Connecting Science Concepts and Engineering Practices: Supporting Student Understanding of Energy Transformation

By

Elizabeth Anne McBride

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Science and Mathematics Education

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Marcia F. Linn, Chair
Professor Michelle H. Wilkerson
Professor Robert Rhew

Summer 2018

Connecting Science Concepts and Engineering Practices: Supporting Student Understanding of Energy Transformation

by

Elizabeth A. McBride

Abstract

Connecting Science Concepts and Engineering Practices: Supporting Student Understanding of Energy Transformation

by

Elizabeth A. McBride

Doctor of Philosophy in Science and Mathematics Education

University of California, Berkeley

Professor Marcia F. Linn, Chair

It is often claimed that engineering projects improve student achievement in mathematics and science, but research on this topic has shown that many projects do not live up to the claim (Teacher Advisory Council, 2009). Ideally, undertaking a science project should be motivating, while also helping students to understand the interplay between science concepts (like energy transformation) and engineering design decisions. This dissertation research investigates ways to integrate engineering practices and science concepts (like energy transformation) in classroom settings. I investigate ways to integrate the Next Generation Science Standards (NGSS) science and engineering practices while simultaneously expanding the knowledge integration theory (Linn & Eylon, 2011). I refine knowledge integration design principles in classroom studies, comparing alternative forms of instruction where students integrate engineering design and science disciplinary concepts. I accomplish this by creating new technologies to support students in building solar ovens while testing their design ideas in an interactive computer model that connects science concepts and design decisions.

When students build a physical model they may neglect the scientific basis for their decisions, instead focusing on details of construction that may be superficial rather than scientifically based. Educational tools, like interactive computer models, can help students connect science principles and design decisions by making mechanisms such as energy transformation visible. The NGSS envision that instruction would combine practices including modeling, data, analysis, computational thinking, and design to enable students to integrate their scientific and engineering ideas (NGSS Lead States, 2013). This research identifies optimal ways to integrate science and engineering practices by taking advantage of interactive models, automated guidance for student short essays, and supports for making evidence centered decisions. The investigations are guided by the knowledge integration theory and the results expand the theory into the engineering domain.

In this dissertation, I present five empirical chapters. Each study uses a solar ovens curriculum in which students use a virtual model to design and explore energy transformation, then build and test a physical solar oven. These studies investigate ways to support students in integrating their ideas about energy transformation with ideas about engineering design. The first empirical chapter investigates how computer models function in hands-on curriculum to aid in the knowledge integration process. The second and third empirical chapters investigate supports for students while they use computer models. These chapters document how students interact with the model. Because the computer model aids in both design and reflection, there are three chapters devoted to investigations of how the computer model aids students in knowledge integration. A fourth empirical chapter investigates the non-normative, yet common, idea that shiny or dark objects “attract” light to them, causing them to heat up. I first collect data about the ideas students present around this non-normative idea, then present a method to automatically score student written responses for the presence of this idea. This automatic scoring algorithm could support the development of automated guidance that could then encourage students to refine their ideas. The fifth empirical chapter investigates two ways to frame the curriculum. Since the goal of this curriculum is to integrate both science content ideas and engineering design ideas, I investigate two different frameworks for presenting the curriculum – science-centered or engineering-centered.

Together, these chapters suggest guidelines for the structure of hands-on projects that aim to teach both science concepts and engineering design. First, creating dynamic computer models that allow students to test their design ideas has proven useful in helping students integrate science disciplinary ideas and engineering practices. However, students need scaffolding to integrate these ideas and practices. To ensure that the virtual models inform student designs in a meaningful way (and vice versa), there should be careful consideration about when during the curriculum they are introduced.

Including science content in a meaningful way and supporting the integration of science ideas is also critical for the success of projects that are intended to support the integration of science and engineering. To help students make sense of key scientific phenomena, designers need to identify ideas that are challenging for students to distinguish among, like that of light propagation (e.g., is light reflected, absorbed, or “attracted”?). Creating opportunities for students to follow the knowledge integration process is important with these types of ideas, in order to give students the opportunity to integrate their disparate and perhaps contradictory ideas. Specifically, students need to generate multiple ideas so that those ideas can be inspected, added to through the use of inquiry activities, and then they can distinguish among their entire corpus of ideas. This process helps students to make sense of their ideas; the addition of an engineering project provides further evidence for students to reflect upon.

It is also important to consider the goals for learning when framing curriculum as either an engineering or a science project. Different ways of framing the same type of project may lead to different learning outcomes. If a project is framed around engineering design, students are likely to develop stronger engineering practices, but their understanding of scientific content may not be as deep. If a project is framed as a scientific investigation, students may integrate their science ideas, but not develop a strong sense of engineering practices.

Table of Contents

Acknowledgments

I would like to thank my advisor, Marcia Linn, for providing me with mentorship and support throughout my time at Berkeley. I have so appreciated the opportunities for learning and conducting new and innovative research, and the WISE community you have built at Berkeley. I would also like to thank the members of my dissertation committee. Michelle Wilkerson for her careful consideration of my work and excellent feedback, and Robert Rhew for his passions for both science and education.

I would also like to thank Jonathan Vitale for his guidance and support through every stage of my journey as a researcher. From pushing my thinking about what social science research is, to talking through new analysis methods, and most importantly, being a great writing partner and mentor - this process would have taken much longer and been much more painful without your immense knowledge and wit. I owe many thanks to past and present members of the WISE research group, particularly Libby Gerard, Jennifer King Chen, Charissa Tansomboon, Lauren Applebaum, Eliane Wiese, Dermot Donnelly, Rachel Jansen, Erin Galloway, Emily Harrison, Jacquie Madhok, Kyle Fricke, and Korah Wiley, for making this work possible through your collaborative efforts. I also owe many, many thanks to the WISE tech team for making everything work and supporting research so well – thank you Jonathan Lim Breitbart, Hiroki Terashima, and Geoffrey Kwan. Of course, many thanks are due to David Crowell for making sure I was able to do research and for providing the comical support necessary to see the finish line at the end of this whole process. I am also extremely grateful to the wonderful teachers who welcomed me into their classrooms. Thank you especially to Crystal Mosteiro, Ben Leech, Gabby Mastro, and Erin Galloway for being excellent partners in research.

I would also like to thank Sirpa Tuomainen (and other teachers and fellow students in the Finnish program), for making Berkeley feel more like a family. Sirpa, you have been the ultimate people connector, helping me find new opportunities for learning at Berkeley. Your enthusiasm is contagious, and you have been the ultimate example of being fierce yet kind.

To my friends and family, especially those in the Bay Area - thank you for your constant love and support during this journey. I could not have made it through this without your support! Especially Kim and David, Jenna, Beatriz, Allie, and Anna – thank you for listening when I needed support and pushing me when I needed a kick to keep going.

To my parents - thank you for instilling in me a love of learning that has lasted through ten years of higher education. Thank you for your love and support, your amazing example of combining your passions for social good, science, engineering, and education, and for coming to visit me to make sure I took time to enjoy life in the Bay Area. To my brothers – thanks for being my partners in crime for life, for keeping me grounded, and for going camping with me in all types of weather. Finally, to Will – your support has been the most important part of this journey. Thank you for always being there when I need someone to talk things out with, making sure I'm eating enough vegetables, and for pretending like you

know what I'm talking about when I get deep into the weeds of whatever problem I'm working through. Your curiosity and passion for life inspire me every day and I cannot thank you enough for your constant support.

Chapter 1: Introduction

This dissertation research examines how students learn from hands-on engineering projects in science classrooms. I conducted this research on a variety of aspects that are especially important in this context. I aimed to understand how students could use computer models to better understand science concepts while also meeting engineering design objectives, I examined students' understandings of energy transformation and developed automated scoring algorithms to assess specific non-normative ideas about energy, and I sought to develop an understanding of differences between the framing of hands-on science projects and engineering projects.

In an ideal world, engineering projects would also improve student achievement in mathematics and science. However, research on this topic has shown that many engineering projects do not live up to this ideal (Teacher Advisory Council, 2009). Ideally, undertaking a science project should be motivating, while also helping students to understand the interplay between science concepts (like energy transformation) and engineering design decisions (Hmelo et al., 2000; Cantrell et al., 2006). This dissertation research investigates ways to integrate engineering practices and science concepts (like energy transformation) in classroom settings. I investigate ways to integrate the Next Generation Science Standards (NGSS) science and engineering practices while simultaneously expanding the knowledge integration theory (Linn & Eylon, 2011). I refine knowledge integration design principles in classroom studies, comparing alternative forms of instruction where students integrate engineering design and science disciplinary concepts. I accomplish this by creating new technologies to support students in building solar ovens while testing their design ideas in an interactive computer model that connects science concepts and design decisions. I worked with teachers to develop curriculum on the solar ovens project, and I implement my research in a variety of diverse 6th grade classroom settings using the Web-based Inquiry Science Environment (WISE).

When students build a physical model they often neglect the scientific basis for their decisions (Crismond, 2001), instead focusing on details of construction that may be superficial rather than scientifically based. Interactive computer models can help students connect science principles and design decisions by making mechanisms such as energy transformation visible (Snir, Smith, & Grosslight, 1993; Wilensky & Reisman, 2006). The NGSS envision that instruction would combine practices including modeling, data, analysis, computational thinking, and design to enable students to integrate their scientific and engineering ideas (NGSS Lead States, 2013). This research identifies optimal ways to integrate science and engineering practices by taking advantage of interactive models, automated guidance for student short essays, and supports for making evidence centered decisions. The investigations are guided by the knowledge integration theory and the results expand the theory into the engineering domain.

This work investigates four research questions related to the integration of science content with engineering design projects and the knowledge integration framework. The studies presented in chapters 2-6 aim to improve our understanding of the design of engineering projects at the pre-college level, as well as provide guidance for how these projects can be carried out in classrooms. I carry out this work using comparison of experimental and control conditions, design-based research, and other methods from the fields of data science and learning analytics. I provide an outline of insights gained from each study at the end of this chapter.

Objectives & Research Questions

To identify optimal ways to integrate science and engineering practices and extend the knowledge integration framework, this dissertation research addresses the following research questions:

1. How do students use interactive computer models to integrate science and design during engineering projects?
2. What sources do students use as evidence for design decisions? How can we support students in making decisions based on scientific concepts or evidence?
3. What design principles guide student use of interactive tools (e.g., project report spaces, photos, notebooks, and automated guidance) to support integrated understanding?
4. In classroom instruction, what is an optimal balance of science concept development and engineering design activities to promote integrated understanding?

Theoretical Framework

Knowledge Integration

To design the instruction and assessment for these studies, I implement knowledge integration design principles because they emphasize the process of connecting design decisions and scientific principles (Kali, 2006; Linn & Eylon, 2011). The knowledge integration framework and principles have proven useful for design of instruction featuring virtual design activities (Chiu & Linn, 2011; McElhaney & Linn, 2011), as well as those featuring dynamic visualizations (Ryoo & Linn, 2012) and engineering design (Chiu et al., 2013; McElhaney & Linn, 2011).

The knowledge integration framework is a constructivist approach to instruction that emphasizes reflection on a student's repertoire of ideas, adding new scientific ideas, using evidence to distinguish accurate and relevant ideas, and forming links between ideas to explain a phenomenon (Linn & Eylon, 2011). Students often begin a classroom activity with preconceived notions, and in the process of learning they may develop multiple, often conflicting ideas (Smith, diSessa, & Roschelle, 1994). Indeed, Chapter 4 of this dissertation offers further support of this idea; students struggle to distinguish among attracting, reflecting, and absorbing, and often use some of these words interchangeably. Students typically respond to instruction by adding the new ideas to their multiple and often conflicting views (diSessa, 2006; Driver, Newton, & Osborne, 2000). The knowledge integration framework describes how students develop an integrated understanding of a domain by linking and connecting ideas. The framework calls for eliciting, adding, distinguishing, and sorting out ideas as they engage in challenging scientific activities (Linn & Eylon, 2011). The process of knowledge integration is iterative; thus, each step of curriculum designed using knowledge integration is meant to encourage the different steps in the knowledge integration process.

In this framework, instruction should first elicit ideas that students have about the topic they are studying. These ideas can be about the discipline and about related personal experiences, as some disciplinary ideas are likely to be bound in personal contexts. This step in knowledge integration helps to make sure new ideas are not isolated from prior knowledge. Prompting students for predictions about phenomena before they engage in investigations can aid in eliciting ideas and has been shown to improve learning outcomes (Crouch, Fagen, Callan, & Mazur, 2004; Linn & Songer, 1991).

Next, instruction in the knowledge integration framework should add normative ideas to students' repertoires. Traditionally this has been done using lectures and assigned texts. However, research indicates that there a number of other methods for adding ideas that may be more impactful for students. For example, using analogies (Coll, France, & Taylor, 2005; Clement, 1993) or pivotal cases (Linn, 2005) to present students with comparative situations or situations that are easier to understand or perhaps more relevant to a student's life.

Instruction must also help students to distinguish among their ideas. Once students have been asked about their existing ideas and have added new normative ideas, they are faced with a mix of new and old ideas that may not always be coherent. To distinguish among the various ideas they have, students are generally required to evaluate their ideas in order to integrate them into a coherent set. Pivotal cases (Linn, 2005) can help students distinguish their ideas by allowing them to make comparisons among their ideas. Critique activities can also help by encouraging students to distinguish between normative and non-normative explanations (Zhang, 2010; Sato, 2015).

Finally, instruction using the knowledge integration framework allows students to reflect upon their understanding. This allows students to both develop coherent and integrated ideas about phenomena and to identify inconsistencies or gaps in their ideas.

Literature Review

Engineering Projects in K-12 Education

For many years, there has been a growing interest in engineering at the precollege level, for example through programs like Engineering is Elementary (Rogers & Portmore, 2004). With the implementation of the NGSS, engineering is now included in science curricula more broadly (NGSS Lead States, 2013).

Engineering is a broad field that includes many topics, from design of systems, to user needs analysis and making data-driven decisions. In the NGSS, engineering is emphasized through the Science and Engineering Practices, a set of practices meant for students to develop over the course of the K-12 education. In these practices, engineering is specifically called out in two areas: defining problems (where the science counterpart is Asking Questions) and designing solutions (where the science counterpart is Constructing Explanations). This emphasis on design is consistent with the most commonly used format for engineering design projects: design, build, test, or DBT (Elger, Beyerlein & Budwig, 2000). Within this format, students first engage in the design aspect of the project, which may include a needs analysis and defining the problem they are trying to solve. Next, students build a prototype, and finally they test their prototype to find out what is working and what is not. This format should be iterative, with students using the results from their tests to redesign parts of their prototype. The design, build, test method can be taken to another level by including rapid prototyping, during which students design on a much shorter timeframe in order to build and test prototypes faster, thereby gaining actual data faster (Noorani, 2006).

These methods for conducting engineering projects differ from the scientific method, often used for hands-on science projects. Often the goal in science is to develop knowledge, while in engineering the goal is to develop a solution (Lewis, 2006; Purzer, et al., 2015). The scientific method teaches students about conducting science as a process, but can often be prescriptive and lead to classroom experiments that guide students' steps too much, distracting them from productive inquiry (Tang, Coffey, Elby & Levin, 2010). While the process, as outlined in many textbooks or instructional tools is one method students may use to investigate phenomena, it is not the only way that is useful.

