Designing for Engineering: A Model for Integrating Engineering and Science NGSS Middle School Benchmark Assessments

Maia K. Binding and Lauren Brodsky

Abstract

This study focuses on a project design approach for developing item sets for the NGSS engineering performance expectations (PEs) for middle school. It further examines how the design approach allows the integration of assessments for engineering PEs and for science PEs and how resulting item sets elicit three-dimensional responses from students for both PEs. It addresses the key question: *Does the project design approach allow for the integration of the NGSS ETS PEs with science PEs such that the resulting items elicit 3D student performances aligned to both PEs?* Findings from three evidence sources are analyzed and discussed: 1) multi-curriculum review, 2) external expert review, and 3) external testing. Results indicate that the project design and development approach allows for full integration of ETS and science PEs, and that resulting assessment items elicit three-dimensional responses in line with both PEs being assessed.

Background and Problem

The advent of the Next Generation Science Standards (NGSS Lead States, 2013) and the Framework for K-12 Science Education (NRC, 2012) has heralded the opportunity for a promising, significant shift in science teaching and learning in the United States. The NGSS move science and engineering education from teacher-centered, content knowledge acquisition to students developing and applying knowledge in new situations (Krajcik et al., 2014). The NGSS address a smaller number of topics than previous standards, but with more depth (NGSS Lead States, 2013), present science as a process rather than a product (NRC, 2012), and incorporate engineering as integral content deeply connected to the Earth, life, and physical sciences. Intertwined with this depth of coverage are three dimensions of content: science and engineering practices (SEPs), crosscutting concepts (CCCs), and disciplinary core ideas (DCIs). This multi-faceted shift in thinking requires a parallel change in how students' understanding of these three dimensions is assessed, and presents a design challenge for assessing engineering standards in an authentic way.

A complete assessment system should include a range of assessments at all levels, from classroom assessments embedded in specific curricula, to classroom benchmark assessments, to state-level tests typically administered every few years (NRC, 2014; Osborne et al 2016; Shepard et al, 2017). As the need for NGSS-designed assessments has become clear, several projects have developed exemplars for three-dimensional assessments, while others have developed three-dimensional assessments specific to their NGSS-designed curricula, but there are limited engineering assessments available outside of those embedded with curricula (Harris et al., 2019, MacPherson et al., 2017, SNAP, 2018, Zaidi, et al., 2018). Meanwhile, schools and districts across the country are utilizing a wide range of curricula and materials that incorporate engineering in different ways. Additionally, standardized assessments are being developed at the district and state level that are not necessarily aligned to a particular curriculum. Given these circumstances, students need to be provided the opportunity to apply their engineering understanding to assessments that are not embedded in or based on their curriculum before they take standardized assessments. The NGSS leaves room for the use of a variety of visual representations, phenomena, and aspects of the presentation of the three dimensions within curricula. This is particularly true of the engineering content, which can be embedded in curricula in a variety of ways. Students will benefit from practice with assessments that require them to transfer their understanding, including that of engineering and design, to contexts that are different from what they have experienced with their particular curriculum. In 2019 the Lawrence Hall of Science launched a two-year project, funded by the Carnegie Corporation of New York, to develop summative, curriculum-neutral, NGSS-designed, three-dimensional assessments for all of the middle school NGSS Performance Expectations (PEs) in Earth, Life, and Physical Sciences and Engineering. These model assessments are designed to serve as a bridge between the curriculum and the standardized assessments. Within this context, the project team tried two models for developing items for the Engineering Design (ETS) PEs. We developed items for the ETS PEs on their own using an existing, tested approach. Wanting to create items that have students apply more science content through engineering, we then developed an approach to assessing the ETS PEs where the ETS PE dimensions are layered onto an existing science PE item set, as a proposed method for integrating science content with the engineering content within the assessment. This paper aims to outline the new design processes, item development, and results of this project specific to the integrated ETS PEs, sharing the lessons learned thus far. In this way it will address the question: *Does the project design approach allow for the integration of the NGSS ETS PEs with science PEs such that the resulting items elicit 3D student performances aligned to both PEs?*

