* (1), 1-29.
Authentic intellectual work	Newman and his colleagues (Newman & Associates, 1996; Newman, Bryk, & Nagaoka, 2001) found that authentic intellectual work involved construction of knowledge, disciplined inquiry, and value beyond the classroom. Authentic intellectual work can be shown in three places: instruction, assessment, and student work. (To what degree does instruction involve students in manipulating information and ideas to arrive at conclusions that produce new meanings for the student? To what extent does instruction address the central ideas of the discipline? To what degree are students involved in exchanges with the teacher and/or their peers about disciplinary ideas? To what extent do the assessment and instruction provide connections to the larger context in which students live?) When students in elementary and middle school classrooms engaged in authentic work, quality of academic performance increased, regardless of SES, ethnicity, gender, or prior achievement level.	Newman, F. M., & Associates (1996). *Authentic achievement: Restructuring schools for intellectual quality.* San Francisco: Jossey-Bass. Newman, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence?* Chicago: Consortium on Chicago School Research.
Student self-assessment	National Survey of Student Engagement (NSSE), an annual survey of student engagement at the college level, was developed to measure student engagement so that colleges and universities could self-evaluate and improve the quality of instruction in order to increase student engagement (Kuh, 2001, 2003, 2004). NSEE assesses students' involvement with activities and conditions likely to generate learning (Coates, 2006). It also looks at "the policies and practices that institutions use to induce students to take part in these activities" (Kuh 2003). Using longitudinal and cross-sectional data, NSEE found that level of engagement was a good predictor of undergraduate GPA.	Kuh, G. D. (2001). Assessing what really matters to student learning: Inside the national survey of student engagement. *Change, 33*(3), 10–17. 66. Kuh, G. D. (2003). What we're learning about student engagement from NSSE. *Change, 35*(2), 24–31. Kuh, G. D. (2004). *The national survey of student engagement: Conceptual framework and overview of psychometric properties*. Bloomington: Indiana University Center for Postsecondary Research and Planning.

Factor	Brief Summary of Research	Sources
Social-emotional context of the classroom	Research has shown that the social, instructional, and organizational climate of schools influences both students' engagement and their academic achievement (e.g., Eccles et al. 1998; Patrick et al. 2007; Ryan and Patrick 2001). For example, children who feel a sense of belonging and social support are more likely to be engaged and participate in school (Deci and Ryan 1985; Wentzel 1997). Ladd and Burgess (2001) found that, when teacher-child conflict was greater, students were less engaged in the classroom, were less likely to enjoy school, and were at increased risk for poor academic performance. Further, Baker (2006) found that teacher-child conflict was associated with lower report card grades and standardized test scores. Whereas, classrooms rated as having a positive climate were associated with children being more engaged in classroom activities and higher achievement.	Baker, J. A. (2006). Contributions of teacher–child relationships to positive adjustment during elementary school. *Journal of School Psychology, 44,* 211–229. Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum. Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3: Social, Emotional, and personality development* (5th ed., pp. 1017–1095). Hoboken, NJ: Wiley. Ladd, G. W., & Burgess, K. B. (2001). Do relational risks and protective factors moderate the linkages between childhood aggression and early psychological and school adjustment? *Child Development, 72,* 1579–1601. Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology, 99,* 83–98. Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal, 38,* 437–460. Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology, 89,* 411–419.
Integration of subject-matter content	Connecting language arts, math, science, and arts through integrated instructional units focused on real-world problems increased student engagement and achievement. In authentic problems, students were presented with a scenario in their first session. They were required to identify issues, research the principles underlying the issues, and learn material in context; scenarios were intended to motivate the students to engage in learning and understand the issues that underlie the problem.	Racknor, W., & Drake, S. M., (2011, Fall). *Curriculum integration: One school district's journey.* Toronto, Canada: Education Canada, Canadian Education Association.