One response to the prescriptive nature of the scientific method has been scientific inquiry learning, whereby students are able to explore phenomena without following a specific method (e.g., Krajcik, Blumenfeld, Marx & Soloway, 1998; Slotta & Linn, 2009). Using engineering design, students can gain many similar insights into scientific concepts, while also building design and analysis skills. An example of this is the Learning by Design approach to curriculum design within project-based learning (Kolodner, et al., 2003), wherein students learn both science concepts and skills in engineering design.

Often, in science the goal is to develop knowledge, while in engineering the goal is to develop a solution (Lewis, 2006; Purzer, et al., 2015). This tension carries over into instructional methods that differ between science-focused projects and engineering-focused projects (Lachapelle, et al., 2013).

Project-Based Learning

In developing our engineering projects, two features of project-based learning have proven useful: starting with a driving question (or problem), and engaging students in authentic, situated inquiry about that problem (Blumenfeld et al., 1991). Research on project-based learning reveals the advantage of engineering design to generate student interest and motivation in science topics (Hmelo et al., 2000; Cantrell et al., 2006). This type of learning can also be supported effectively using technology (Krajcik et al., 1998; Krajcik et al., 2000).

Due to the focus on engineering practices in the curriculum used in this research, we also draw on other principles from project-based learning, including the creation of artifacts, collaboration, and the use of technology tools to support learning (Krajcik & Blumenfeld, 2006). This work differs from project-based learning (Krajcik & Blumenfeld, 2006) due to the focus on engineering practices instead of on scientific concepts. The creation of artifacts in this work is a main part of the learning experience, as well as an external model of students' knowledge.

Learning by Design (Kolodner, et al., 2003) was also important in the development of the project-based learning presented in this research. In the Learning by Design framework, students are introduced to a design challenge and must work with their peers to design, test, and iterate on a solution to that challenge. Students engage with inquiry activities and scientific content on an as-needed basis. We build on this framework by incorporating more technology in the design process and allowing students to use online resources to find useful information. In Learning by Design, the teacher acts as a facilitator for students as they engage in research and design; this is similar to the role we expect teachers to take on in the Solar Ovens project. One marked difference is in the methodology students are encouraged to take; in Learning by Design, students are encouraged to use the scientific method, running controlled tests to determine the role of variables. We focus on an engineering method, in which students are encouraged to find solutions instead of producing knowledge. On the way to finding those solutions, students must gain new scientific knowledge for themselves, but the final goal of the project is on designing a solution to the problem, not producing new knowledge.

Using dynamic visualizations and interactive models to promote knowledge integration

Dynamic visualizations capture aspects of scientific phenomena that are difficult to describe using text or static images. While students often regard science as a series of unrelated facts, in most cases scientific content is characterized by complex systems of inter-related components and unseen dynamic processes (Hmelo-Silver, Marathe, & Liu, 2007; Linn & Eylon, 2011). These characteristics make complex systems particularly challenging to understand (Hmelo-Silver & Azevedo, 2006) and an important target for scaffolding (Linn et al., 2014). Often, students are unsure about how to conduct an experiment using a simulation, and engage with models in an unsystematic way, even when asked to articulate predictions before engaging with the model (McElhaney & Linn, 2011). Visualizations can overwhelm students with the amount and complexity of information presented, hindering them from utilizing the visualizations for learning. Dynamic visualizations of complex processes, like energy transformation or climate science, can also give students the illusion of understanding and therefore discouraging them from engaging in distinguishing among their ideas (Chiu & Linn, 2008).

Even with all of these potential drawbacks, research on dynamic visualizations has shown that they are a definite improvement on static visualizations for communicating complex science concepts (McElhaney, et al., 2015). In a review of dynamic visualizations research, McElhaney (et al., 2015) found that the most impactful scaffolds for helping students realize the potential of dynamic visualizations in learning were using prompts for reflection, prompts to distinguish among parts of the visualization, visual cues that identify salient features, using multiple visualizations presented sequentially, and using interactive features that govern the pacing of activities. Recent research offers similar findings and promising ways to guide student interactions with complex dynamic visualizations (Krajcik, Blumenfeld, Marx & Soloway, 2000; Ryoo & Linn, 2014; Wilkensky & Reisman, 2006).

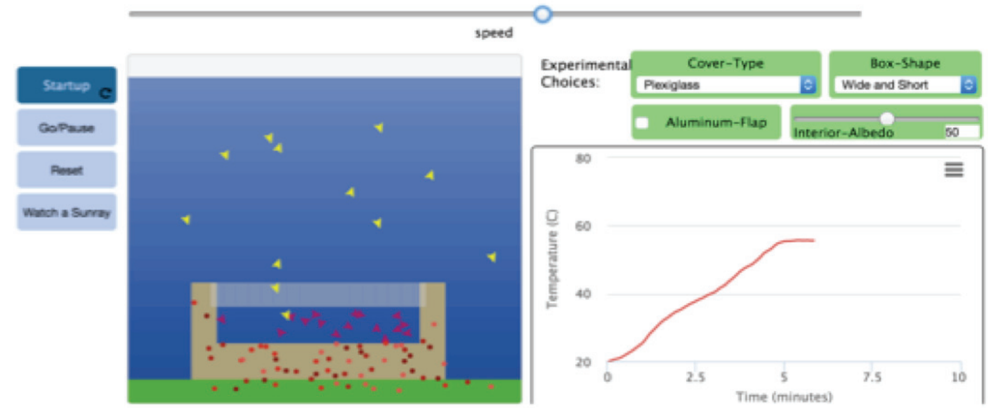
Much research in the area of modeling and dynamic visualizations is done in the area of complex systems, where these models can help students make sense of the world (e.g., Wilensky & Resnick, 1999). Initially, many modeling environments built on science models used for advanced research, and were not always easy for students to make sense of. However, now that visualizations and models have become more ubiquitous and interfaces have become easier to use, they are being used in new instructional ways, such as for engineering education (McElhaney & Linn, 2011; Dym, et al., 2005) and for conducting laboratories (De Jong, Linn, & Zacharia, 2013).

Solar Ovens Curriculum & Assessment

All but one of the studies presented in this dissertation use the solar ovens curriculum. The solar ovens curriculum familiarizes students with the way energy transforms from solar radiation to heat (MS-PS3-3) using a hands-on project and interactive models, emphasizing the modeling aspect of the science and engineering practices of the NGSS, as well as the standards associated with energy (NGSS Lead States, 2013). This curriculum draws on all eight of the science and engineering practices in the NGSS, focusing on the using models, developing solutions, and engaging in argument from evidence. Students engage with the curriculum through WISE (Web-based Inquiry Science Environment), which allows for the utilization of a variety of instructional and assessment tools (Linn & Eylon, 2011). The curriculum was developed as a collaborative effort between teachers and researchers, making sure teachers were comfortable with all activities and revising based on classroom implementation.

In the curriculum, students follow the design, build, test cycle for two iterations. Throughout the curriculum, students use interactive simulations to design their solar ovens (Figure 1.1). These simulations provide students with a framework for thinking about how, when, and where reflection and absorption happen in their ovens. The curriculum takes 10-15 classroom hours to implement, and students work in groups of 2-3 during the curriculum. In experimental comparison studies, WISE automatically assigns each working group to one of the conditions during instruction.

The solar ovens project implements the knowledge integration framework across the whole project. Each group is frequently prompted to make predictions (eliciting) about what will happen, then conduct a test using a simulation or their physical project (adding). The results of these tests or trials allow students to consider the ideas that they implemented in the project and make appropriate revisions (distinguishing). At the end of the project, students are asked to reflect on how their oven works and why it works the way it does (reflecting). In addition to this project-level implementation of the knowledge integration framework, each step uses the framework to guide students through, for example, the use of simulations or making evidence-based revisions during their second iteration of design.



#	Box Shape	Cover Type	Flap?	Int Albedo	Temp at 0m	Temp at 1m	Temp at 5m	Temp at 10m
1	Wide and S	None	No	50	20	24	30.2	31.5
2	Skinny and	None	No	50	20	26.2	29.7	31.1
3	Wide and S	Plexiglass	No	50	20	25.6	55.4	53.4

Figure 1.1: Interactive simulation used by students to design their solar ovens

The research presented in this dissertation uses pre- and posttests to determine the efficacy of the curriculum and of conditions within experiments. Students work individually on all pre- and posttests. For this curriculum, the pre- and posttests are made up of the same questions; there are generally five questions, some with multiple parts. Some questions measure student ability to integrate science concepts, others measure ability to integrate science concepts with engineering design decisions. To measure this integration, items are scored using knowledge integration rubrics to assess the links between multiple normative ideas (Linn & Eylon, 2011; Liu et al, 2008). Knowledge integration rubrics for a sample questions are shown in Table 1.1 and Figure 1.2. Knowledge integration rubrics capture the number and quality of normative conceptual links between scientific ideas. In order to adapt these rubrics to the engineering design arena, we evaluate students responses based on the number of links between the design description and the scientific basis for the decision. Adapting knowledge integration rubrics to fit engineering questions has been a relatively simple process, as the connections between engineering and science being evaluated lend themselves well to knowledge integration. However, developing appropriate engineering design questions has been a more challenging task.

Table 1.1: Scoring rubric for the *David's Claim* item, which asks students to explain to a fictional student (David) whether he should use a tall and skinny box or a short and wide box based on the results from a computer model

Score	Level	Examples
1	**Off Task**	*I don't know.*
2	**Irrelevant/Incorrect**	*David is correct because I chose the skinny and tall one and the heat went up really fast.*
3	**Partial** Normative isolated ideas without a valid link	*David's claim is not correct because in the model it show solar radiation stayed trapped inside the wide and short one making heat easily trapped inside.*
4	**Basic** Elaborate a scientifically valid link	*David's claim is incorrect because the skinny box got to 33.8 in 2 minutes and the wide box got to 44.7 in 2 minutes. The wider box could keep a lot of energy because of the space and the skinny box doesn't have a lot of space. So, this means David was wrong.*
5	**Complex** Elaborate two or more scientifically valid links	*David's claim is incorrect because the more area for radation to come the more radiation can get trapped and turn into heat.there is less of the when you have a skiny box.*

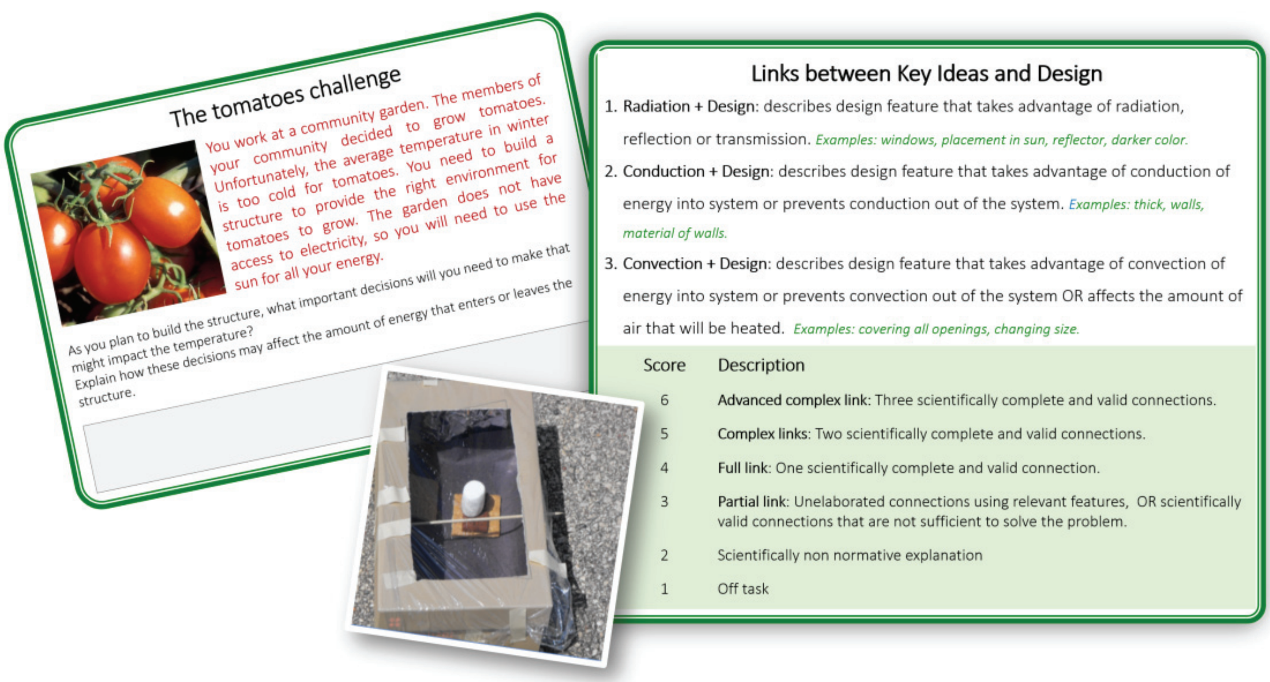


Figure 1.2: Tomatoes item to assess integration of design ideas with science concepts on pre- and posttests.

Dissertation Outline

This dissertation is written such that each chapter can stand alone, therefore there may be some repitition among chapters. This will be most apparent in the methods and curriculum sections of each chapter. Each of these empirical chapters is in preparation for journal submission. Table 1.2 provides an outline of each chapter, along with the key findings from that chapter. Each chapter does rely on independently conducted studies, so the data collected for each chapter is unique to that study.

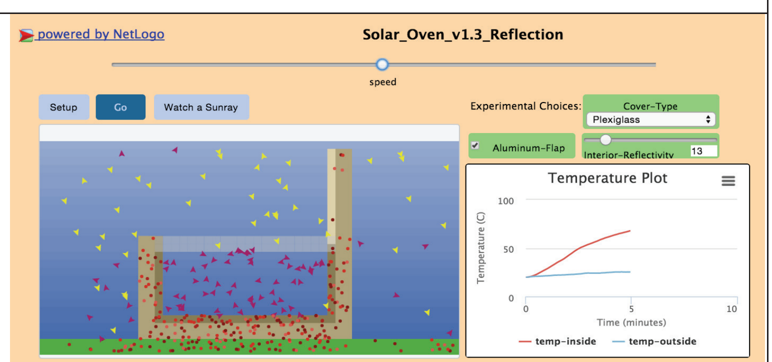
Table 1.2: Outline of key findings and methodologies used in each chapter

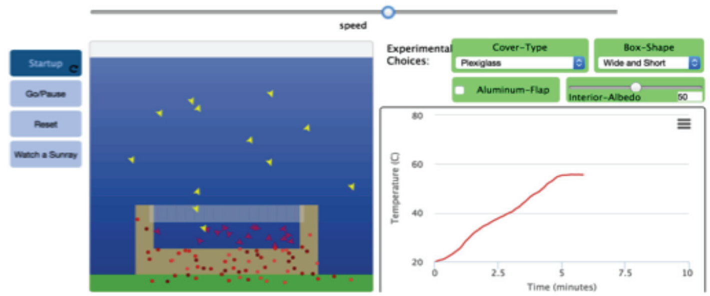
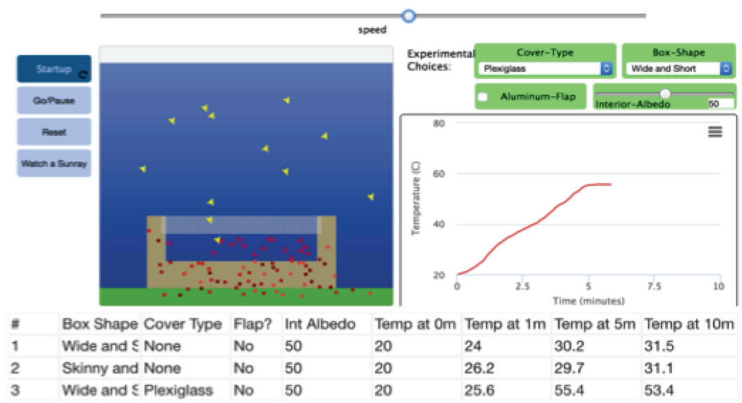
Chapter	Key Findings	Methodologies Used
2	• Students integrate science ideas better when they use a virtual model to *plan* their designs because they add more ideas earlier in the project • Using a virtual redesign activity is just as effective for science learning as doing a physical redesign and rebuild activity	Two separate comparison studies: • Use of the virtual model to *plan* or to *reflect* on the solar oven design • Second iteration of redesign done using *physical* or *virtual* activity
3	• Students who spend more time observing the virtual model (instead of clicking) made greater learning gains. We measure this by counting period of student observation lasting 15 seconds or longer. • We added scaffolds, like an automatically generated table and reflection questions to the virtual model • Students who conduct more trials in the model make greater learning gains; using the control of variables strategy does not improve learning gains	Log data analysis from virtual models in a climate change unit and the solar ovens unit, analysis of pre/posttest learning gains
4	• Students often do not trust the results from computer models, prefering results from physical testing • Students come to see the value in features of the virtual model by the end of a curriculum that uses it (e.g., speed of testing, generated graphs, visualization of energy transformation)	Pre/posttest analysis of one question on student opinions in conjunction with pre/post learning gains

5	• Students have inconsistent ideas about light propagation, thinking that light is “attracted” instead of or before being reflected by shiny objects • Developing automated guidance for student writing on this specific idea is possible with good accuracy, but other instructional tools are also necessary to help students add enough new ideas and falsify ideas about “attraction”	Analysis of embedded questions and development of an algorithm for automated guidance using natural language processing and machine learning
6	• Students learn material differently depending on how the project is framed: ○ The *engineering* condition made greater gains in developing engineering practices ○ The *science* condition made greater gains in integration of science ideas	Comparison study: • Curriculum framed as an *engineering* or *science* project

Further, many of these studies rely on design iterations of a virtual model. Table 1.3 outlines how the design of the virtual model changed throughout these iterations, and which version was used for each study. Chapters 4, 5, and 6 use the same version of the solar ovens virtual model designed for chapter 3.