Assessment Item and Study Design

This work builds on work from a collaborative project between The Lawrence, the American Museum of Natural History, and the University of Connecticut, funded by the National Science Foundation (NSF). This project developed a model NGSS-aligned curriculum unit, with embedded and benchmark assessments (*Disruptions in Ecosystems,* DRL 1418235). Several of the end-of-chapter assessments were identified as model assessments by the Achieve *Task Annotation Project in Science* (personal communication, Achieve, 2019). A supplement to the NSF grant provided the project team with the opportunity to score and analyze student responses to these assessments. This work provided significant insight in determining which assessment features and characteristics best elicited three-dimensional responses and should be carried into the current project. The project has also reviewed published model items and research on specific content areas and practices, when available (Badrinarayan et al., 2019; McElhaney et al., 2019).The assessment development approach for the current project has been further informed by the work of others in the field, including the practice of evidence-centered design (Almond, et al., 2002; Mislevy & Haertel, 2006) and research and design tools and approaches for three-dimensional assessment development. These design tools and approaches include the Next Generation Science Assessment (NGSA) approach described in Harris et al., 2019, and *STEM Teaching Tool #29* (Penuel et al, 2016).

The project team consists of curriculum and assessment developers, the majority of whom are former classroom teachers. The project approach for assessment development for the science PEs has been to create an item set to address each PE on its own, to allow for maximum flexibility in terms of how different curricula might have grouped PEs into different units. For the ETS PEs, the project team began with a similar approach, addressing each ETS PE on its own with a unique item set. It became clear that there was significant overlap between the ETS item sets and item sets for science PEs that had the Science and Engineering Practice of Designing Solutions and that included the same ETS DCIs as secondary DCIs. Based on the significant overlap and the potential for more robust items that integrated science ideas with engineering, the project team tried a new approach for layering ETS dimensions onto existing science PE item sets. This study describes that approach and our results.

For the development of items that integrate science with engineering, for each ETS PE, a team member began by unpacking each dimension of the PE and used the results of the unpacking to create ETS Knowledge, Skills, and Abilities (KSAs). They then used the KSAs to draft initial learning performances (LPs) that were science-content agnostic, describing general ways that students might show understanding of the PE (e.g., from ETS1-4, LP1: Students select the most promising solution out of a set of possible solutions. LP2: Students develop (or suggest revisions to a provided) model that can be used to test a proposed solution. LP3: Students describe a plan for using a model to test a proposed solution in a way that can lead to modifications and further iteration). The unpack, KSAs, and general learning performances were then reviewed by the project team. The developer next identified existing assessment tasks designed for science PEs that include aspects of engineering as the practice (Science and Engineering Practice 6: Constructing Explanations and Designing Solutions), and that include ETS DCI elements as secondary DCIs to the PE. Identified science tasks were analyzed against the ETS KSAs to determine which KSAs from the ETS PE were missing from the existing science task. Once those target KSAs were identified, the developer took the general learning performances that incorporated those KSAs, and created design patterns that took into account the context of the existing science task. In practice this resulted in small adjustments and additions to the existing science task, often in the form of one or two additional prompts. In this way the unique ETS PE aspects were layered onto the science PE item set, and the engineering content was integrated with science content. These changes were again reviewed by the full team. Once revised item drafts were complete, sample student responses and scoring guides were developed and went through a similar review and revision process.

A typical item set that integrates one science PE and one ETS PE includes a range of 2–5 items. The items begin with scenarios or phenomena for students to respond to. Most of the item sets include simple diagrams to enhance accessibility. Item sets cover all three dimensions of the science and ETS PEs being assessed. While some individual items are three-dimensional, other items are two-dimensional when developers determined that more targeted items were needed to assess student understanding of specific aspects of the dimensions of the PEs.