Factor	Brief Summary of Research	Sources
Active learning, collaboration, and instruction focused on students' success	Study of trends in NSSE over four years for freshmen and seniors suggest that high engagement scores are associated with more focus on active and collaborative learning experiences, better student-faculty interactions, and a supportive learning environment. Supportive environmental factors include focus on student learning, helping students set personal goals, and giving students an opportunity to close the gap between themselves and those who are more advanced.	McCormick, A. C., Kinzie, J., & Korkmaz, A. (2011, April). *Understanding evidence-based improvement in higher education: The case of student engagement.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
Visualizing	Students' ability to develop and use visual models of mathematics and science concepts improves engagement and performance.	Cifuentes, L. (2004). Visualization for middle school students' engagement in science learning. *Journal of Computers in Mathematics and Science Teaching, 23,* 109-137.
Decreasing stress	High-achieving students in constant high-stress contexts have decreased motivation and engagement and may resort to illegal drugs to manage the demands; decreases in stress increase motivation. Students' engagement increases when they believe that the educators care about their success. Strategies for decreasing stress include: block schedules, creating a climate of care, and focusing on understanding of the big ideas rather than many discrete bits of knowledge.	Pope, D. (2010). Beyond doing school: From "stressed out: to "engaged in learning." Toronto, Canada: Canadian Education Association.
Inquiry-based learning	Introducing enquiry-based learning in a first-year seminar course at a university level significantly affected the learning behaviors of students, led to greater motivation to succeed, and to enhanced reasoning and processing skills that were transferred to other courses throughout undergraduate experience; all students benefited from EBL in terms of achievement, engagement, and ability to access and use resources to support learning (Murray & Summerlee, 2007; Summerlee & Murray, 2010). Studies of inquiry-based classrooms in science showed that students developed deeper conceptual understanding as well as increased engagement, interest, and positive attitudes toward science (Chang & Mao, 1999; Schwartz, Lederman, & Crawford, 2004). However, focus on teaching to standardized tests tended to undermine focus on student learning, trivialize the targets of learning, and discourage inquiry teaching. Studies of classroom assessments found that assessments in low-income schools mirrored the types of test questions found on standardized tests and focused on low-level skills while assessments in middle- to upper-income schools focused on reasoning and inquiry (Madaus, 1999; Madaus, et al., 1992).	Chang, C., & Mao, S. (1999). Comparison of Taiwan science students' outcomes with inquiry-group versus traditional instruction. *The Journal of Educational Research, 92*, 340–346. Madaus, G. F. (1999). The influence of testing on the curriculum. In M. J. Early & K. J. Rehage (Eds.), *Issues in curriculum: A selection of chapters from past NSSE yearbooks.* Chicago: The University of Chicago Press. Madaus, G.F., West, M. M., Harmon, M.C., Lomax, R. G., & Viator, K. A. (1992). *The influence of testing on teaching math and science in grades 4-12.* Chestnut Hill, MA: Boston College; Funded by National Science Foundation and Center for the Study of Testing, Evaluation, and Educational Policy. Murray, J., & Summerlee, A. J. S. (2007). The impact of problem-based learning in an interdisciplinary first-year program on student learning behaviour. *Canadian Journal of Higher Education 37*(3), 87-107. Schwartz, R., Lederman, N., & Crawford, B. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education, 88,* 610–645. Summerlee, A. & Murray, J. (2010). Impact of inquiry-based learning on academic performance and student engagement. *Canadian Journal of Higher Education, 40*, 78-94.

Factor	Brief Summary of Research	Sources
Assessment as Learning: Student engagement during testing	Students' engagement during testing increased when given an opportunity to use a "cheat sheet" or to consult with peers to clarify problems.	Skidmore, R. L., & Aagaard, L. (2004). The relationship between testing condition and student test scores. *Journal of Instructional Psychology, 31*, 304-313.
Assessment tasks that allow higher-order reasoning and "sense-making"	Tasks that are well designed and support meaningful conjectures contribute to students' motivations to participate and cultivate positive dispositions toward mathematics.	Mueller, M., & Maher, C. A. (2009). Learning to reason in an informal math after-school program. *Mathematics Education Research Journal, 21*(3), 7–35. Mueller, M., Yankelewitz, D., & Maher, C. (2010). Promoting student reasoning through careful task design: A comparison of three studies. *International Journal for Studies in Mathematics Education, 3*(1), 135–156. Yankelewitz, D. (2009). *The development of mathematical reasoning in elementary school students' exploration of fraction ideas.* Unpublished doctoral dissertation, Rutgers, The State University of New Jersey, New Brunswick. Yankelewitz, D., Mueller, M., & Maher, C. (2010). Tasks that elicit reasoning: A dual analysis. *Journal of Mathematical Behavior, 29*, 76–85.
Peer learning (learning together) and feedback	Boud, Cohen, and Sampson (1999) define peer learning as "the use of teaching and learning strategies in which students learn with and from each other without immediate intervention from the teacher" (p. 413). In their study, Chan and Leijten (2012) found that students involved with peer learning were deeply engrossed in examining other student's approaches to the welding tasks and then evaluating the process just undertaken: *"The feedback strategies proposed are not difficult to introduce to learners and teachers but lead to improved student engagement, improved student meta-cognition and enhanced skill practice and learning." (p. 23)*	Boud, D, Cohen, R. and Sampson, J (1999) Peer learning and assessment, *Assessment and Evaluation in Higher Education, 24*(4), 413-126. Chan, S., & Leijten, F. (2012). Using feedback strategies to improve peer learning in welding. *International Journal of Training Research, 10* (1), 23-29.
Hands-on learning in small groups	Hands-on learning experiences in small groups – even with worksheets and textbooks – with high-quality tasks increase engagement while lectures and whole class instruction decrease student engagement (Cooper & Speece, 1990; Greenwood, 1996, 1991; Greenwood, Delquadri & Hall, 1989). The best-fitting statistical model shows engagement as a mediating variable between instruction and test scores.	Cooper, D.H., & Speece, D. L. (1990). Instructional correlates of students' academic responses: Comparisons between at risk and control students. *Early Education and Development, 3,* 270-300. Greenwood, C. S. (1996). The case for performance-based instructional models. *School Psychology Quarterly, 11*, 283-296. Greenwood, C. S. (1991). A longitudinal analysis of time to learn, engagement, and academic achievement in urban versus suburban schools. *Exceptional Children, 57,* 521-535. Greenwood, C. R., Delquadri, J., & Hall, R. V. (1989). Longitudinal effects of peer tutoring. *Journal of Educational Psychology, 81*, 371-383.