Table 1.3: Description of changes made to virtual model

Chapter	Description	Screenshot
2, study 1	• Used in WISE software version 4 • Virtual model designed in NetLogo • Graph shows temperature inside and outside of oven • Experimental Options: Cover Type, Albedo (slider), Aluminum Flap	

2, study 2	• Used in WISE software version 5 • Virtual model redesigned in HTML using NetLogo • Graph shows only temperature inside oven • Experimental Options: Same as Ch. 2, Study 1 plus option for box shape/size	
3	• Same as version used in Ch. 2, study 2 • Added automatically generated table: rows appear after students let the trial run until the graph reaches 10 simulated minutes	

Conclusions

These studies investigate ways to support students in integrating their ideas about energy transformation with ideas about engineering design. The first empirical chapter investigates how computer models function in hands-on curriculum to aid in the knowledge integration process. The second and third empirical chapters investigate supports for students while they use computer models, and how exactly students are interacting with the model. Because the computer model aids in both design and reflection, there are three chapters devoted to investigations of how the computer model aids students in knowledge integration. A fourth empirical chapter investigates the non-normative, yet common, idea that shiny or dark objects "attract" light to them, causing them to heat up. I first collect data about the ideas students present around this non-normative idea, then present a method to automatically score student written responses for the presence of this idea. This automatic scoring algorithm would allow for the development of automated guidance that could then encourage students with this non-normative idea to reconsider what they have written. The fifth empirical chapter investigates two ways to frame the curriculum. Since the goals of this curriculum are to integrate both science content ideas and engineering design ideas, I investigate two different frameworks for presenting the curriculum – science-centered or engineering-centered.

Together, these chapters lead to specific findings about the structure of hands-on projects that aim to teach both science content and engineering design. Using computer models has proven useful in helping students integrate science and engineering. However, these models require scaffolding to help students integrate their ideas and there should be careful consideration about when during the curriculum they are introduced. In addition, it is critical to encourage students to address and sort out the non-normative ideas they may have around given science concepts, instead of focusing all the student effort on the engineering aspects of a project. Including science content in a meaningful way is critical for the success of hands-on projects, but instructional designers need to anticipate student dilemmas and design adequate supports. In addition, different ways of framing the same project may lead to different types of learning gains. If students are introduced to the project as an engineering project, they will make greater gains in engineering design skills than if they are introduced to the project as a science project with a hands-on component. It is important to consider the goals for learning when considering how to structure the curriculum for hands-on projects integrating science and engineering.

Chapter 2: Use of Interactive Computer Models for Design

This chapter investigates three specific uses for computer models within hands-on curricula, and the ways these models can be used to connect science concepts with engineering design. We investigate the use of a computer model for planning of a design, for reflecting upon the results of testing a physical model, and for making revisions to a design for a solar oven. Two studies are presented to investigate how the computer model functions at different locations within the curriculum. This first study investigates how the model functions as either a tool for planning during design or a tool for reflection upon the results of testing a solar oven. Each of these tasks is valuable in knowledge integration, but the tasks are quite different. We find that students interact with the model more at the planning stage, and thus produce significantly greater learning gains.

A second study investigates whether students achieve comparable learning gains by using the computer model as a revision activity as opposed to revising a design, then rebuilding and testing a physical solar oven. From this study we find that the model functions as well as a physical rebuilding cycle, while being far more time-efficient.

Together, these two studies reveal important uses for the computer model within the hands-on Solar Ovens curriculum. Each of these studies also evaluates student learning of science concepts and engineering practices. Later chapters investigate student actions within the computer model and how those actions may lead to greater integration of science content and engineering practices.

Introduction

It is often claimed that engineering projects improve student achievement in mathematics and science, but research on this topic has shown that many projects do not live up to the claim (Teacher Advisory Council, 2009). While hands-on projects may generate more student interest and engagement (Hmelo et al., 2000; Cantrell et al., 2006) than typical science curricula, they often fall short on developing science concepts. Ideally, undertaking a science project should be motivating, while also helping students to understand the interplay between science concepts (like energy transformation) and engineering design decisions.

Interactive computer models can help students explore alternative design decisions and illuminate difficult energy concepts. In this research, we study an interactive computer model designed to help students connect science concepts and design decisions while carrying out a hands-on design project in a classroom setting. Often when students build a physical model they neglect the scientific basis for their decisions (Crismond, 2001), instead focusing on details of construction that may be superficial to the learning goal. Consistent with the Next Generation Science Standards (NGSS), students use a computer model that helps them connect science concepts to their design decisions by focusing on critical variables and using engineering practices (NGSS Lead States, 2013). Interactive computer models can help students connect science principles and design decisions by making mechanisms, such as energy transformation, visible (Snir, Smith, & Grosslight, 1993; Wilensky & Reisman, 2006).

The model used in this research shows the results for each design in a graph of temperature vs. time, a potential method for capturing unobservable processes (Wilkerson-Jerde & Wilensky, 2015). Research also demonstrates that students need scaffolding to utilize the advantages of models. In a review of dynamic visualizations research, McElhaney (et al., 2015) found that the most impactful scaffolds for helping students realize the potential of dynamic visualizations in learning were using prompts for reflection, prompts to distinguish among parts of the visualization, visual cues that identify salient features, using multiple visualizations presented sequentially, and using interactive features that govern the pacing of activities. We selected some of these features to use as scaffolding in our simulation based on what made sense with the content and structure of the model. The model itself is also a scaffold for learning about making engineering design decisions that are based on scientific evidence. In this study, we scaffold the use of the model mainly through the use of guiding questions. The interactivity within the model also uses many scaffolding principles (McElhaney, et al., 2015) to guide student use. Further development of scaffolding for student use of the model is also discussed in later chapters.

The computer model plays an important role in linking science concepts with the design process because students are able to manipulate design alternatives while seeing how their choices impact energy flow. The process of integrating science concepts with engineering design is an important, yet challenging, part of developing projects (Johnson et al, 2015).

We used the knowledge integration framework to create a unit about solar ovens, because the framework focuses on building coherent understanding (Linn & Eylon, 2011). The framework offers instructional design principles to enhance connections between design decisions and scientific principles. The knowledge integration framework has proven useful for design of instruction featuring dynamic visualizations (Ryoo & Linn, 2012) and engineering design (Chiu et al., 2013; McElhaney & Linn, 2011). The framework emphasizes linking of ideas by eliciting all the ideas students think are important and engaging them in testing and refining their ideas. When students build a physical artifact they can often only test a few of their ideas due to time and material constraints. Modeling allows students to explore many more ideas. In addition, the model offers suggestions for tests to conduct, in the form of the pre-programmed options for students to select from, while providing a visualization of the mechanism of energy transformation. Using the model as a way for students to add new ideas to their repertoire before deciding on a design for their solar oven is also important, as this functions similarly to brainstorming. If students do not engage in this idea-adding, they may stop adding new ideas once they have settled on a design, instead focusing on the building aspect of the project. The process of knowledge integration is iterative and each step of the curriculum is designed to encourage a different step in the process.

In this unit, students use models much as scientists and engineers use models to rapidly explore diverse hypotheses about a complex system. Students also use the model to practice designing their solar ovens using scientific evidence for their design decisions. We implement the engineering design process using three stages – designing, building, and testing (DBT). Students learn how to improve their solar oven by analyzing the results of testing and using the results to inform the next iteration of their design. In this unit, students are introduced to the DBT process and to other features of engineering design, including using constraints and specifications. Students were required to use a budget while designing and building their ovens (Figure 1). The budget was small enough that students were required to make deliberate choices in which materials they used and how they used them, with the goal being that students would consider science mechanisms when making their choices. Students conducted two iterations of designing, building, and testing their solar oven.

The solar oven engineering design unit incorporates ideas from project-based learning and combines them with the knowledge integration framework. It offers students the opportunity to construct their own understanding of a phenomenon while working with and using ideas (Krajcik & Blumenfeld, 2006). This work draws on principles from project-based learning, including the creation of artifacts, collaboration, and the use of technology tools to support learning (Krajcik & Blumenfeld, 2006), while using the knowledge integration framework for the overall structure of the curriculum and the development of questions. Within the curriculum, each component is linked to one of the tenets of knowledge integration: eliciting ideas, adding normative ideas, distinguishing ideas, and reflecting. It also enables students to develop an understanding of science concepts through the design and development of product (Silk et al, 2009). It offers the opportunity for students to explore a problem that has no single correct solution, giving students agency over their own learning (Hmelo-Silver, 2004).

In the two studies presented in this paper, we discuss the interactive computer model used in the solar ovens curriculum. This model could be used for a number of different steps in the knowledge integration process: students can use the model to add and test new ideas to their repertoire, to test ideas and distinguish which ideas fit together, or even to reflect on how their solar ovens design performed. We present two experimental studies in which we investigate how students use the computer model at different stages in the knowledge integration process.

Each of these studies aims to understand how interactive computer models can support students during the knowledge integration process in a hands-on design project. The first study investigates how students use the model for either *planning* their designs or for *reflecting* upon how their solar ovens performed. We find that students who use the model for *planning* their design are more likely to run more trials and test more ideas in the model. This leads to greater learning gains between pretest and posttest. In the second study, all students use the computer model at the *planning* phase. In the second study we investigate use of the computer model as a way to guide the redesign process. Students either redesign, rebuild, and retest their *physical* solar oven during a second iteration of the design, build, test process or they use the computer *model* to redesign and test their ideas. We find no overall difference between these two conditions. While students may be more motivated to rebuild a physical solar oven, using the computer model to lead the redesign process is much more efficient when considering time and materials, and students make similar learning gains.

General Methods

In the two studies described below all instructional materials were presented in the Web-based Inquiry Science Environment (WISE), which is an open-source, online platform for developing inquiry materials and assessments (Linn, Clark, & Slotta, 2003). Each of the two studies are comparison studies; each uses two conditions to examine differences in student use of an interactive computer model.

Curriculum

Each study was implemented in a curriculum module entitled *Solar Ovens and Solar Radiation* (referred to as *Solar Ovens* in this paper). The goal of the unit is to familiarize students with the way energy transforms from solar radiation to heat through a hands-on project that utilizes interactive models. The curriculum covers the modeling aspect of the science and engineering practices of the NGSS, as well as the standards associated with energy, specifically standards related to the transfer of thermal energy (NGSS Lead States, 2013). The curriculum utilizes a variety of instructional and assessment tools (Linn & Eylon, 2011). WISE allows for assessment items to be embedded within the curriculum and for student pairs to be randomly assigned to condition.

The solar ovens curriculum within WISE has been designed and refined with the collaboration of multiple expert teachers and researchers to help students test and refine their ideas about energy transformation. The curriculum seeks to help students utilize their ideas about how radiation works in various contexts, like in the atmosphere and inside solar ovens.

Within the curriculum, students follow the design, build, test cycle with two iterations (Figure 2.1). During each design phase, students use the interactive model to test different features on a virtual solar oven. This model is shown in Figure 2.3.

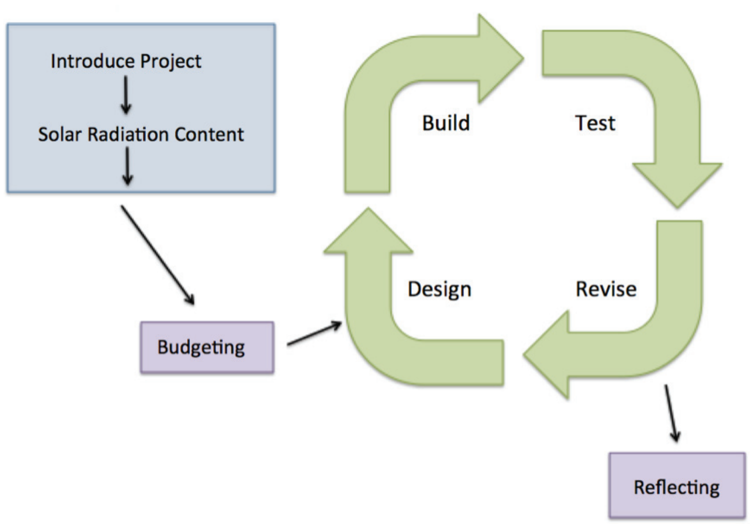
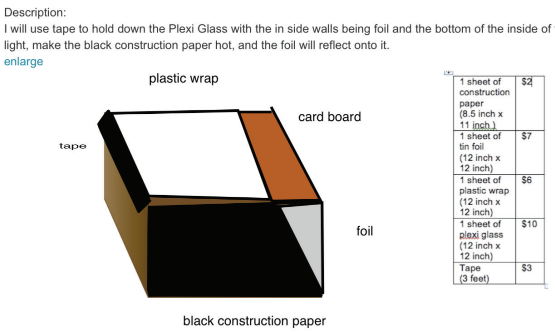


Figure 2.1*:* The outline of the solar ovens curriculum. Students go through the "design, build, test" cycle two times.