Study Design

To address the research question, *Does the project design approach allow for the integration of the NGSS ETS PEs with science PEs such that the resulting items elicit 3D student performances aligned to both PEs?,* data was collected from three different sources: multi-curriculum group review, external expert review, and external testing. The assessments will be available as open resources, for teachers to use and adapt to meet their specific needs. As teachers select and modify the assessments, most validity claims, especially any claims about test performance and certain aspects of inferential validity, will no longer apply. Given these conditions, the project has focused on a limited evaluation of the validity of the assessments to support principled adaptation of the assessments by educators for their student populations.

Multi-curriculum group review. The full project team includes representatives from three established curriculum programs: Full Option Science Systems (FOSS), the Learning Design Group (LDG), and the Science Education for Public Understanding Program (SEPUP), which is also the project lead. Each step of the item design process (unpack, learning performances, design patterns, items, sample student responses and scoring guides) received a general review from all team members for alignment to both science and ETS PEs and item clarity. Individuals from each program specifically reviewed for the alignment of each step of the item design process with their curriculum interpretation of the ETS PEs. Knowing that the different curricular programs had varying interpretations of the PEs, reviewers documented instances where this occurred in the unpack of the PE. They also noted aspects of design patterns and items where significantly different graphic representations or vocabulary were used, for example the use of straight arrows versus wavy arrows for showing energy transfer.

External Expert Review. Item sets for both stand-alone ETS PEs and integrated science and ETS PEs were also sent for review by an external expert group, then underwent a final round of review and revisions by the project team. The expert review group analysis focused on the ETS dimensions for all the item sets, and additionally on the science PE dimensions for one sample set where the PEs were integrated. The expert group rated items against a rubric that included a category for ability to elicit evidence about the three dimensions of the PE (*To what extent does the item prompt students to provide a response that integrates the three dimensions sought by the PE?*). Scores were given on a scale of 1-5 and comments were provided for each item within a set and also for the overall item set.

External Testing. Due to the Covid 19 pandemic, classroom field tests were not able to be administered per the project's IRB protocol. In lieu of this, both stand-alone and integrated ETS PE items were tested with external staff members from each of the three curriculum projects involved who were otherwise not a part of this assessment project. To better emulate students who had experienced an NGSS-aligned unit on the relevant PEs, external staff were chosen who had familiarity with the NGSS and varying science backgrounds and classroom teaching experience. Participants were instructed to do their best to limit their response to content that would likely be covered in an NGSS-aligned middle science course, i.e., not to rely on extensive background knowledge. Responses were analyzed for three-dimensionality with regard to both the science PE and the ETS PE.

Analyses and Findings

Does the project design approach allow for the integration of the NGSS ETS PEs with science PEs such that the resulting items elicit 3D student performances aligned to both PEs? The extent to which the integrated design approach led to the design of items that elicit 3D student performances aligned to the ETS and science PEs was analyzed by looking across the multi-curriculum group review, external expert review, and external testing.

The analysis by the multi-curriculum review found that all integrated item sets were designed to elicit 3D performances aligned with both the ETS and science PEs. The three curriculum groups integrated the ETS PEs and associated dimensions with different science content in their programs, and had slight variations in interpretations of the ETS PEs. For example, one curriculum integrated all four ETS PEs with life science PEs in a bioengineering unit, whereas another curriculum integrated the same ETS PEs with Earth science PEs over multiple Earth science units. Developers ensured that the scenarios, phenomena, and contexts used in the item sets differed from those in all of the project-related curricula. Despite these differences, each group found that the item sets were able to elicit understanding of the ETS content in line with their respective curricula. The analysis found that the integrated approach of layering ETS onto science items elicited 3D responses for both the science PE and the ETS PE. All three groups additionally noted that even though the science content integrated with the ETS PEs may have differed from what was used in their curricula, the engagement with engineering content when science content was included was stronger and more similar to what was supported in their curricula, as compared with the stand-alone ETS items.