Factor	Brief Summary of Research	Sources
Timely, usable feedback	If our interactions with students are to be pedagogically effective, students must engage with them. This is particularly the case with formative assessment feedback, which relies for its effectiveness on being applied at some point in the future (Handley & Williams, 2011). Three factors influence students' engagement with feedback. First, students must be able to make sense of the feedback. If it is written in disciplinary language that is not familiar to students or uses terms that differ across disciplines, understanding is hampered. Second, if feedback has no bearing on what students do next, they are unlikely to attend to it. Finally, students cannot effectively use feedback if they don't understand the criteria (expectations) for the work. Feedback should help students close the gap between current performance and the desired performance level. Two effective ways to give feedback so that students engage with it are to give feedback on drafts and to give students exemplary work so that they understand what the ultimate goals of instruction area. For feedback to be effective, it has to be used in a structured and meaningful manner. Hattie and Timperley (2007) advocate the use of three forms of feedback: • Feed up – which is to answer the question of whether the learning objectives are being met • Feedback – providing the learner with the actual performance level on the learning activity • Feed forward – to the learner on what needs to be done to improve learning or move to the next objective The feedback cycle requires all three forms of feedback to be present for effective feedback to be completed.	Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessments. *Phi Delta Kappan, 92* (1), 81-90. Black, P., & Wiliam, D. (2009). Developing a theory of formative assessment. *Educational Assessment, Evaluation, and Accountability, 21*, 5-31. Boud, D. 1995. *Enhancing learning through self-assessment.* London: Routledge. Handley, K., & Williams, L. (2011). From copying to learning: Using exemplars to engage students with assessment criteria and feedback. *Assessment & Evaluation in Higher Education, 36*, 95-108. Hattie, J., and Timperley, H (2007) The power of feedback, *Review of Educational Research, 77*(1), 81-112. Sadler, D. 1989. Formative assessment and the design of instructional systems. *Instructional Science 18,* 119–44.

Factor	Brief Summary of Research	Sources
Content that influences Differential Item Functioning	In her chapter on differential item functioning (DIF), Arbuthnot (2011) reviewed research on math and verbal test DIF and ability. Among studies that examine the items that showed performance differences between African American and white students, one study found that particular item content genres favored African American students over white students. In math, purely quantitative math problems and those involving symbols favored African American students. While on verbal subject tests, reading items that featured humanities, human relations, and African American history topics favored African American students. In the same chapter, Arbuthnot attempted to answer the question of *why* one group of students performs differently than another. Arbuthnot examined the explanations that were not typically considered. The first was cultural differences. The cultural familiarity hypothesis tests the assumption that culture influences the performance on assessments. Researchers found that African American students performed differentially worse than white students on items involving words typically identified as high-frequency. Arbuthnot suggested that this may be explained by cultural differences because not all students are exposed to the same high-frequency words; cultural differences may lead students to be exposed to a different set of words. The second explanation Arbuthnot suggested was the degree to which a test item is engaging and interesting to a student or groups of students. Specifically, the more interested a particular group of students is in the topic of the test item, the more likely that item will differentially favor that particular group. Arbuthnot examined student perceptions and behaviors within the testing environment. "Test-wiseness" refers to a student's ability to use their test knowledge to their advantage – for example, knowing test-taking strategies to boost one's test score (e.g., most SAT preparation courses give student insight to the structure of the test and how to eliminate wrong answers). She found that white students were more likely than African American students to have this test-wiseness. In particular, African American students were more likely to skip questions on the SAT, where white students would guess because there is a penalty of skipped questions and no penalty for guessing.	Arbuthnot, K. (2011). *Filling in the Blanks: Understanding Standardized Testing and the Black-White Achievement Gap,* pp. 50-58. Charlotte, NC: Information Age Publishing.