Materials & Costs

1 sheet of construction paper (8.5 inch x 11 inch)	$2
12 inch x 12 inch Sheet of tin foil	$7
12 inch x 12 inch Sheet of plastic wrap	$6
You can RENT: 12 inch x 12 inch sheet of Plexiglas (thick plastic)	$10
3 feet of tape (Duct, masking, clear)	$3

Figure 2.2: Student budget (left) and example of student design (right); students were given $20 for their first design iteration and $13 to add to their oven for the second iteration

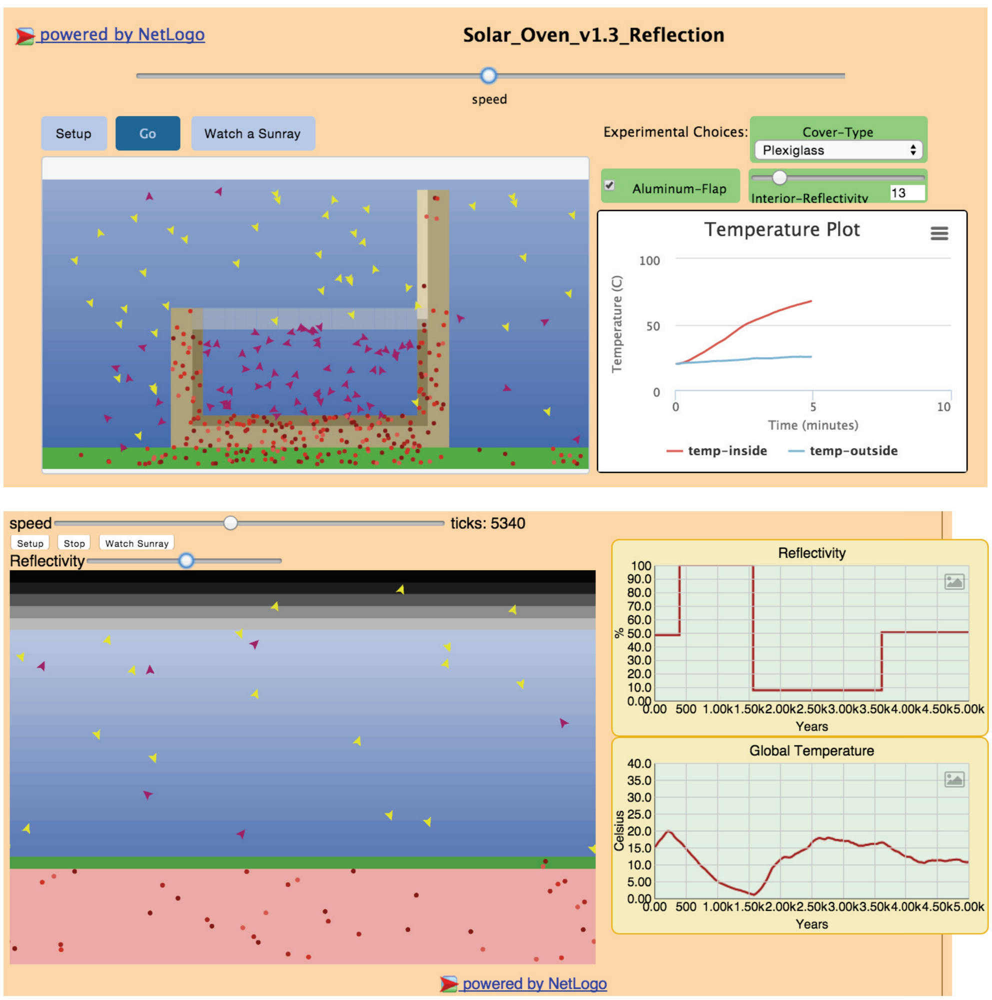


Figure 2.3: Screenshot of solar oven model (top) and solar radiation and the atmosphere model (bottom)

The solar oven model, designed using NetLogo (Wilensky, 1999) is similar to the models of radiation in the atmosphere students interacted with earlier in the unit, as shown in Figure 2.3. Each time students view a model, they make a prediction about the model or what causal relationship will be present (e.g. Greenhouse gases increase global temperature, or putting tin foil inside a solar oven will increase the temperature), and then interact with the model to test their prediction, and finally write about whether their prediction was correct or incorrect and why. Each model is associated with a graph that is generated dynamically in order to give students a way to connect energy mechanisms with the big-picture idea of the temperature inside the oven.

Assessment

To measure knowledge integration, the items were scored using knowledge integration rubrics to assess links between multiple normative science ideas (Linn & Eylon, 2011; Liu et al, 2008). The knowledge integration rubric for *Cars* shows how links are scored (Table 2.1).

Table 2.1: KI scoring rubric for "Car on a Cold Day" pre/post open response item

Score	**Level**	**Examples**
0	**No Answer**	
1	**Off Task**	*I don't know.*
2	**Irrelevant/Incorrect**	*The inside air and the outside air are the exact same temperature because the sun is not enough to heat the inside if the car.*
3	**Partial** Normative isolated ideas without a valid link	*The solar radiation would go through the metal and would stay in the car when the outside air wouldn't be able to get inside.*
4	**Basic** Elaborate a scientifically valid link	*it would be warmer than the outside air because if the car hasn't been driven for a week and its been in the sun the whole time the car will absorbe the heat and scence there is know way the heat can get out of the car the heat will just keep building up.*
5	**Complex** Elaborate two or more scientifically valid links	*The sun produces solar radiation which heats up the car and the infrared radiation gets trapped in the car which leads to the temperature rising.*

The pre- and posttest assessments measure student ability to link concepts about energy. Typically the items offer a choice among options and ask for a written explanation of the choice, consistent with the knowledge integration emphasis on linking ideas. For example the, *Car on a Cold Day (Car)* item prompted students to explain what would happen to a car left in the sun during a cold day. In another item, *Laura's Car,* students are prompted to write about what color interior and exterior Laura should have on her car in order to keep it the coolest on a sunny day (*Laura1*), and to explain whether or not Laura should use a sun shield to keep her car cool (*Laura2).* In another set of items, students are shown two pictures, of greenhouse gases in the atmosphere and of a greenhouse structure, and asked to compare them *(GHG1)*, then asked to compare the atmosphere and a greenhouse with a solar oven in *Greenhouse Gases 2 (GHG2)*. One item, *Model,* asked students to use a basic solar oven model to answer help a fictional student determine whether a tall, skinny box or a short, wide box would heat up faster. The pretest is given the day before the curriculum, and the posttest on the day following the curriculum.

Study 1: Knowledge Integration Design – Planning vs. Reflecting

Besides testing the overall advantage of modeling for knowledge integration, we also investigate whether it is more effective to use modeling to connect design decisions and principles prior to building a physical model or following the model construction and testing. Modeling before building the physical oven could help students distinguish among alternatives such as whether to line the inside of the solar oven with black paper or with foil. Modeling after building a physical model could enable students to test conjectures that arose during the construction of the oven.

In this study, we explore student use of the interactive computer model for two different purposes: *planning* or *reflecting*. In the *planning* condition, students were asked to compare three or more designs for their oven by using the model. They could link their designs to energy principles such as transfer and transformation. In the *reflecting* condition, after they design, build, and test their own solar ovens students also compare three or more designs typically including their own design. As in planning, they can distinguish between the designs using energy principles.

Each of these conditions is important within the knowledge integration framework (Linn & Eylon, 2011). During the *planning* phase of the project, students are adding their ideas and potentially testing and distinguishing those ideas. During the *reflecting* phase of the project, students are also distinguishing their ideas to understand which worked and which didn't. This study is designed to understand how students are supported in different stages of the knowledge integration process by the computer model.

Methods

Participants and procedures

Two teachers from one middle school serving a diverse population (42% reduced lunch, 13% ELL) chose to participate in this study. A total of 267 sixth grade students participated in some part of this study. Out of these students, 252 students completed a pretest, (some part of) the curriculum unit, and a posttest. The pretest was conducted one day before beginning the unit, and the posttest was conducted one day after finishing the unit. Both the pretest and posttest were administered to students individually. Pairs, or in some cases triads, of students were assigned to collaborative workgroups by their teacher to work on curriculum. Workgroups were randomly assigned to a condition (*planning* or *reflecting*) by the software and received the same activities in different orders. All students received the same curriculum curricular content, but activity order varied by condition.

Curricular materials

Students used the curriculum as described earlier in the general methods section. Table 2.2 outlines the differences between the *planning* and *reflecting* conditions in study 1.

Table 2.2: *Solar Ovens* Curriculum. Students used the model EITHER for planning or for reflecting.

Activity	**Topics, Resources, and Assessments**
Introduction to Solar Ovens	Elicit initial student ideas about energy transformation
Solar Radiation and the atmosphere	Energy comes as radiation from the sun; energy can be absorbed or reflected. Students use a virtual model to investigate energy (Figure 2.3).
Solar Radiation and Greenhouse Gases (GHGs)	Describes how energy interacts with greenhouse gases. Students use a virtual model to investigate how addition of GHGs impacts energy.
Model for planning condition	**Students use an interactive model (Figure 2.3) to investigate how solar radiation works to heat a solar oven [*Trials* item]**
Design, Build, Test 1	Design oven under budgetary constraints using a draw tool, build, test under a heat lamp using a temperature probe to collect data
Design, Build, Test 2	Students reflect on what was learned from the first iteration; use new budget constraints to repeat process [*Learn* item]
Model for reflecting condition	**Students use an interactive model to investigate how radiation works in a solar oven [*Trials* item]**
Reflect	Students describe how their solar ovens work using energy from the sun; make connections between solar ovens and the atmosphere [*Atmosphere* item]

In the context of this study we highlight those steps after the conditions diverge, specifically the embedded *Trials, Learn,* and *Atmosphere* items. In the *Trials* item students were asked to run at least three trials on the solar oven model, then write about what settings they used, how hot the oven got, and how long it took for the oven to get that hot. In the *Learn* item, which occurred between DBT iterations, students were asked what they learned from their first trial and how they will improve their design during the second iteration based on what they learned. In the *Atmosphere* item, students were asked to compare and contrast how radiation works in the atmosphere and in a solar oven.

Test materials

The pretest is made up of the *Car, Laura1, Laura2, GHG1,* and *GHG2* items, described in the general methods section. The posttest is made up of these same items with the addition of the *Model* item (also described in the general methods section). We analyze the embedded items *Trials, Learn,* and *Atmosphere*, which are unique to study 1.

Assessment

Items embedded in the unit (*Trials, Learn, Atmosphere*) were also scored using rubrics. The *Learn* and *Atmosphere* items were scored using knowledge integration rubrics (on a scale from 1-5), while *Trials* was evaluated using an adjusted knowledge integration rubric, since the question asked students to write about the trials they completed using the model instead of for an explanation, as in a typical knowledge integration question. *Trials* was evaluated for use of numerical evidence, mention of energy transformation mechanisms, and completion of *Trials* (rather than simply a description of the oven design), with students earning one point for each piece included in their response for a possible total of three points. The pre/posttest items were scored using knowledge integration rubrics, as described in the general methods section. The total pre/posttest score (adding knowledge integration scores for each item together) is used for analysis.

Results & Discussion

Students in the *planning* condition outperformed students in the *reflecting* condition on posttest [*planning:* M=15.54, SD=0.32; *reflecting:* M=14.83, SD=0.26]. A t-test of pooled pre- and posttest data across conditions (*Car, Laura1, Laura2, GHG1, GHG2* items) revealed a significant effect of testing session [$t(473) = -5.81$, $p < 0.001$], demonstrating that across both conditions students made gains from pre- to posttest. A regression model showed that posttest scores were significantly influenced by condition when controlling for pretest scores [$F(247) = 27.11$, $p=0.05$] (Figure 2.4), suggesting a benefit from interacting with the model for planning. This may be because students considered more ideas for their solar ovens design if they used the model for planning, rather than using the model for reflecting and using the model to retroactively consider ideas for the solar oven design.

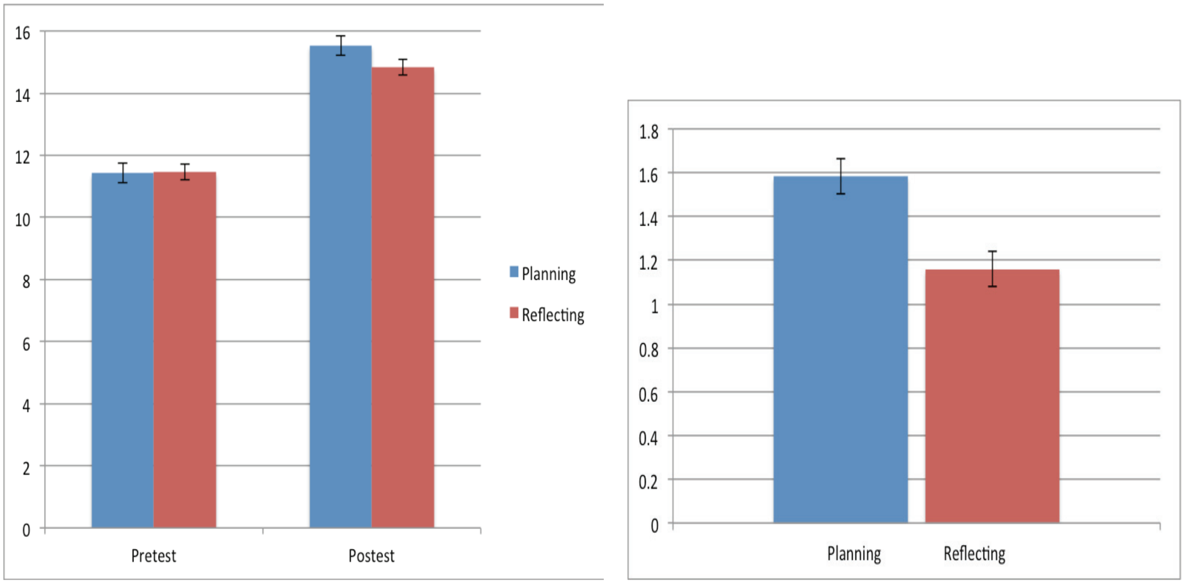


Figure 2.4: (left) Mean scores for pretest and posttest, by condition; difference between conditions at posttest shows significance [p < 0.05]; (right) Average scores on *Trials* item, by condition [p < 0.001]

These results suggest that both conditions were effective in terms of promoting student understanding of how energy transforms, but that interacting with the solar oven simulation before DBT might be most beneficial, if time constraints allow for only one modeling phase. Additional support for this claim comes from the *Model* posttest item, which involved writing an argument based on findings from a computer model of a solar oven. Students in the *planning* condition performed slightly better than students in the *reflecting* condition [$t(240) = 1.88$, $p < 0.06$]. This is surprising since students in the *reflecting* condition would have the added benefit of interacting with a similar computer model recently (during the end of the unit), while approximately one week had elapsed for students in the *planning* condition since they interacted with the embedded model.

Students in the *planning* condition also performed better on scored items embedded in the unit [*Learn:* $t(248) = 2.43$, $p < 0.02$; *Atmosphere:* $t(248) = 1.83$, $p < 0.06$; *Trials*: $t(248) = 4.10$, $p < 0.001$]. Higher scores on the *Learn* and *Atmosphere* items, which were scored using a knowledge integration rubric, indicate that students were able to add more normative ideas and connect their ideas together. Higher scores on the *Trials* item (Figure 2) indicates students used the model to run trials of their existing or future ovens, wrote about the results using numerical values, and connected the energy concepts with their design choices. These results from the embedded items support the idea that using the model for planning broadened the ideas students considered while designing their solar ovens. Using the model for planning helped students carry more ideas through the unit, giving them more ideas to distinguish among on reflection questions, like the embedded *Learn* and *Atmosphere* items.