The analysis by the external expert review team focused on the ETS PE dimensions and one sample of the science PEs. Analysis indicated all of the item sets elicited 3D performances aligned with the NGSS ETS PEs. Three of the four ETS item sets reviewed rated 4 or 5 for the category of all rubric categories, with a score of 4 indicating minimal revisions were suggested to improve 3D alignment of the items, such as small changes to vocabulary or slight improvements to clarity in scenarios or prompts. The fourth item set received mostly scores of 4 or 5, with a score of 3 for the integration of crosscutting concepts, with suggestions for simple modifications to address this issue. Additional reviewer comments were particularly positive regarding the potential for student engagement and accessibility. Comments consistently indicated that while the stand-alone ETS PE items did elicit the dimensions of the ETS PEs, the integrated item allowed for deeper sensemaking in addition to eliciting the dimensions of both PEs. An example item that integrated ETS1-4 with PS2-1 received a score of 5 out of 5 for the category of *How well does the item set probe students' learning related to the three elements of the PE?* and associated comments included, "...this version with PS integrated better reflects the spirit (if not the letter) of the PE's intentions for sensemaking…" with the conclusion that, "Ultimately it will come down to opportunity to learn for these tasks - if students have had the opportunity to learn relevant PS content, they will be better of with this [integrated] version."

The analysis of the external testing found that the integrated items did elicit 3D responses for both ETS and science PEs. In moderated scoring sessions, scorers found that participants from different curriculum groups consistently provided responses that demonstrated understanding of all relevant dimensions. In discussion, it was noted that scorers identified the same aspects of responses as evidence for a given dimension. Comparison to the stand-alone ETS PE items indicated that the integrated items elicited equally complete responses to all dimensions of the ETS PEs. Figure 1 below shows the dimensions associated with one item set in which ETS1-1 is layered onto an existing item set for science PE PS3-3. Layering the ETS dimensions onto the PS item set about designing a container to minimize thermal energy transfer entailed adjusting the final item to incorporate attention to the ETS CCC of *Influence of Science, Engineering, and Technology on Society and the Natural World.* In this case, an additional prompt was added to the second item to elicit student thinking about how human activity draws on natural resources.

PE	SEP	DCI	CCC
ETS 1-1	Asking Questions and Defining Problems	ETS1.A Defining and Delimiting Engineering Problems	Influence of Science, Engineering, and Technology on Society and the Natural World
PS 3-3	Constructing Explanations and Designing Solutions	PS3.A Definitions of Energy PS3.B Conservation of Energy and Energy Transfer ETS1.A Defining and Delimiting Engineering Problems ETS1.B Developing Possible Solutions	Energy and Matter

Figure 1: NGSS Dimensions for MS-ETS1-1 and MS-PS3-3

Given each participant's curriculum background, variation was found in more superficial features of the items, such as vocabulary and visual representations used in responses. For example, participants from one curriculum group were more likely to use "energy transfer" while those from a different curriculum group were more likely to use "heat" in their responses. This variation was used to update scoring guides to make them more able to capture differences in responses, and to add additional administrative guidance for teachers who may want to modify the assessments for their students. For example, alternatives were suggested for language used around iteration and optimal design, how the engineering design cycle is referred to, and standard symbols used in design diagrams. Appendix A shows an integrated item set for ETS1-1 and PS3-3 with example responses from one participant. The responses are annotated to show the dimensions of each PE as identified by the scorers. Appendix B shows the accompanying scoring guide.

Item sets are undergoing a final round of review and revisions based on the analysis discussed in this paper. Upon completion of this process, all item sets, along with supporting documents (sample student responses, scoring rubrics, and teacher support documents) will be published and made freely available on an online platform.

Contribution and General Interest

In addition to the specific item sets and analysis described thus far, the work presented above provides support for classroom practitioners, school districts, and curriculum and assessment researchers working to assess the ETS PEs.

Much of the analysis described above supports the integration of ETS DCIs and PEs with science PEs whenever possible. Our findings suggest that this leads to better opportunities for student sensemaking and for more authentic assessment of student knowledge and understanding. Additionally, integration of ETS DCIs and PEs with science PEs is common practice across many curricula. If ETS and science PEs are being "bundled" for classroom instruction, it is logical to integrate them in the summative assessments. As has been demonstrated above, this is both possible and practical. It should be noted that particular attention must be paid to all dimensions across both PEs. As described, simple adjustments to item prompts can often allow for better integration of dimensions. If teachers are bundling PEs for instruction, a careful examination of the dimensions for each PE should provide insight into what adjustments are necessary. This will often lead to stronger assessments, and will elicit more complete student responses.