Factor	Brief Summary of Research	Sources
Text complexity for ELL's in math	Martiniello (2008) explored the literature on text and lexicon complexity to understand DIF between ELL student and non-ELL student performance. The literature revealed the importance of not having unnecessarily complex words or sentence structure in math word problems to give ELL student test-takers proper accessibility. She chose to analyze six items that had a DIF between ELL and non-ELL students. Part of the analysis used a think-aloud protocol where ELL Spanish-speaking students deconstruct their process of reading and answering each item. The article included how students interpreted items differently from how they were intended to be read and how this affected their performance on the item.	Martiniello, M. (2008). Language and the Performance of English-Language Learners in Math Word Problems. In *Harvard Educational Review,* 78:2, 2008.
Cultural Validity	Solano-Flores (2011) explained that the way "students interpret test items and respond to them are mediated by cultural factors" and may not relate to the knowledge or skills being assessed. Solano-Flores suggested that in order to be valid, student scores should only reflect the knowledge being assessed and not other factors. This chapter described four aspects of cultural validity that help examine testing practices from a cultural perspective: theoretical foundations, population sampling, item views, and test review. *Solano-Flores and Trumbull (2003) analyzed the different ways students interpreted 1996 4th grade NAEP mathematics items. The "lunch money" item read as follows:* *Sam can purchase his lunch at school. Each day he wants to have juice that costs 50¢, a sandwich that costs 90¢, and fruit that costs 35¢. His mother has only $1.00 bills. What is the least number of $1.00 bills that his mother should give him so that he will have enough money to buy lunch for 5 days? (p. 6)* This item was intended to measure proficiency with addition, multiplication, and rounding. However, interviews of students' interpretation of the question of three student subgroups (high-SES suburban white; low–SES, inner city African American; and low-SES, rural Native American) revealed great variation. 84% of white students read the question as intended, whereas only 56% and 52%, respectively, of Native American and African American students read the sentence as intended. Solano-Flores and Trumbull also found that 10% and 18%, respectively, of the Native American and African American students interpreted the word *only* as restricting the number of dollars ("His mother has only one dollar"); however, this interpretation was not observed by white students in the study.	Solano-Flores, G. The Cultural Validity of Assessment Practices. In Basterra, M. del R., Trumbull, E. & Solano-Flores, G. (Eds.). (2011). *Cultural Validity in Assessment: Addressing Linguistic and Cultural Diversity,* pp.3-21. New York: Routledge. Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher, 32*(2), 3-13.

Factor	Brief Summary of Research	Sources
Equity and themes	The content of the themes in an assessment item can make a difference in a student's performance. When a student is more engaged in a reading passage, he or she is more likely to comprehend it than when reading a passage of similar complexity that he or she finds less engaging. Not only is comprehension easier, but students are more likely to persist in the face of difficulty when engaged. Similarly, writing prompts that are engaging generate more elaborated and developed writing than others. While engagement can boost the ability of all students to demonstrate what they know on an assessment, engagement is particularly important for the students least well-served in schools. More affluent students often find it easier to compartmentalize and to carry out academic tasks related to themes that they find dry and trivial. Some of the racial and cultural groups that currently do the worst in school also have a cultural orientation to learning through tasks that are meaningful and authentic and have difficulty with pursuing tasks on themes that seem irrelevant. The more barriers a student faces, the more the student needs and benefits from the boost of finding the themes that an assessment item addresses to be interesting and meaningful. Not all student populations will be equally engaged by all themes. When thinking about the issue of engagement in large-scale assessments, it is important to find themes that would not compound disadvantage. Themes should either be (a) engaging to all student subgroups or (b) engaging to groups least well-served in schools while not hurting the performance of other student populations. The issue of engaging themes is relevant to four types of items that will be included in the English language arts and math assessments aligned to the common core standards: (1) reading passages, (2) writing prompts, (3) projects and tasks involving research, and (4) math word problems.	Darling-Hammond, L., Barron, B., Pearson, D. P., et al. (2008). *Powerful Learning: What We Know About Teaching for Understanding,* pp. 74-76. San Francisco: Jossey-Bass.

Factor	Brief Summary of Research	Sources
Linguistic and cultural strengths	Meier (2008) establishes that, in the literature about reading comprehension, understanding is facilitated by a connection that a student has with his or her own experiences; this could be a text-to-self connection, a text-to-text connection, or a text-to-world connection. In this section, Meier provides teachers with strategies to help students make these connections so that students can then draw their own connections when they read materials on their own in other settings. Meier makes a point about the content of reading materials in classrooms: it is important for available reading books in a classroom to have content that reflects the student demographics in that class. It is also the duty of teacher to expose their students to literary texts that feature many cultural backgrounds. In this section, Meier models themes from African American literature for elementary school-aged children and outlines the ways that these cultural books have underlying themes that speak to most children's experiences. Meier reviews studies and literature that illustrate the cultural aspects of being raised in predominately African American communities and the linguistic skills that are developed from an early age. This review highlights vivid examples of non-academic contexts where young African American children learn how to use cognitively complex language behavior like building arguments, taking context into account, understanding multiple meanings of words, and using metaphorical and figurative language. These are the examples that are particularly applicable to reading cognitively complex texts.	Meier, T. (2008). *Black Communications and Learning to Read: Building on Children's Linguistic and Cultural Strengths*, pp. 119-121. New York: Lawrence Erlbaum.
Interpreting Student Performance	Smitherman (2000) analyzed 17-year old African American student NAEP writing samples and the scores they received between 1969 and 1989. The purpose of the analysis was to test a hypothesis of a decrease in the use of Black language in student writing over time and to understand how the use of Black language in student writing affected test scores. The study found the use of Black language had decreased in a 10 year period, but only in narrative essays. The study also found that between 1969 and 1979, student writing that included Black language was more likely to receive a higher score with the primary trait scoring rubric (a rubric that analytically scored writing with minimal weights based on grammar and syntax). In 1984 and 1988, student writing that included Black language was more likely to receive a lower score when scored according to a holistic rubric (a rubric that includes an assessment of grammar, mechanics, and syntax without specified weights for each element).	Smitherman, Geneva. (2000). African American Student Writers in the NAEP 1969-88/89[1992] and 'The Blacker the Berry, the Sweeter the Juice' [1994]. In *Talkin that Talk: Language, Culture and Education in African America*, pp. 163-191. New York: Routledge Press.