Out of 137 students in the *reflecting* condition, 81 (59%) scored a 0 or 1 on the *trials* item, while only 43 out of 113 students (38%) in the *planning* condition scored 0 or 1. A regression model showed that posttest score across conditions was by influenced by score on the *trials* item, when controlling for pretest score [$F(247) = 30.57$, $p < 0.02$], indicating that students who scored higher on the *trials* item were more likely to score highly on the posttest.

Case studies

To illustrate how students responded to the unit, we describe case studies of 3 students. Cases were selected to show typical interactions with the embedded computer model. Student A was in the *planning* condition, while students B and C were in the *reflecting* condition.

Student A

Student A used the computer model during the *planning* phase of the project to conduct exploratory trials, experimenting with different materials and noting the temperatures of each trial. Student A did not explicitly connect science concepts to her writing about the trials, which was common for students in the *planning* condition. Below are the notes from student A (and her partner) about their interactions with the model.

> ***Trial 1****: We used plastic wrap and an aluminum flap high reflectivity and the temperature was at 22, and it took around 2 minutes*
> ***Trial 2:*** *We used aluminum foil, and a reflectivity of 80. The highest temperature was 23 and it took about 2.5 minutes.*
> ***Trial 3:*** *We used Plexiglas, an aluminum flap, and a reflectivity of about 30. The highest temperature was about 43 and it took about 2.5 minutes to get that high.*

Based on the notes student A wrote, we can see that she was exploring the use of different design features. These trial notes do not indicate a systematic changing or comparison of one variable (e.g. the covering of the solar oven), but instead student A changes multiple features between each trial. This use of the model is not systematic, but exploratory in nature, as we would expect it to be.

Student A scored a 9 on the pretest, which is in the 15th percentile for all pretest scores. However, by the posttest, student A increased her score to a 17, which is in the 70th percentile for all posttest scores. Student A made great gains between the pre- and posttest, possibly aided by her interaction with the computer model, since the main source of information about the mechanism of energy transformation in the unit is the model.

Student A's use of the model gives us an example of how unstructured exploration of the model could lead students to make connections between their design decisions and energy concepts. However, this example also shows one area where we could provide further scaffolding. The balance between allowing students to freely explore the model environment and scaffolding their interactions in order to connect more science ideas is one that still needs to be researched in this context.

Student B

Student B and her partner used the embedded computer model during the *reflecting* phase to confirm the design of the solar ovens they designed and built during the unit, but did not connect science concepts to their designs, and likely did not use the model to test the ideas about energy they gained during the unit. Notes written by student B and her partner about their interactions with the computer model are below.

> ***Trial 1:*** *We used alumanum foil which we used as a reflector and black colored paper. Our oven got 29.7 degrees celcius. It took us 300 seconds(or about 5 minutes) to get it that hot.*
> ***Trial 2:*** *On trial 2 we used aluminum foil as a reflector, black colored paper, and plastic wrap as an absorber. It got to 44.6 degrees celcius. It took 300 seconds(or about 5 minutes) to get hot.*
> ***Trial 3:*** *N/A*

During the unit, students only had the opportunity to design/redesign and test their ovens two times. Since student B only ran two trials, it indicates that she thought the modeling activity during the *reflecting* phase should be used to confirm the design of her previously built solar ovens. Further evidence that student B did not conduct new trials using the model comes from the language used in the *Trials* notes (above); aluminum foil was not an explicit choice in the computer model, instead we implemented a reflectivity slider. Student A wrote about settings used in the reflectivity slider, while student B did not, and because of this we presume student B only interacted with surface-level features of the model, not making the effort to make improvements to the ovens she designed or connecting the oven to the science concepts made explicit in the model. While the confirmation of physical models could be a useful activity for students to do with the model, students miss out on making the important connections between energy and design choices that comes from a more exploratory interaction. Student B missed out on these concepts by only testing the designs that were already explored in the physical design, build, test process.

Student B earned a score of 11 (30^{th} percentile) on the pretest, and also earned a score of 11 on the posttest (8^{th} percentile), indicating that student B did not gain the energy concepts simply from engaging in the design, build, test process.

Student C

Student C also used the model during the *reflecting* phase of the unit, but in a more exploratory manner. Student C (and her partner) conducted trials in a more exploratory way than student B, but in a more systematic way than student A. Notes taken by student C during the computer modeling activity are shown below:

> ***Trial 1:*** *We used plastic wrap on top. It got to about 46°C in about 3 minutes.*

***Trial 2:** We used Plexiglas on top and it shows that some of the solar radiation gets released er it enters. It got to about 38°C in 2 in a half minutes.*
***Trial 3:** We are using aluminium foil to cover the top of the oven and the top is reflecting the solar radiation and the cardboard is absorbing it. It has gotten to 26°C in about 2.5 minutes.*

Student C changed features of the modeled solar oven systematically, only altering the one aspect of the model. In Trial 1, student C used a plastic wrap cover, and trial 2 a Plexiglas cover, and in trial 3 an aluminum foil cover. Each trial also reported the final temperature and time in a somewhat systematic way. This pattern of interactions shows that student C was probably interested in how the different types of materials that could be used to cover a solar oven impacted the temperature, so student C explored this feature by conducting trials for each option. Student C also connects energy concepts to her design choices, writing about where solar radiation is let in, absorbed, and released.

Student C earned a score of 12 (53^{rd} percentile) on the pretest, and a score of 21 (97^{th} percentile) on the posttest. This progress between pretest and posttest was likely helped by student C's interaction with the computer model and shows that using the model can be beneficial for making connections during the *reflecting* phase as well as during the *planning* phase, as we predicted. Student C's interaction with the model was different than many other students in the *reflecting* condition [Student C *Trials* score: 3, Average *Trials* score in *reflecting* condition: 1.16, SD: 0.94]. However, this interaction represents an ideal in how we would like students to use the model to test new designs and connect energy concepts and design choices during the *reflecting* phase.

Study 1 Conclusion

Study 1 shows one way students can use computer models in conjunction with hands-on activities. This combination allows students to connect science concepts to their design decisions, since the computer model gives a visualization of energy transformation while encouraging students to make new design choices.

Integrating science and engineering concepts is a challenging task in both curriculum and assessment. During this solar ovens unit, we offered one method for integrating concepts about energy through the design and construction of a solar oven. Students first engaged with science concepts through online instruction in WISE, then designed, built, and tested physical solar ovens, interacting with a computer model either before or after the DBT process. The modeling activity created links between the science concepts and design decisions, helping students to visualize energy while also offering them a space to plan and test their designs in the *planning* condition, or to confirm their results and make further connections after they engaged in design process in the *reflecting* condition. In the *planning* condition, the model offered students the opportunity to add even more ideas to their repertoire, strengthening the distinguishing and reflecting activities later in the unit and leading to greater learning gains overall. This finding is supported by the knowledge integration framework.

Students who used the interactive model for *planning* during the unit were more likely to make gains on the integration of energy concepts between pre- and posttest. Many students in the *reflecting* condition used the model for very simple confirmatory analysis of models they had already built, which was not as helpful as using the model during the *planning* phased to connect science and design ideas. However, using the model during the *reflecting* phase to make further explorations or systematic tests could be useful for students while *reflecting*. In addition, students in the *planning* condition scored higher on embedded items (*Learn* and *Atmosphere*), which measured the number of normative links they made in each written response. This provides evidence that the *planning* condition allowed students to add more ideas to their repertoires, which stayed with them throughout the curriculum and they were then able to distinguish among those ideas on the embedded reflection questions (*Learn* and *Atmosphere*). Students did not spend as much time writing using the model to explore ideas in the *reflecting* condition, as evidenced by lower average scores on the *Trials* item (shown in Figure 2.4). Students in the *reflecting* condition may not have attended to the mechanism for energy transformation shown in the virtual model, or may not have considered as many variables while using the model, stunting their grasp of the complexity of energy transformation.

This study suggests the value of adding embedded assessment items to characterize more specifically how students use the model to plan the design of their solar oven. This could also be done using click-stream data collected from students' use of the modeling environment. Collecting this data would provide a clearer picture of what information the model is providing to students, and how different levels of interaction with the model may impact learning outcomes. For example, students might run the same trial multiple times without paying attention to the results; this may, however, help students learn about the mechanism of energy transformation. Understanding how students are using the model and what learning gains students make from certain actions within the modeling environment will help us to develop further scaffolding that will improve the learning experience for all students.

These results also suggest the value of using the model for both *planning* and *reflection*. To strengthen the *reflecting* phase, it may be advantageous to add critique activities (Chang & Linn, 2013; Zhang, 2010), such as asking students to evaluate each other's designs using the computer model or critiquing a fictional student's written response to a question by using the model to refute or agree with the fictional student. These types of critique activities help students to distinguish among the many ideas they may have about a topic in a way that feels relevant to students. Helping to critique another student's ideas or design feels helpful and allows for distinguishing and reflection on a student's own ideas, all while offering students a way to critique their own ideas in a non-threatening way. Critique may also give students a way to add even more ideas to their repertoires by allowing students to consider variables from the model in new and different ways.

In the next study, we build on the findings from study 1 by using the model at the *planning* phase for all students. However, we also test the use of the model as an activity for revising the design of the solar oven.

Study 2: Physical vs. Virtual Revision Activity

Study 1 showed that the computer model developed for this unit helps students make connections between science concepts and the decisions made during the design process and that students who interact with the model earlier during the curriculum make greater improvements than students who interact with the model later (McBride et al., 2016).

Here, we investigate how the computer model can be used as an alternative to rebuilding a physical solar oven prototype. Making changes to improve a design, testing them, and reflecting on the results are important opportunities to learn during the engineering design process. Many engineering projects include some form of testing and iterating. We investigate whether a focus on integrating science concepts and design in the model or in the physical design will help students improve their understanding of the underlying science ideas. Within the knowledge integration framework, students add new ideas, distinguish their ideas, and reflect during the redesign process. Thus, it is an important place to support students during engineering projects.

We aim to test the impact of replacing a second iteration of rebuilding and testing the physical prototype with a virtual modeling activity. Student groups were assigned randomly to one of two conditions – *modeling* or *physical* revisions. The virtual modeling activity may help students connect their revisions to underlying science principles. However, students may be more interested and invested in rebuilding and testing a new version of their physical prototype.

Here, the model is assessed as a tool for making revisions. The revision process is an important learning opportunity during engineering projects, as students are able to evaluate how their first prototype functioned. This evaluation process can encourage deep thinking about the underlying science concepts that impact the prototype.

Methods

Materials and data source

Three teachers from two different schools participated in this study, along with their students (*N*=283). Teacher A (*N*=124) teaches 6th grade at a school in the suburbs of a large U.S. city that serves a mostly middle SES community (32% free and reduced lunch, 5% ELL). Teachers B (*N*=80) and C (*N*=79) teach 12th grade at an urban school serving a mostly lower-middle SES community (73% free and reduced lunch, 19% ELL), and teach honors physics (teacher B) and physics (teacher C). We conduct analyses on the entire corpus of data (all students), as well as separating by grade level. In this study, students completed pre- and posttests individually. During the unit, students worked in pairs or triads.

Curriculum

The two conditions in this study are outlined in Table 2.3. The main difference between conditions is that in the *modeling* condition, students revise and test revisions using a virtual model, while in the *physical* condition, students revise and test their physical prototype. Class time was equalized between the conditions. Students in the *modeling* condition could conduct more revision cycles than those in the *physical* condition during the same amount of class time, and students in the *physical* condition only need to revise their model, not build an entirely new oven.

Table 2.3: Curriculum outline, with conditions

Activity	**Description & Items of Interest**
Introduction to Solar Ovens	Elicit initial student ideas about energy transformation
Solar Radiation and the atmosphere	Energy comes as radiation from the sun; energy can be absorbed or reflected. Students use a simulation to investigate energy.
Solar Radiation and Greenhouse Gases (GHGs)	Describes how energy interacts with greenhouse gases. Students use a model to investigate how addition of GHGs impacts energy.
Model 1	Students use an interactive model to investigate how radiation works in a solar oven, the impact of various design decisions
Design, Build, Test 1	Design oven under budgetary constraints using a draw tool, build, test under a heat lamp using a temperature probe to collect data
[*Physical* Condition] Design, Build, Test 2	Students reflect on what was learned from the first iteration; use new budget constraints to repeat process
[*Modeling* Condition] Model 2	Students use an interactive model to decide how they would redesign their ovens
Reflect & Posttest	Students describe how their solar ovens work using energy from the sun; make connections between solar ovens and the atmosphere, and take the posttest
Model/Build	Students do the opposite activity from what they did earlier in the curriculum (modeling or design, build, test)

Assessment

To assess student progress, we used responses from pre- and posttests, as well as responses embedded within the curriculum. The pre- and post-test assessments measure student ability to link concepts about energy. The pre- and posttests were each made up of the same items as in study 1: *Car, Laura1, Laura2, GHG1, GHG2,* and *Model.* In addition, a new item, *Tomatoes,* prompted students to think about what they would need to consider when building a structure for growing tomatoes during colder months, and how their decisions would impact the flow of energy. The *Tomatoes* item also appeared on both the pretest and the posttest.

Embedded within the unit, we examine two items, also scored using knowledge integration rubrics. The first item, *Revise,* asks students to explain how they would revise their oven (*Model* condition) or how they did revise their oven (*Physical* condition) and to explain why they made those changes. The second item, *Compare*, prompts students to make comparisons between the atmosphere (the greenhouse effect) and their solar ovens.

Results & Discussion

Of the 283 students participating in the study, 249 completed the pretest, some part of the unit, and the posttest.

Overall, all students learned from this unit (Pre: M=20.45, $s.d.$=3.81; Post: M=22.13, $s.d.$=4.24). Across all students and items there was no effect for condition. Examining the pre-/posttest items in separate groups (science, modeling, design), also showed no significant effect of condition. On examining the embedded items using a regression model, there was no significant difference found between conditions on the *Compare* item, but there was a significant difference on the *Revise* item in favor of the *modeling* condition (β=0.33, p=0.01).

Looking at high school and middle school students, we explore the effects of the conditions. As expected, the high school students had higher pretest scores than the middle school students on average (High School: M=22.55, $s.d.$=4.18; Middle School: M=18.95, $s.d.$=2.65). High school students have spent more years learning science, and are therefore more familiar with the topics of energy transformation than middle school students.

For high school students, there was not a significant difference between conditions from pre- to posttest. Of interest, the trend for high school favored the *physical* condition (β=-1.07, p=0.11), though the difference was not statistically significant. There was also not a statistically significant difference between conditions when considering embedded items.

For middle school students, there was also not a significant difference between conditions from pre- to posttest. Of interest, the trend for middle school slightly favored the *modeling* condition (β=0.73, p=0.13), though this was not statistically significant. There was not a significant difference by condition for the *Compare* item, but there was a significant difference favoring the *modeling* condition for the *Revise* item at the middle school level (β = 0.40, p =0.01).

Examining all students together using a regression with an interaction term for grade level and condition, the interaction term was a significant predictor of posttest score (β=1.50, p<0.05). This shows the effect of condition differs for the two grade levels.

The differences between high school and middle school students suggest that the modeling activity may be more useful for students in middle school, while the building activity may be more useful or interesting for high school students. Some factors that may impact this difference in effect of condition are prior knowledge, motivation, and metacognitive skills. Overall, the fact that there was no difference between conditions is of interest because it shows that doing a short *modeling* activity (students spent an average of ~7 minutes using the model during the revision) can be just as beneficial as the longer *physical* revision. The significant difference for the *Revise* item shows that the *modeling* activity may help students use science concepts to describe their revisions more than the *physical* revision.