Additionally, for science PEs that have ETS DCIs aligned as secondary DCIs, our work suggests that these science PEs can often be fully integrated with an ETS PE to allow for a more robust assessment of student understanding.

However, it is also important to recognize that integration of ETS and science PEs is not always appropriate. For example, a standardized state-level benchmark assessment that will be given across districts where different curricula are being used would be better served with stand-alone ETS and science item sets. This would prevent any bias for students who might have had a curriculum with PE integration that happens to align with the benchmark assessment. In this instance, assessment developers will need to balance having item scenarios that are engaging and elicit three-dimensional responses without bringing in science content students may not have been exposed to. Additionally, developers will need to ensure that the resulting assessment is not simply an exercise in reading comprehension, but truly assesses students' understanding of all of the dimensions of the PE in question.

The aim of this project is to design item sets that assess single PEs, bridging the space between curriculum, curriculum-embedded assessments, and standardized exams. The lessons learned from this work have provided valuable insight into ETS assessment development in many critical ways, and should serve as a model for both researchers and practitioners developing items.

Acknowledgements

The authors would like to acknowledge the project team at the Lawrence Hall of Science for their continuing hard work, all of the teachers and students who have participated and shared their knowledge, and their colleagues at WestEd for their expert reviews and advice.

This work was made possible by a grant from the Carnegie Corporation of New York. The statements made and views expressed are solely the responsibility of the authors.

References

Almond, R. G., Steinberg, L. S., & Mislevy, R. J. (2002). Enhancing the design and delivery of assessment systems: A four-process architecture. Journal of Technology, Learning, and Assessment, 1 (5). Retrieved from http:// ejournals.bc.edu/ojs/index.php/jtla/article/view/1671

Badrinarayan, A., Wertheim, J., Penuel, B., Krajcik, J, and Smolek, T.J. (2019). Reconceptualizing Alignment for NGSS Assessment. Paper presented at NARST 2019 Annual International Conference, Baltimore, MD.

Harris, C. J., Krajcik, J. S., Pellegrino, J. W., & DeBarger, A. H. (2019). Designing knowledge-in-use assessments to promote deeper learning. *Educational Measurement: Issues and Practice*, 38(2), 53–67.

Krajcik, J., Codere, S., Dahsah, C., Bayer, R., & Mun, K. (2014). Planning instruction to meet the intent of the Next Generation Science Standards. *Journal of Science Teacher Education*, *25*(2), 157–175. http://doi.org/10.1007/s10972-014-9383-2

MacPherson, A., Kastel, D., Nagle, B., Willcox, M., Jackson, W., Hariani, M., and Howarth, J. (2017). NGSS-designed assessments for a middle school ecosystems unit. Paper presented at the NARST 2017 Annual International Conference, San Antonio, TX.

McElhaney, K. W., Basu, S., Wetzel, T., and Boyce, J. (2019). Three-dimensional assessment of NGSS upper elementary engineering design performance expectations. Paper presented at NARST 2019 Annual International Conference, Baltimore, MD.

Mislevy, R., & Haertel, G. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*(4), 6–20.

National Research Council (2012). *A Framework for K-12 Science Education: Practices, crosscutting concepts, and core ideas*. Committee on a Conceptual Framework for New K-12 Science Education Standards, Board on Science Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council. (2014). Developing Assessments for the Next Generation Science Standards. Committee on Developing Assessments of Science Proficiency in K–12. J.W. Pellegrino, M.R. Wilson, J.A. Koenig, and A.S. Beatty, eds. Board on Testing and Assessment and Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

NGSS Lead States. 2013. *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.