Factor	Brief Summary of Research	Sources
Interpreting student writing responses	Escamilla and Coady (2001) found that understanding writing conventions and patterns of ELL students' native language can be used in the assessment of writing samples written in English in order to fully understand student mastery of writing and ELA standards. Additionally, if ELL student writing is assessed in both English and their native language or through "contrastive analysis," the evaluator can analyze the writing skills that a student has mastered that cannot yet be expressed in English because of the student's level of language acquisition. Samples of student writing and the approach to this method for writing assessment are provided in the article.	Escamilla, K. & Coady, M. (2001). Assessing the writing of Spanish speaking students: issues and suggestions. In J. Tinajero and S. Hurley (Eds.). *Handbook for Literacy Assessment for Bilingual Learners*. Boston: Allyn & Bacon.
Interpreting ELL Student Performance in Math	Moschkovich (2008) examined third grade ELL group problem-solving discourse during a math class. Her data, recorded conversations among students and their teacher, revealed the verbalization of student math thought processes and reasoning patterns. She found that multiple meanings existed behind reasoning expressions. Moschkovich illustrated the way that recognizing these multiple meanings helped a teacher bridge student understanding to formal vocabulary and concepts. This recognition also helped the teacher not dismiss student understanding and grasp of math concepts.	Moschkovich, J. (2008). "I Went by Twos, He Went by One": Multiple Interpretations of Inscriptions as Resources for Mathematical Discussion. *Journal of the Learning Sciences*, 17:4, 551-587.

Appendix B: Original and Revised Performance Task to Increase Student Engagement

Heartbeats Classroom Activity ORIGINAL

Resources needed:

- Each student should have access to a piece of paper and writing tool
- Projector or some manner to display images
- A timing device for measuring a 20 second interval

Setting the Context

Facilitator says: "The performance task you will complete allows you to explore the body weights and pulse rates of different animals."

Facilitator says: "Let's start by talking about the body weights of different animals. Imagine a chicken, a dog, a horse, and a rat. On your paper, write the animals in order from lightest to heaviest according to their body weights. [Display the animals' names for students in the order listed: chicken, dog, horse, rat.]

Facilitator asks: "Which two animals do you think are closest in weight?" [Wait for responses. Responses may include the rat and the chicken, or some students may think the dog and the horse, depending on which breeds they are considering or because they have never seen the live animals. Students may explain why they made the choice they did.]

Facilitator asks: "How did you order the animals from lightest to heaviest?" [Record the orders that students provide, writing down different responses if they arise.]

Facilitator says: "Based on the average body weight, the order from lightest to heaviest is rat, chicken, dog, horse."

Modeling a Process

Facilitator says: "Next, we are going to think about pulse rate. This is the rate at which your heart beats. It is often measured in beats per minute. Today, you are going to find your pulse on your neck or your wrist. Using your middle and index fingers, gently touch your wrist to find the pulse beating. Do not use your thumb to take the pulse because it has a pulse of its own." [Display the image on the resource documents showing how to take a pulse via wrist.]

Facilitator says: "Everyone, try to find your pulse. Raise your free hand once you have found it." [Help students find their pulse on the wrist or show them another method for finding pulse.]

Facilitator says: "We are going to count the number of beats in 20 seconds. If you were not able to find your pulse, keep trying. You will not be graded on whether you found your pulse. When I say begin, start counting the number of beats. When I say stop, stop counting." [Wait for an appropriate starting point.]

Facilitator says: "Begin." [20 seconds pass.] "Stop."