Conclusions

Together, the results of these two studies suggest that students benefit from using the interactive computer model to think about design during multiple sections of a design, build, test curriculum. The context of energy transformation lends itself particularly well to using computer models, since the science phenomenon being studied and applied by students during the design project is invisible. Often, students may think that conducting a physical experiment will help them to see how energy transformation takes place, but that is not true. While a physical experiment can, in this case, allow students to connect their decisions with the outcomes of experiments, they will never be able to visualize how energy transformation happens without the aid of a computer model.

The results of this work offer guidelines for utilizing interactive computer models in hands-on design projects. These models, if designed to help students visualize science concepts, can help students plan, reflect upon, and redesign physical models. However, the models must also be interesting and realistic enough to hold student attention. If the computer model shows a phenomenon that seems too simple to students, they may not feel compelled to use the model. In this case, students may have assumed that they already understood the process of energy transformation, and so they did not need to use the model as much. Further, scaffolding student interaction with the computer model can also be useful to ensure that students are using the model to test ideas they have about their designs and gain useful information. Further analysis on model scaffolding to promote learning and knowledge integration will be discussed in chapter 3.

Since time constraints are often a concern when conducting hands-on engineering projects in classrooms, the results of study 2 are especially promising. They suggest that using the model for a redesign and reflection can help students to benefit from considering how they would improve their design just as much as conducting a redesign, then rebuilding and testing the physical prototype.

Chapter 3: How Students Think About and Use Models

This chapter investigates how students interact with computer models. Two studies are presented, with data from two different but related curriculum units. The first uses student meta-data from interactions with a computer model within a Global Climate Change curriculum unit. We investigate how students use the model without built-in scaffolds for structuring in the modeling environment. We find that students exhibit periods of observation and periods of activity when using the model. Students with more periods of observation when using the model make greater science learning gains from pretest to posttest. Given these findings, in a second study we developed a table that automatically generates when students run trials in the Solar Ovens curriculum unit, encouraging more periods of observation and giving the added benefit of structuring information for students to look back on when making design decisions. We investigate how students use this automatic table scaffold, finding that students who generate more rows (more trials) in the table make greater science learning gains from pretest to posttest.

Together, the results of these studies point to a need for scaffolds that structure student use of computer models as well as further emphasis on the benefits and uses of computer models.

Introduction

Interactive simulations are commonly used tools in technology enhanced education, since they have the ability to make complex phenomena more concrete. This is especially true in the case of energy. However, in order for these interactive simulations to be successful learning activities, students must be able to engage with them in a manner that allows them to glean the correct information.

Helping students use interactive simulations is also important in light of the Next Generation Science Standards (NGSS Lead States, 2013). One of the science and engineering practices, which span grade levels and topics, focuses on developing and using models. Thus, helping students to develop good practices around using models is important in many areas of science for students across grade levels.

Simulations are often used as inquiry activities, especially in science disciplines. In inquiry activities, students are able to develop their own knowledge by asking scientific questions, answering those questions using evidence, developing explanations, and connecting explanations to scientific knowledge (Olsen & Loucks-Horsley, 2000). Often, students can make incorrect assumptions about what they see in a simulation because they may not know some of the underlying rules that govern how the simulation plays out.

Tools and methods exist to measure certain scientific practices associated with computer models, for example controlling variables in a simulation (De Jong & Van Joolinger, 1998; Lin & Lehman, 1999). However, developing ways to measure practices, like deliberate observation of how a simulation plays out, are not readily available. This also relates to work on autonomy and agency; how students are able to independently determine enough governing rules of the simulation that they are able to use appropriate techniques to extract information from the simulation. For example, in the climate change simulation, students are able to add or remove greenhouse gases to the atmosphere, but there is a significant time lag between when students add/decrease the greenhouse gases and when the temperature visibly changes in the graph that accompanies the simulation. If students do not allow the simulation to run for a period of time, they may not see the connection that more greenhouse gases in the atmosphere leads to a higher temperature in the atmosphere.

Across both studies we use the knowledge integration framework to guide our analysis and design. Curriculum designed using the framework has been shown to increase the coherence of students' scientific ideas (Linn & Eylon, 2011). This framework helped us to develop simulation scaffolds so that interactions lead to a more integrated understanding. Following the framework, we elicited students' ideas, guided exploration of new ideas, and supported the process of organizing and distinguishing among ideas to achieve a coherent understanding of a topic.

Previous work in educational data mining has examined the positive actions taken that correspond to scientific practices, like controlling variables through multiple trials (Sao Pedro et al., 2013) or using inquiry skills (Gobert et al., 2013). Work has also been done to classify student use of models (McElhaney & Linn, 2011; Linn & Hsi, 2000). We augment this work by focusing on scientific practices like making deliberate observations, especially in K-12 settings. Deliberate gathering of evidence can be measured by video records of learners or by log files. Using log files allows efficient analysis of large groups of student interactions.

Study 1: Periods of Observation and Activity

In this study we examine student use of a relatively simple simulation about greenhouse gases that was a part of a web-based science curriculum on climate change. Since students only have two options for controlling the simulation, we examine their action logs to understand how students watch the simulation and whether students who allow the simulation to play out more often are able to understand concepts more thoroughly than those who take more actions.

Although simulations can reveal hidden phenomena, feature design can impact whether or not students attend to subtle mechanics of the interactive system. The simulation we use, about climate change, focuses on helping students understand the relationship between greenhouse gases, the energy that comes to earth from the sun, and the energy emitted by earth toward space. These simulations also offer the benefit of allowing students to make connections between small scale and large scale phenomena, for example, the interaction of energy with greenhouse molecules and global warming (Wilensky & Reisman, 2006).

Methods

Materials and data source

This study uses data collected from middle school student interactions with an inquiry simulation on climate change. Students interact with this simulation at four different points throughout a 5-7 day online curriculum unit on climate change. This curriculum was run using Web-based Inquiry Science Environment (WISE). Major components of the curriculum were solar radiation and the atmosphere, greenhouse gasses, and the ozone layer. We use the action log data from the simulations, which includes clicks of buttons in the simulations and timestamps. We also use scored items from pre- and posttests that were conducted one day prior to beginning the unit and one day after finishing the unit, respectively. Pretest and posttest questions asked students about energy and the greenhouse effect and were a combination of multiple-choice and essay items.

Essay items were scored using a knowledge integration rubric (Linn & Eylon, 2011) to score responses based on the number of correct links made between ideas. The scale for knowledge integration rubrics is 1 (low) to 5 (high). An example rubric can be found in Table 3.1.

Table 3.1: Example Knowledge Integration Rubric for an item that asks students if a car will be warmer or colder than the surrounding air after sitting in the sun on a cool day

Score	Level	Examples
0	**No Answer**	
1	**Off Task**	*I don't know.*
2	**Irrelevant/Incorrect**	*The inside air and the outside air are the exact same temperature because the sun is not enough to heat the inside if the car.*
3	**Partial** Normative isolated ideas without a valid link	*The solar radiation would go through the metal and would stay in the car when the outside air wouldn't be able to get inside.*
4	**Basic** Elaborate a scientifically valid link	*it would be warmer than the outside air because if the car hasn't been driven for a week and its been in the sun the whole time the car will absorb the heat and scence there is know way the heat can get out of the car the heat will just keep building up.*
5	**Complex** Elaborate two or more scientifically valid links	*The sun produces solar radiation which heats up the car and the infrared radiation gets trapped in the car which leads to the temperature rising.*

Students interacted with a model, built using NetLogo (Wilensky, 1999), at four different points during the curriculum. Each instance of the modeling activity allowed students to manipulate a different feature of the model. The first allowed manipulation of the reflectivity of the surface, the second of concentration of greenhouse gases, the third of concentration of ozone (in the stratosphere), and the fourth of both ozone and greenhouse gases. We choose to use the second instance of the model only during our analysis for two reasons. First, students were often less confused the second time they interacted with the model than during the first interaction, so their actions have more meaning. Second, based on analysis of student thinking about the ozone layer's impact on global temperature at pretest and at posttest, students did not adequately learn what they were supposed to from the model – that the ozone layer has little to no impact on global temperature. However, students did make gains in understanding the role of greenhouse gases during the course of the curriculum. Therefore, we use only the second model, in which students manipulate and make sense of greenhouse gases' role in global warming.

The model (Figure 3.1) allows students to add or decrease greenhouse gases and change the speed at which the simulation runs. Two graphs to the side of the simulation show the concentration of greenhouse gases and the global average temperature as functions of time.

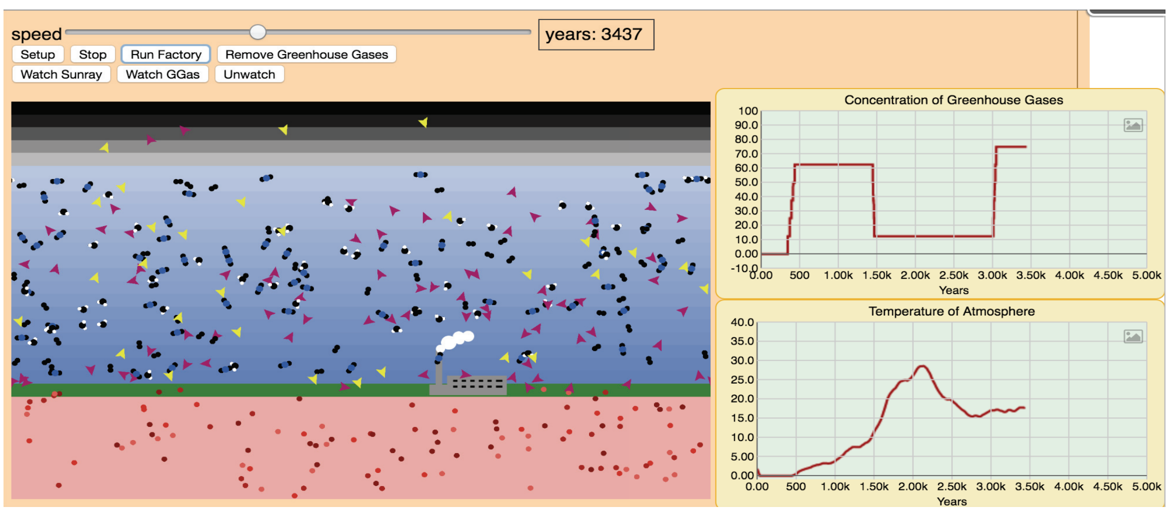


Figure 3.1: Image of Netlogo greenhouse gas simulation used in curriculum

The data comes from 422 students in three classrooms from two different schools. The students at these schools come from generally middle socio-economic status communities. During simulations and the unit, students worked in pairs, but students were assessed individually at the pre- and posttests. Because of this, we use averages of pre/posttest scores across student groups in this analysis. There were 188 student groups that interacted with the model.

In this dataset, we remove step visits when the only logged actions are loading and exiting the page, thus counting only step visits during which students made some interactions with the model.

Descriptive statistics

Students were able to return to steps during the curriculum. Students revisited this model step an average of 1.98 times (*s.d.*=1.28, *min*=1, *max*=7). Students were allowed to spend as much time as they wanted using the model. Across all visits, students spent an average of 585.65 seconds using this model (*s.d.*=305.55, *min*=1, *max*=1771) (Figure 3.2). Students made improvements overall from pretest to posttest (Figure 3.3).

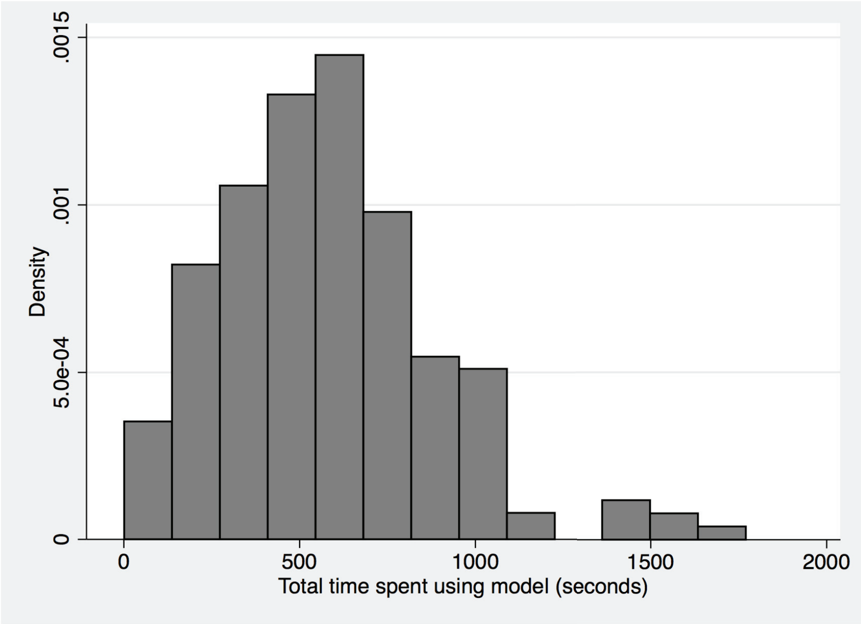


Figure 3.2: Histogram of the time spent on the simulation by student groups

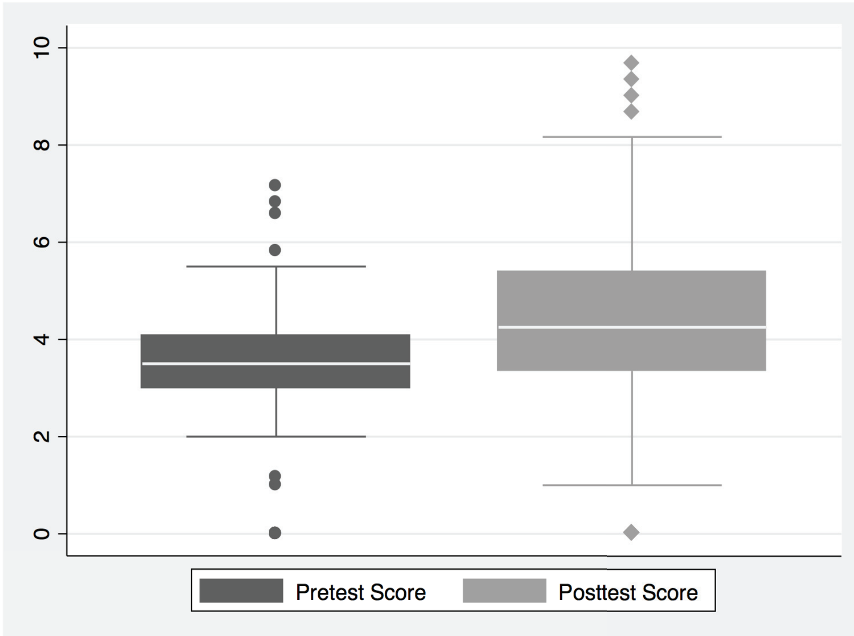


Figure 3.3: Box plots of pretest and posttest scores, showing that students made gains between pretest and posttest

Analysis

We define an observation period as one in which the student does not interact with the model after an initial period of interaction. This observation period could last any amount of time, but based on our classroom observations and the amount of time it takes for the model to show variable relationships when at a medium speed setting, we choose to define an observation period as lasting a minimum of 25 seconds. We do not set an upper bound for the observation period, since students work in pairs and longer observation periods may indicate they were talking with each other about the model. We also investigate minimum observation periods of 15, 20, and 30 seconds to find out whether the same outcomes hold. For each student, we count the number of observation periods across all visits to the step. We calculate the proportion of the total time spent on the model that the student is observing, based on our definition.