Osborne, J., Pecheone, R., Quinn, H., Holthuis, N., Schultz, S., Wertheim, J., and Martin, P. (2016). A system of assessments for the next generation science standards (NGSS) in California: A discussion document. Prepared by the Stanford NGSS Assessment Project Team (SNAP). Available at https://snapgse.stanford.edu/snap-reports/snap-reports

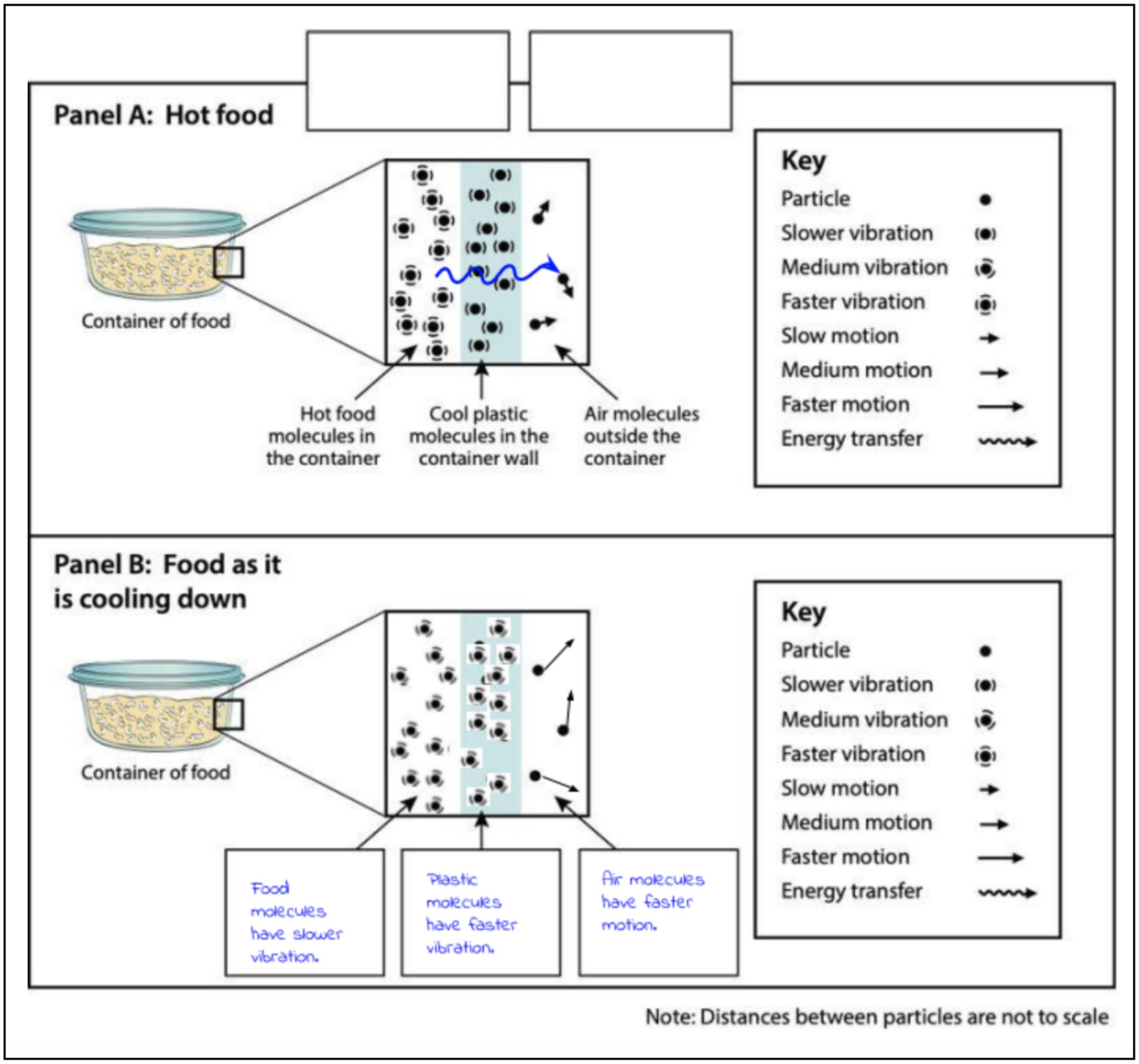
Penuel, W.R., Van Horne, K., and Bell, P. (2016) Steps to Designing a Three-Dimensional Assessment. UW Institute for Science + Math Education. Available at stemteachingtools.org/brief/29.