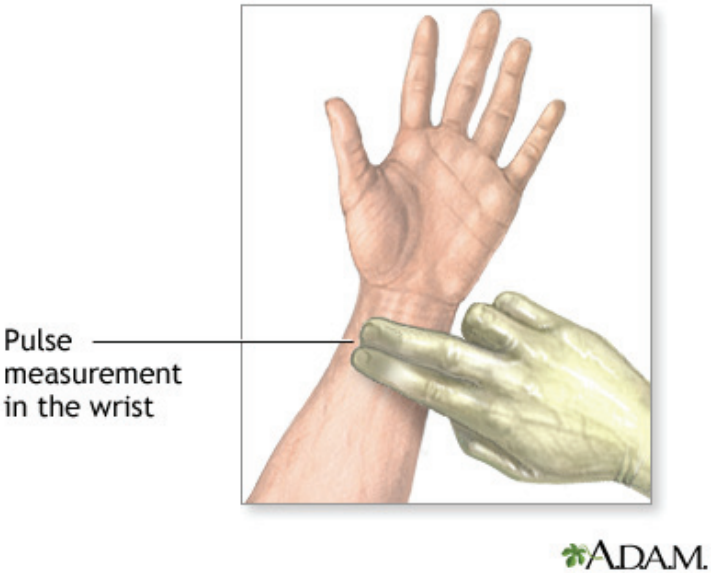
Facilitator says: "How many pulse beats did you count in 20 seconds?" [Collect and record different student responses.] "Notice that not everyone has exactly the same pulse rate. There are many factors that can affect pulse rate, such as age, body temperature, or exercise."

Facilitator says: "Earlier, I told you that pulse rate is often measured in beats per minute. What do you think we should do to the pulse rate we found to convert it to beats per minute?" [Students may have different responses based on their background in ratio and proportion. One possible solution is to multiply the rate by 3 since one minute is equal to 3 x 20 seconds, so this is the same as performing the counting exercise for three 20 second intervals, or one minute.] "Try converting your pulse rate to beats per minute." [Ask for a four or five conversions.] "How might you be able to tell if somebody had an incorrect conversion without checking each one individually?" [Students should notice that, if the original numbers were about the same, then the output should not be widely disparate.]

Facilitator says: "Now that we have talked about pulse rate and body weight, you are ready to complete the task."

[Begin performance task]

Resource Documents



Source: http://www.nlm.nih.gov/medlineplus/ency/imagepages/9799.htm

Heartbeats Performance Task ORIGINAL

In this task, you will develop a model of the relationship between the body weight and pulse rate of animals. You will examine additional data to evaluate the initial model.

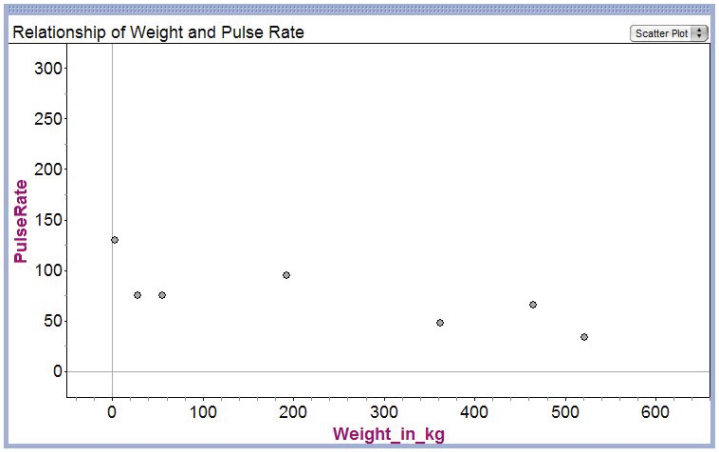
A study states that the relationship between an animal's pulse rate and body weight is approximately linear. The study data are below.

Table 1. Average Body Weight and Average Pulse Rate of Seven Animals

Animal	Average Body Weight (in kilograms)	Average Pulse Rate (in beats per minute)
Cat	3	130
Goat	28	75
Sheep	56	75
Pig	192	95
Ox	362	48
Cow	465	66
Horse	521	34

Item 1.

The data from Table 1 are plotted below. Use the Connect Line tool to create a linear model of these data.



Item 2.

What is the equation of the line you drew in Item 1?

Item 3.

Interpret the slope of the line from Item 1 in the context of the situation.

Item 4.

Based on the equation from Item 2, predict the pulse rate in beats per minute, of an animal that weighs 6000 kilograms.

Explain whether the predicted pulse rate in part (a) is reasonable in the context of the situation.

Item 5.

The body weight and pulse rate of a guinea pig and rabbit are given in the table below.

Animal	Average Body Weight (in kg)	Average pulse rate (in beats per minute)
Guinea Pig	1	250
Rabbit	2.5	265

If the study had included these data, would this change the model relating average body weight and average pulse rate? How do you know?

Heartbeats Classroom Activity REVISED

A Classroom activity introduces to students the topic or key vocabulary of the performance task. The activity provides an opportunity for activating students' prior knowledge and generating student interest in further exploration of the topic. It also provides students with an opportunity for interaction with the topic and with each other. The Classroom activity may be up to 30 minutes in length, but should be simple and easy to implement with clear instructions.

Classroom Activity

Resources needed:

- Each student should have access to a piece of paper and writing tool
- Projector or some means to display images
- A timing device (e.g., timer, stopwatch, or watch with second hand)

The purpose of the classroom activity is to teach students to determine their heart rate and to clarify some vocabulary that students will need to understand to complete the Performance Assessment they will complete today or in the coming days.]