Results & Discussion

We use a regression model to predict a student's posttest score based on the number of observation periods per pair, controlling for the pretest score. This regression model shows that across minimum observation periods from 15-30 seconds, the number of periods of observation per pair is significant, even across different definitions of how long the minimum observation period must last (15 seconds: $p<0.01$, $\beta=0.069$; 20 seconds: $p<0.01$, $\beta=0.096$; 25 seconds: $p<0.01$, $\beta=0.124$; 30 seconds: $p<0.01$, $\beta=0.130$). These results help validate our choice to use a 25 second minimum observation period, given the significance of this cutoff point as a minimum for an observation period in predicting posttest scores, as well as the observational evidence gathered from both the classroom and the design of the model. While the regression results are useful, we use this cut-off based empirically on the time it takes for the model to generate useful data trends.

As predicted, the proportion of the total time spent on observation periods is also a significant predictor of posttest performance, when controlling for pretest score (β=1.42, p<0.01). The relationship between the gains made between the pretest and the posttest and the number of observations made can be seen in Figure 3.4, and the relationship between the proportion of time spent in "observation periods" and score gains in Figure 3.5.

Our analysis of the data revealed several patterns. Students:

1. Make rapid changes, then have periods of observation,
2. Make a single change, with periods of observation between each change
3. Make rapid changes with no periods of observation

Often, students exhibited all three of these behaviors over the course of their interactions with the simulation. Defining the boundaries of each of these three activity types will be necessary before analyzing how the proportion of each activity type impacts student learning. Examples of 8 student pairs and their actions are shown in Figure 3.6, showing the complexity of making these distinctions among groups of actions.

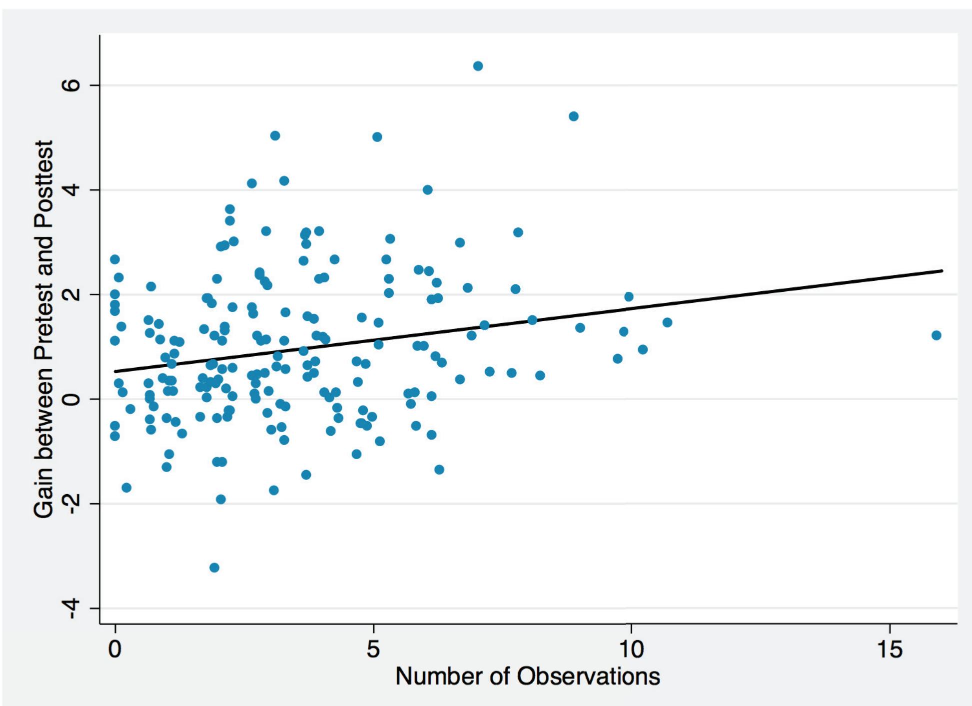


Figure 3.4: Scatterplot of the gains made between the pretest and posttest and the number of observation periods for a group

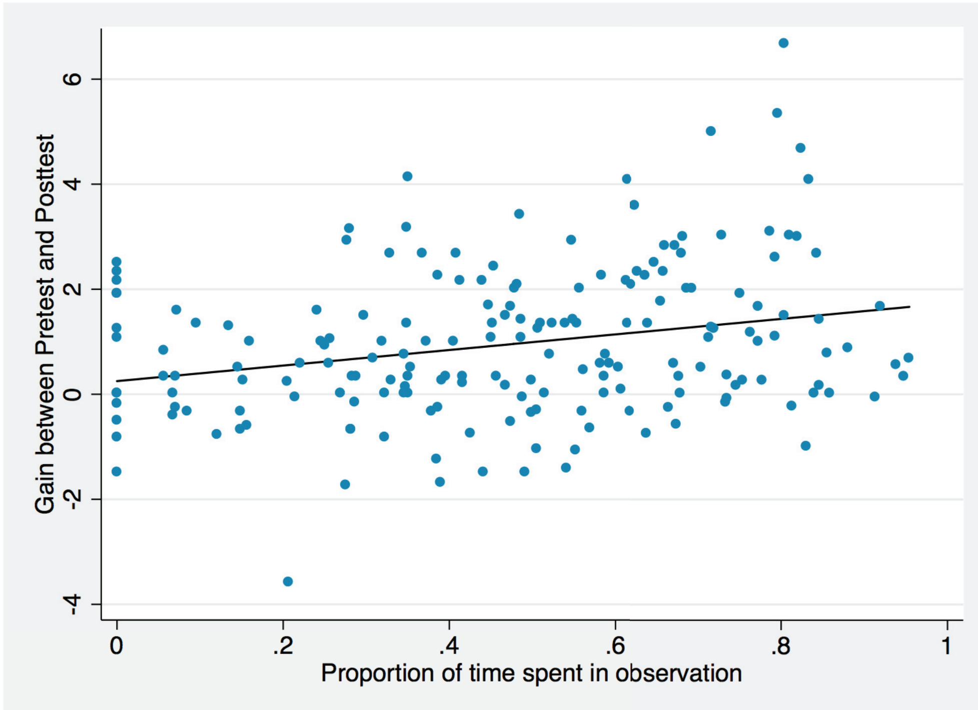


Figure 3.5: Scatterplot of gains made between the pretest and posttest and the proportion of time spent on observation periods of the total time spent on the simulation

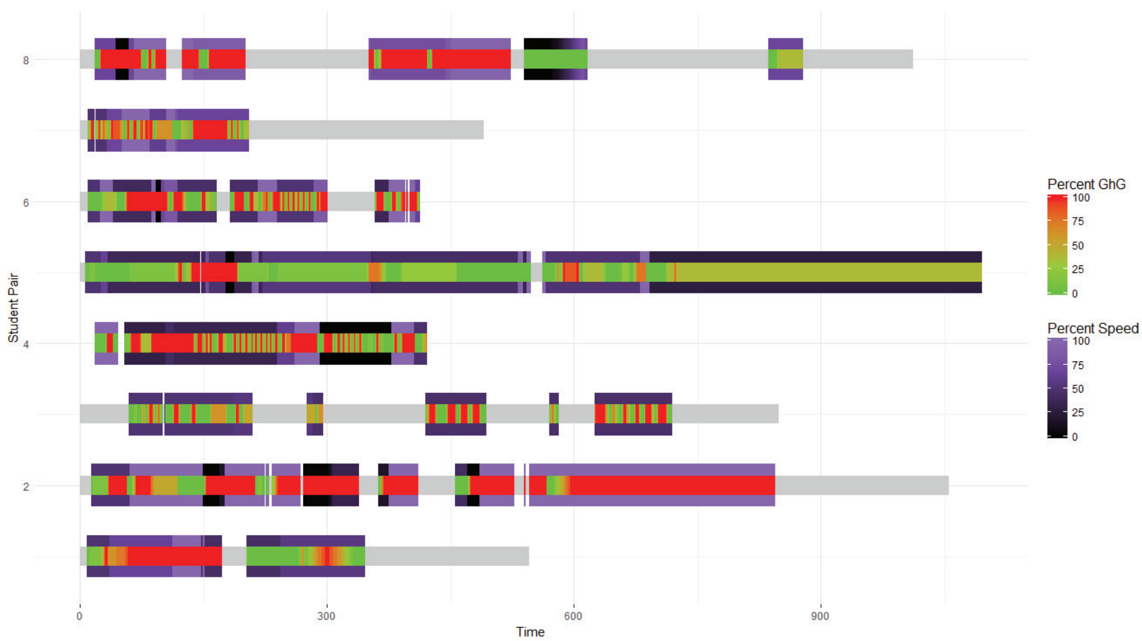


Figure 3.6: Eight student pairs and their actions while interacting with the computer model. Each gray bar shows the amount of time students used the model, while the green-red bars show when students were making changes to the Greenhouse Gas levels and the purple gradient bars show when students were adjusting the speed of the model.

Study 1 Conclusion

This study sheds light on the implementation of the NGSS practice of modeling, by clarifying how students use models and identifying promising ways to guide the process of interacting with models. Students often explore the available buttons and features of a model in quick succession without analyzing the results. While exploration is important for identifying options, that is only the first step. To discern nuanced relationships in models, students need to deliberately investigate their features. This may include trying many combinations to refine understanding of the scientific principles that underlie the model.

This study may also help us to understand how best to determine when students are sense-making and when they may be disengaged (e.g. Gobert et al., 2015; Joseph, 2005). Prior research suggests that a continuous stream of interaction is a good indicator of sense-making. This study goes beyond analysis of a stream of interaction to look at the nature of pauses. Since simulations often require students to make sense of complex concepts, students may need extra time to examine the simulation, the actions they took, and the output in order to engage in sense-making. Thus, detectors that require students to take action to be engaged may inadvertently miss meaningful student interactions.

The results of this study lead us to the next study, where we implement an additional scaffold for structuring information in the model. Students are able to run trials that automatically generate rows in a table, giving them added time to do sense-making after they have collected information from their trial.

Study 2: Student Use of Computer Models to Run Trials

Simulations can be powerful tools for allowing students to engage in inquiry, especially in science disciplines. To succeed, these simulations generally benefit from scaffolds that guide students to keep track of their investigations and reach meaningful insights (McElhaney & Linn, 2011). In this study, we examine guiding questions and recording of trials in a table as scaffolds. We use a simulation of a solar oven that allows students to investigate the multiple variables at play in energy transformation and gives representation to invisible phenomena that can be especially confusing for students (Wilensky & Reisman, 2006; McBride, Vitale, Applebaum & Linn, 2016). We report on how students utilize the simulations for learning and how scaffolds function to alter the learning experience. We investigate how systematic versus exploratory testing with a simulation impacts learning of science concepts.

To study learning from the use of computer models, we use a curriculum in which students learn about energy and then design, build, and test their own solar ovens. During physical construction of ovens, students are guided through an online curriculum to support scientific reasoning about the process. This curriculum is unique because it combines online and hands-on learning to take advantage of the unique affordances of each. In addition to the physical ovens, students use an interactive simulation that models solar oven mechanics while designing and redesigning their solar ovens. We use a table and guiding questions to scaffold students' interactions, and we study the trials students run in the simulation and their relation to learning. These trials are automatically added to the table so students can keep track of what they have tried.

As in study 1, we use the knowledge integration framework to create the curriculum about solar ovens, because the framework focuses on building coherent understanding (Linn & Eylon, 2011). This framework offers instructional design principles to enhance connections between design decisions and scientific principles. The knowledge integration framework has proven useful for design of instruction featuring dynamic visualizations (Ryoo & Linn, 2012) and engineering design (Chiu, et al., 2013; McElhaney & Linn, 2011). When students build a physical artifact, as in this curriculum, they can only test a few of their ideas due to time and material constraints. Using the interactive model during the design phase allows students to explore many more ideas.

Previous research has shown that students tend to use simulations in diverse ways. Students may intentionally seek specific goals, engage in exhaustive testing of features, or test in an unsystematic manner (McElhaney & Linn, 2011). Unsystematic testing, in particular, typically does not contribute to learning. However, by providing scaffolds that support intentional choices during inquiry, simulations may foster stronger learning gains (Reiser, 2004).

Various scaffolding methods have been used with interactive simulations. Often, these scaffolds are implicit, or built into the system with the simulation (Podolefsky, Moore & Perkins, 2013). For example, guiding questions are often used in inquiry simulations to direct students' attention toward certain features of simulations (Hmelo & Day, 1999). Other tools, like concept maps and note-taking spaces can also assist students in making sense of inquiry simulations (Kali & Linn, 2008; Rye & Rubba, 1998; Svihla & Linn, 2012). Students are often encouraged in science to run multiple trials and control variables between trials (only change one variable between trials). A control of variables strategy can help students to determine the effect of a single variable on a more complex system, although in some cases students may benefit from more exploratory strategies (McElhaney & Linn, 2011).

To give students a record of their trials, we developed a table that automatically populates after a student lets the simulation run for a certain period of time. Although we do not describe the control of variables strategy during the curriculum in any way, we analyze spontaneous use of the strategy as a potential indicator of engagement and learning. Students who make more changes between trials may be considering their choices more carefully than those who run many trials but only change one variable between each trial. Indeed, engineers often experiment based on their informed intuitions and do not run control of variables trials. Using log files from student interactions with the curriculum and output from the automatically generated tables (simulation scaffolding), we use feature engineering to identify how students use the model and whether these uses have an impact on learning. We specifically develop features that have to do with the control of variables strategy, such as the number of trials a student runs and the percent of those trials that are systematic. These types of techniques have also been used with more complex simulations and micro-worlds (e.g., Gobert, Kim, Sao Pedro, Kennedy & Betts, 2015; Conati, Fratamico, Kardan & Roll, 2015)).

Materials and data source

Curriculum

This study focuses on a curriculum about solar ovens that is run using the Web-based Inquiry Science Environment (WISE). During this curriculum, students design, build, and test a solar oven. They go through the design, build, test process two times to get an idea of how engineers iterate on their designs based on results from testing (Figure 3.7). Students are allowed to use only a certain set of materials (e.g., tin foil, black construction paper, plastic wrap, Plexiglas, tape), in addition to a cardboard box they bring from home (Figure 3.9). Students use an interactive computer simulation to test the different materials in their oven. This simulation helps to elicit student ideas before they get to the building process, consistent with the knowledge integration framework. The testing portion of the project allows students to distinguish their ideas. In this way, the (potentially) more exciting building activity is bookended by activities that will help the building to be an enriching learning activity.

Throughout the project, students respond to short response style questions about the choices they are making in their design and how their ovens work. This curriculum is unique, since it is guided by an online platform, but students also design, build, and test their solar ovens in a hands on portion of the project. An image of two of the solar ovens built by students is in Figure 3.8. Solar ovens are built using cardboard boxes and take advantage of light from the sun. They convert the solar radiation to heat and infrared radiation, and trap that heat and infrared radiation inside the oven. The combination of the online and hands-on curriculum aims to take advantage of the unique affordances of each style of curriculum. Using online curricula offers the opportunity for dynamic simulations and visualizations in the same space as other features like writing activities and concept mapping activities. Hands-on curriculum may be motivating for students and provide opportunities for real life observations.