Shepard, L. A., Penuel, W. R., & Davidson, K. L. "Design principles for new systems of assessment," *Phi Delta Kappan*, v.98, 2017.

Stanford NGSS Assessment Project (SNAP). (2018). Available at https://snapgse.stanford.edu/snap-assessments-ngss.

Zaidi, S.Z., Ko, M., Gane, B.D, Madden, K., Guar, D., and Pellegrino, J.W. (2018). Portraits of teachers using three-dimensional assessment tasks to inform instruction. Paper presented at the 2018 NARST Annual International Conference, Atlanta, GA.

Appendix A - Sample Item Set for ETS1-1 and PS3-3



Scoring note: Demonstrates understanding of DCIs PS3.A and PS3.B from PS3-3, and the CCC of Energy and Matter from PS3-3

b. How would you improve the design of Kira's container to keep her macaroni and cheese hot? Use science ideas to explain why it would be an improvement.
Kira needs a container that prevents the transfer of heat to the outside. The molecules that make up the container should not absorb the heat from the food.

Scoring note: Demonstrates understanding of DCI from PS3-3, and SEP 1 Asking Questions and Defining Problems from ETS1-1, and the CCC Energy and Matter from PS3-3

Item 2

Your friend Kira decides to get a new lunch container and asks you for help in designing a good one. She wants one that will have room for both macaroni and cheese and a juice box. She wants to keep the macaroni and cheese hot and the juice box cold at least until lunch time. She is having trouble finding one in a store and she asks you to design one for her. She only has $5.00 to give you to use for it, and she needs to be able to carry it in her backpack.

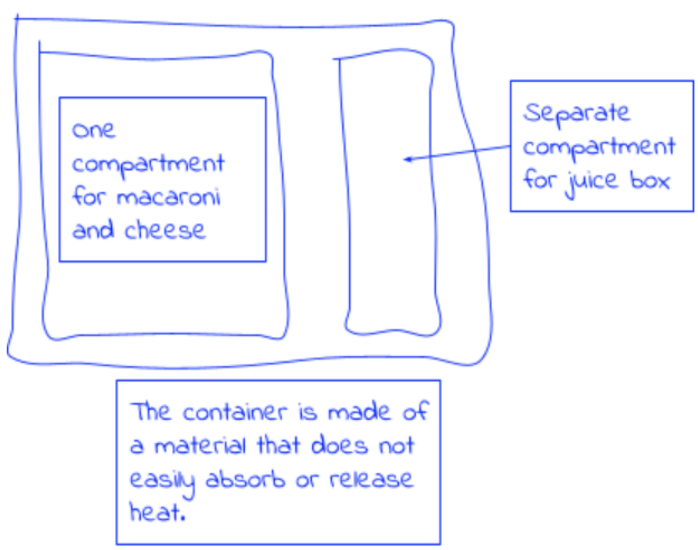
a. Identify the criteria and constraints of your design.

Criteria:
Keeps macaroni and cheese hot until lunch
Keeps juice box cold until lunch

Constraints:
Cost $5 or less
Must fit in backpack

Scoring note: Demonstrates understanding of DCI ETS1.A from PS3-3 and ETS1-1

b. Sketch a design of a new container based on science ideas about energy transfer. Be sure to label your design.



Scoring note: Demonstrates understanding of DCI PS3.B from PS3-3, SEP 6 Constructing Explanations and Designing Solutions from PS3-3, and CCC Energy and Matter from PS3-3

c. You will need to run a test to determine if your design meets the criteria and constraints. Answer the questions below about your design test.