Teacher asks: "Does anyone in the class know a simple way to measure your heart rate?"

[Teacher calls on students who have familiarity with taking a pulse and ask them to demonstrate and explain the process. The teacher verifies the process described by the student or students.]

[If no student knows how to take a pulse, the teacher demonstrates and explains using the Figure 1. Finger Placement for Pulse Rate below.]

Modeling Process for Teacher

Teacher says: "To measure your heart rate, you are going to find your pulse on your wrist. Using your middle and index fingers, gently touch your wrist to feel your pulse beating. Do not use your thumb to take the pulse because it has a pulse of its own." [Display Figure 1. Finger Placement for Pulse Rate showing how to take a pulse via the wrist.]

Students Measure Resting Heart Rate

Teacher says: "We are going to think about what we mean by a resting heart rate. Based on the name, what do you think it means?"

[Students share ideas. If no one comes up with a definition, the teacher can share the following definition: A resting heart rate is your heart rate when your body is not moving. A resting heart rate is usually taken when you first wake up in the morning or after you have been lying down for a while.]

Teacher says: "You will work with a partner to find both of your resting heart rates."

[The teacher helps students find their pulse on the wrist and then helps students find their pulse for one minute. Students may do this a variety of ways. For example, some pairs may have one person keep time for 60 seconds while the other counts heart beats. Others may count for 10 seconds and multiply by 6, or count for 20 seconds and multiply by 3, etc.]

Teacher asks: "What are your resting heart rates is in beats per minute."

[The teacher can ask students to record their heart rates on the board or on a visual display; or the teacher can collect and record different student responses. The teacher highlights that not everyone has exactly the same heart rate and that there are many factors that can affect heart rate, such as age, body temperature, or exercise.]

Teacher asks: "How did you and your partner count and figure out your beats per minute?"

[Students may have different responses based on their background in ratio and proportion. One possible solution is to multiply the rate by 3, since one minute is equal to 3 x 20 seconds; so this is the same as performing the counting exercise for three 20-second intervals, or one minute.]

Teacher says: "Now you will learn about another type of heart rate. When you exercise so hard that you think your body will collapse, your heart is beating at its maximum capacity. If you measured your heart rate at that very moment, it would be called your maximum heart rate. Many professional athletes push their bodies to the limit and they reach their maximum heart rate and collapse."

[The teacher explains that it is important to exercise, but you can seriously hurt your body if you exercise at the maximum heart rate. To avoid hurting your heart, doctors recommend that you exercise at your target heart rate.]

Teacher asks: "Who in the class can explain what a target heart rate is?"

[The teacher calls on students who want to explain target heart rate. If no students volunteer, teacher can explain that a target heart rate is a percentage of your maximum heart rate. It is why people recommend that, when you exercise, you should still be able to talk.]

Teacher says: "You will complete a task that is about the three different types of heart rates: resting, maximum, and target heart rate. Do you have any questions about these three types of heart rates?"

Teacher says: "You are ready to complete the Heartbeats performance task. There are many different ways to accomplish the task. This task has several parts. Look over the entire task to understand the goal. Use what you know. If you get confused, review the questions and information provided to see what you have already answered. If possible, skip one part and see if other parts can help you with a part you might have missed. Please remember to try your best on every part of the Performance Task so that you will get credit for all of your thinking."

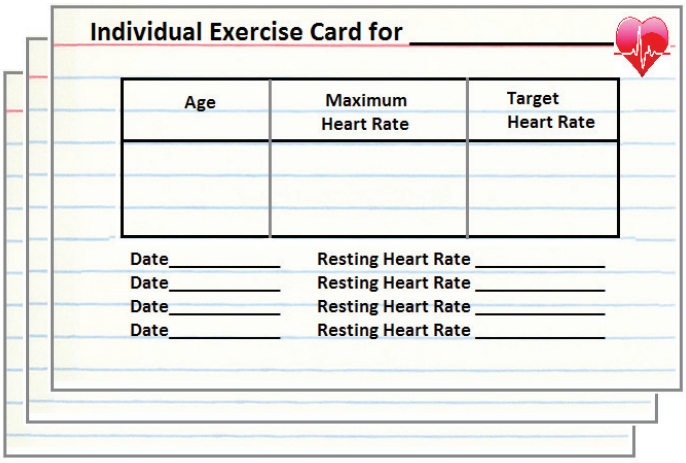
Heartbeats Performance Task REVISED

Purpose

Your community is organizing an exercise challenge day to motivate people to exercise more as a way to improve their health. People who come to the event will learn about different types of exercise and how to exercise safely. You have volunteered to help people of all ages (children, teenagers, adults, and senior citizens) learn about their resting, maximum, and target heart rates so they will know how to exercise without putting too much strain on their hearts and bodies.

Task

In this task, you will first review the different types of heart rates and evaluate data to learn more about the relationship between heart rate and age. Then you will complete your own and help an adult complete his or her Individual Exercise Card (see below) and explain to him or her what these numbers represent. (There are other individual health considerations, but for this task you will only use age to give exercise guidance.)



Key Terms

Below are definitions of the different types of heart rates. These distinctions are important to understand so that participants at the event do not get injured or dizzy during exercise.

Glossary

Resting heart rate	The number of heart beats per minute when you are at rest (when you wake up in the morning, for example).
Maximum heart rate	The highest number of beats per minute an individual can achieve without causing severe problems to the body. This rate depends on age.
Target heart rate	The number of beats per minute that burns calories and is still safe for your body. This number is usually 50% to 70% of a person's maximum heart rate.

Data

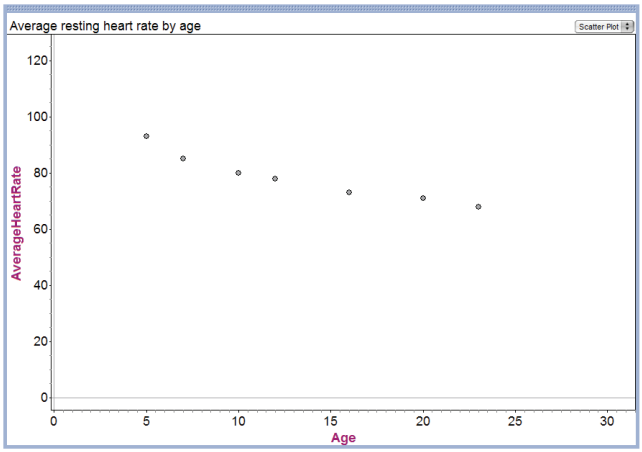
While researching the topic, you learn that heart rate is an important indicator to monitor during exercise. To learn more about the relationship between heart rate and age, you decide to examine Table 1 and Graph 1 below.

Table 1 is a data set of the average resting heart rate (in beats per minute) for healthy people at different ages and Figure 1 is a graph of the same data set.

Table 1. Average Resting Heart Rate by Age[5]

Age	Average resting heart rate (in beats per minute)
5	93
7	85
10	80
12	78
16	73
20	71
23	68

Graph 1: Average Resting Heart Rate by Age



[5] Overall estimates for U.S. national resting pulse rate: National Health and Nutrition Examination Survey, 1999–2008 http://www.cdc.gov/nchs/data/nhsr/nhsr041.pdf

Table 2. Average Resting Heart Rate by Age

Age	Average resting heart rate (in beats per minute)
1	116
2	112
3	105
4	100

Table 3. Equations for Calculating Maximum Heart Rate[6]

Original equation (1970) Maximum heart rate = 220 – Age
Revised equation (2001) Maximum heart rate = 210 - [.5 x Age]

1. Estimate your resting heart rate, in beats per minute, based on the data provided to you. Enter your estimate and explain how you figured this out.

2. Table 1 and Graph 1 present the same data on average resting heart rate by age. Describe the relationship between age and average resting heart rate.

3. Using the data in Table 1 and/or Graph 1, what is your prediction of the average resting heart rate for a 90-year-old person in beats per minute?

   Explain whether you think your prediction is reasonable and support your reasoning with evidence.

4. As you continue your research, you find heart rate data for younger children. Look at the data in Table 2. Evaluate how well your model for predicting the average resting heart rate of a 90-year-old (Item 3 above) works for predicting the average resting heart rate of young children, ages 1 through 4. Explain the results of your evaluation using at least two specific examples to support your answer.

5. In your research, you also learn that there are two established equations used by doctors to help people of different ages estimate their maximum heart rate. Look at the two equations in Table 3. People will want to know which equation they should use to estimate their maximum heart rate.

[6] http://www.digifit.com/heartratezones/maximum-heart-rate.asp

6. Explain whether the two equations give similar results for people of all ages. Include specific examples to support your evaluation by finding the maximum heart rate using both equations for at least two different ages. Discuss the similarities and differences between these sets of results.

Now that you have completed your research about heart rates, you can complete the last part of the exercise card showing the maximum and target heart rates.

- The *maximum heart rate* can be determined using one of the equations from Item 5.
- The *target heart rate* is equal to the range between 50%-70% of a person's maximum heart rate.
- Use this information to complete the exercise card. Enter the information for yourself and for Ms. Jones, a teacher in your school (age 49).

Name	Age	Maximum Heart Rate	Target Heart Rate Range
Your name			
Ms. Jones	49		

Use specific examples to explain to Ms. Jones why her target heart rate range is not the same as yours.

STANFORD CENTER FOR
OPPORTUNITY POLICY IN EDUCATION
505 Lasuen Mall
Stanford, CA 94305
Email: scope@stanford.edu
http://edpolicy.stanford.edu

STANFORD CENTER FOR ASSESSMENT,
LEARNING AND EQUITY
1705 El Camino Real
Palo Alto, CA 94306
https://scale.stanford.edu/