How would you set up your test?
`I will place hot water in the macaroni and cheese compartment and cold water in the juice box compartment.`
 	
What data will you collect?
`Temperature of water over time`
`Cost of materials`
`Dimensions of the container`
 	
What results would show that you met your criteria and constraints?
`The hot water will stay hot for multiple hours and the cold water will stay cold for multiple hours. The cost of the materials is less than $5 and the dimensions are small enough to fit in a backpack.`

Scoring note: Demonstrates understanding of DCI ETS1.B from PS3-3, and SEP 6 Constructing Explanations and Designing Solutions from PS3-3

d. Kira cares about the environment. Identify one specific criterion she could add if her goal is to reduce the impact of her lunch bag on the environment. Explain why it would make a difference for the environment.
`Her lunch box could be made recyclable. When her lunch box is no longer used, it could be recycled into something new instead of becoming trash. This would reduce adding trash to the environment`

Scoring note: Demonstrates understanding of DCI ETS1.A from PS3-3 and ETS1-1, and CCC Influence of Science, Engineering, and Technology on Society and the Natural World from ETS1-1

Appendix B - Sample Scoring Guides for ETS1-1 and PS3-3

Scoring Guide - Item 1	
Level	**Item-Specific Description**
4 Complete and Correct	Student's response includes all four of the following components: ● Panel A shows thermal energy arrows going from the food to the plastic wall and from the plastic wall to the air. ● Panel A: The captions correctly describe the direction of thermal energy transfer. ● Panel B shows that the particles in the food have slowed down, the particles in the plastic wall have sped up, and the particles in the air have sped up. ● Panel B: The captions correctly describe the temperature and/or motion of the particles in all three locations.
3 Almost There	Student's response includes three of the following: ● Panel A shows thermal energy arrows going from the food to the plastic wall and from the plastic wall to the air. ● Panel A: The captions correctly describe the direction of thermal energy transfer. ● Panel B shows that the particles in the food have slowed down, the particles in the plastic wall have sped up, and the particles in the air have sped up. ● Panel B: The captions correctly describe the temperature and/or motion of the particles in all three locations.
2 On the Way	Student's response includes two of the following: ● Panel A shows thermal energy arrows going from the food to the plastic wall and from the plastic wall to the air. ● Panel A: The captions correctly describe the direction of thermal energy transfer. ● Panel B shows that the particles in the food have slowed down, the particles in the plastic wall have sped up, and the particles in the air have sped up. ● Panel B: The captions correctly describe the temperature and/or motion of the particles in all three locations.
1 Getting Started	Student's response includes one of the following: ● Panel A shows thermal energy arrows going from the food to the plastic wall and from the plastic wall to the air. ● Panel A: The captions correctly describe the direction of thermal energy transfer. ● Panel B shows that the particles in the food have slowed down, the particles in the plastic wall have sped up, and the particles in the air have sped up. ● Panel B: The captions correctly describe the temperature and/or motion of the particles in all three locations.
0	Student's response is missing, illegible, or irrelevant.
X	Student had no opportunity to respond.

Scoring Guide - Item 2		
Level	**General Description**	**Item-Specific Description**
4 Complete and Correct	Student's design: ● meets all the criteria within the defined constraints, AND ● has further improved on design, AND ● uses relevant scientific concepts to explain why any revision were made to optimize the design.	Student's design ● suggests appropriate criteria and constraints, AND ● has improved on the original container's design, AND ● explains how this new design helps to prevent the transfer of thermal energy from the macaroni and cheese to the cold drink and the room and from the room to the cold drink. AND ● suggests appropriate environmental criterion.
3 Almost There	Student's design ● Meets all the criteria within the defined constraints AND ● Explains the relevant scientific concepts.	Student's design ● suggests appropriate criteria and constraints, AND ● has improved on the original container's design, AND ● provides a partial explanation for how this new design helps to prevent the transfer of thermal energy AND ● suggests appropriate environmental criterion.
2 On the Way	Student's design ● meets all the criteria but exceeds the defined constraints OR ● meets some of the criteria within the defined constraints.	Student's design ● suggests only one or two criteria and constraints AND/OR attempts improvements on the design, BUT ● they are not complete, and the response provides no explanation or an incorrect explanation for how this new design helps to prevent the transfer of thermal energy.
1 Getting Started	Student's design does not meet any of the criteria	
0	Student's response is missing, illegible, or irrelevant.	
X	Student had no opportunity to respond.	