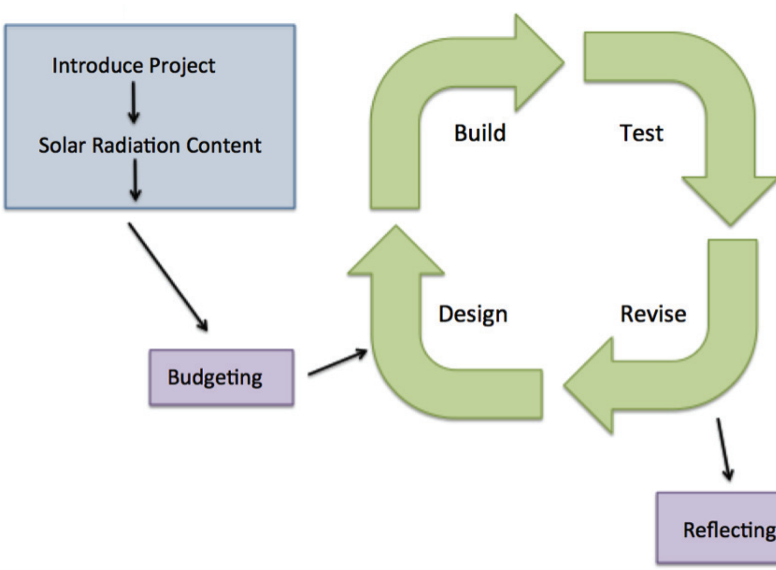


Figure 3.7: Outline of the solar ovens curriculum

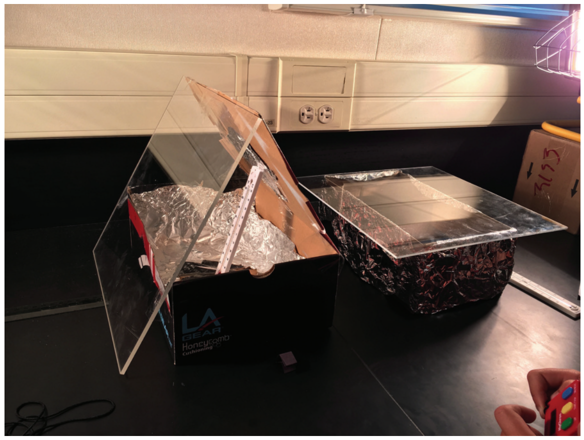


Figure 3.8: Image of two solar ovens built by students during the testing phase

Materials & Costs

1 sheet of construction paper (8.5 inch x 11 inch) 	$2
12 inch x 12 inch Sheet of tin foil 	$7
12 inch x 12 inch Sheet of plastic wrap 	$6
You can RENT: 12 inch x 12 inch sheet of Plexiglas (thick plastic)	$10
3 feet of tape (Duct, masking, clear) 	$3

Figure 3.9: List of materials and costs for the building of the solar oven. Students are allowed to spend $20 during iteration 1 and then add $13 worth of materials during iteration 2.

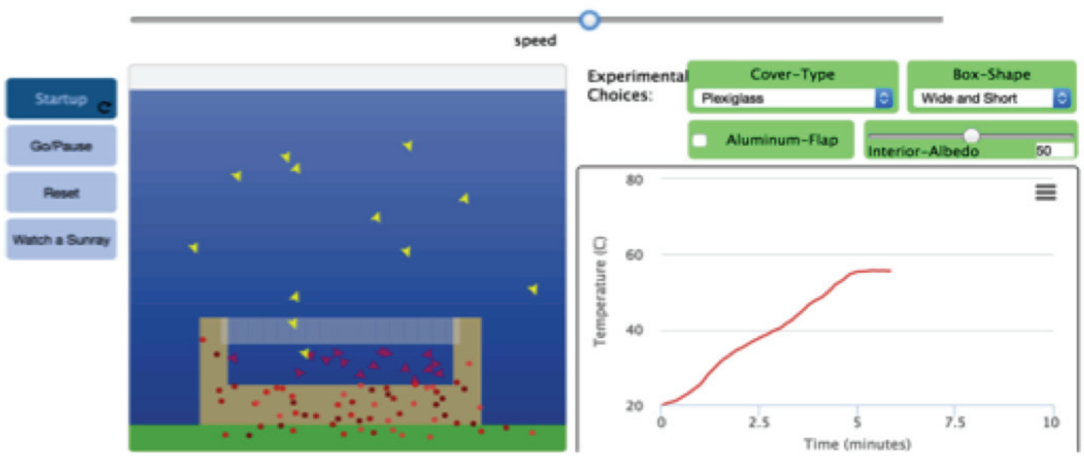
This curriculum takes between 10-15 class periods (~45 minutes per class period), depending on how the teacher manages the classroom and project. Students complete this project in groups of 2 or 3 students. Students also complete a pretest the day before the project begins and a posttest the day after completing the solar ovens project. Students do the pretest and posttest individually.

Computer simulation

The interactive simulation was built using NetLogo (Wilensky, 1999). In this simulation visualization (Figure 3.10, left side), solar radiation (yellow arrows), emerge from the top of the display and travel down towards the virtual solar oven. Some of this radiation is absorbed by the solar oven and transforms into heat (red circles). These are then emitted as infrared radiation (purple arrows) from the material that absorbed them. Students can manipulate the simulation in a number of ways. They can change the cover on top of the oven, whether or not there is a reflective flap on top of the box, the shape of the box (wide and short or skinny and tall), and the albedo (reflectivity) of the inside of the box. Students may also manipulate the speed at which the simulation runs.

Once a simulation runs to the end of the graph (10 simulated minutes), a new row is added to the table below the visualization with the settings and results from the trial. If the students do not allow the simulation to run until the simulated 10 minutes finish, nothing is added to the table.

The scaffolds we developed for the interactive simulation are twofold; short responses style questions direct students to investigate capabilities and limitations of the simulation and an automatically generated table helps students to keep track of trials they have run. The table includes information about all of the settings used in that trial, as well as the results of the trial at certain time points (e.g. 5 minutes, 10 minutes).



#	Box Shape	Cover Type	Flap?	Int Albedo	Temp at 0m	Temp at 1m	Temp at 5m	Temp at 10m
1	Wide and S	None	No	50	20	24	30.2	31.5
2	Skinny and	None	No	50	20	26.2	29.7	31.1
3	Wide and S	Plexiglass	No	50	20	25.6	55.4	53.4

Figure 3.10: The interactive simulation used by students to test solar ovens and visualize energy transformation; below the table simulation is output from the automatically generated table

Data

This data comes from 635 students across three schools and five teachers. During this study, students participated in a pretest and posttest (each lasting one class period), as well as the 2-3 weeklong curriculum. During the curriculum, students worked in teams of 2-3. These 635 students formed 255 teams. The school demographics represent a diverse population of students. We drop 77 of the students in this study because they either did not participate in the pretest, the posttest, or a majority of the curriculum. If a student scored below a 5 (total score) on either the pretest or the posttest, they were dropped from our analysis. Since students receive a score of 2 for attempting a problem in our coding rubrics and there were 7 problems, this meant a student did not even attempt much of the test. These students may not have completed portions of the study due to absence or other school-related disturbances. After dropping these students, there were 558 students and 246 groups or partial groups remaining.

Pretest and posttest

We use pre- and posttest data to assess student understanding of scientific methods and concepts Students answer 7 questions on both the pretest and the posttest, and the questions are exactly the same on each. Five of these questions ask about science concepts (e.g., energy transformation), one question asks students about a different engineering design project using science to justify decisions, and a final question asks students to use a computer simulation to help a fictional student decide what shape to build his solar oven.

We score each question using a knowledge integration rubric. These rubrics score each response on a scale from 0-5 based on the number of links students make between concepts. Students are not penalized for incorrect statements. A score of 0 indicates that a student did not attempt the question, while a 1 indicates that the students' response was off-task.

A score of 2 indicates that the student responded to the question with only incorrect or irrelevant ideas, but was on- task. A score of 3 indicates that a student made one link between two ideas, while a four indicates 2 links and a 5 indicates 3 or more links. We use the sum of both the pretest and posttest, where the maximum score for each would be 35 and the minimum 0. Histograms showing the pretest and posttest scores (summative) are shown in figures 3.11 and 3.12. Figure 3.13 shows the relationship between pretest and posttest scores.

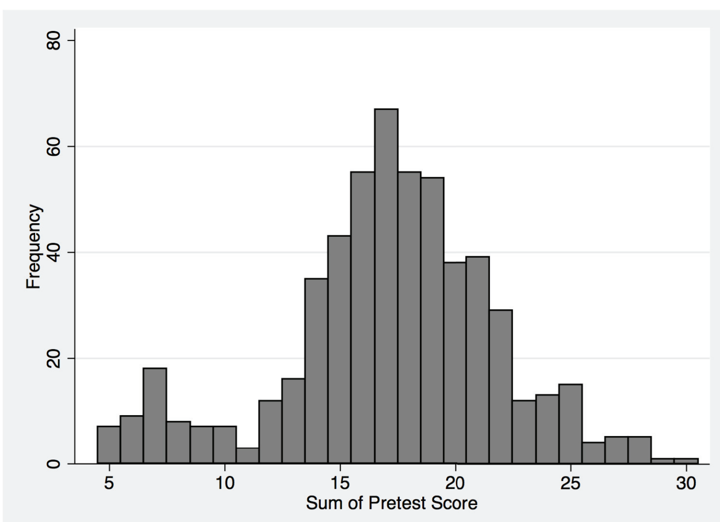


Figure 3.11: Histogram depicting the frequency of pretest scores (Mean: 17.17, Std.dev: 4.69)

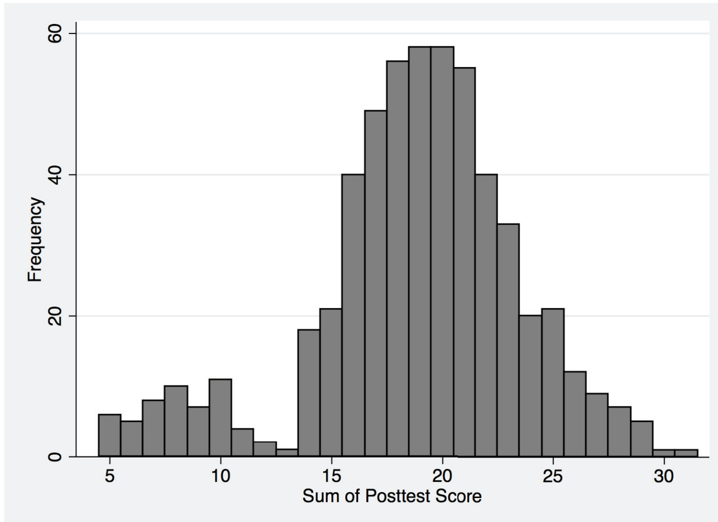


Figure 3.12: Histogram depicting the frequency of posttest scores (Mean: 18.85, Std.dev: 4.77)

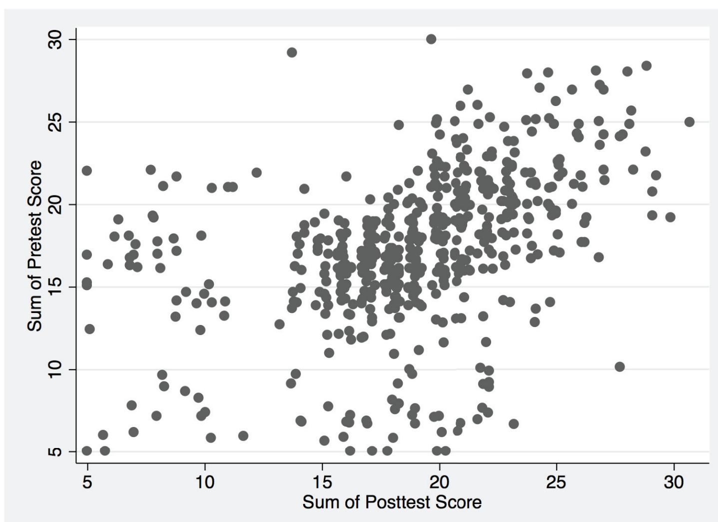


Figure 3.13: Scatter plot depicting the pretest and posttest scores; Mean gain from pretest to posttest: 1.67

Analysis

Descriptive statistics

Of the 246 groups who participated in the curriculum, 216 (87.80%) of the students used the computer model to produce at least one row of data during the first design iteration. We consider each row of data produced to be a trial. During the second design iteration 164 (66.67%) of the groups used the computer model to produce at least one trial. Overall, 152 (61.79%) groups used the computer model during both iterations to produce at least one trial. As seen in figure 3.14, many groups do not use the simulation scaffolds at all and produce zero rows in the automatically generated table. Still more students produce only 1 row in the table, which may mean they are confirming their ideas for a solar oven that they have already discussed and planned prior to using the simulation and without any evidence outside of their intuitions.

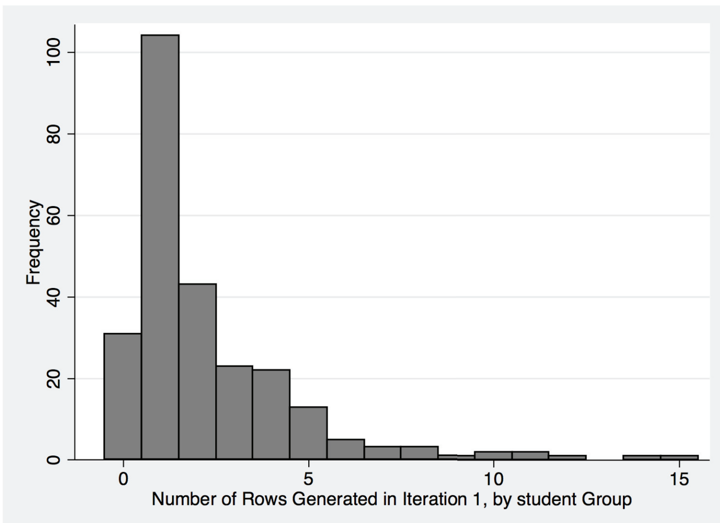


Figure 3.14: Histogram depicting the frequency of the number of trials run by a group of students during the first iteration of using the simulation (Mean: 2.27)

Students also responded to open response prompts before and after the first iteration of the simulation that were meant to provide scaffolding for student use of the simulation. These questions were loaded on the same page with the simulation so students could use the simulation and work on their responses to the questions without leaving the page. Even though 30 of the groups did not generate any rows in the table, only 16 of the groups did not respond to the question occurring before the simulation and only 4 groups did not respond to the question occurring after the simulation. Two of the groups did not answer either question, but did generate rows using the table scaffolding. Based on this data and analysis of log files, all of the 246 groups used the simulation in some way, though not all the groups used the table scaffolding.

Based on this, students used the model in different ways – some to write explanations, some to do trials. Since we provided multiple scaffolds in the model, it was possible for students to use some, but not all, of them and still have a good learning experience while using the model.

Controlling variables

Students are often encouraged in science to run multiple trials and control variables between trials (only change one variable between trials). A control of variables strategy can help students to determine the effect of a single variable on a more complex system. However, a more exploratory strategy in which students complete multiple trials (generate multiple rows of data), but do not necessarily control variables, may be beneficial for learning if students are thinking deeply about each trial they are testing. We define a control of variables strategy as changing a single variable at a time. However, due to the presence of the table when students are using the interactive simulation, they can look across all trials they have run and therefore trials that employ a control of variables strategy do not necessarily have to be in order. For this reason, we look across all trials a student produces when looking for a control of variables strategy.

We use feature engineering to develop a variable, COV Trials, which represents the number of trials a student ran using the control of variables strategy. For example, if a group generates 5 rows in the table and is controlling variables in each of the trials run, the group would have run 4 trials. This is because controlling variables requires that groups run more than one trial so that they can compare across two or more trials. We do not count the first row generated in the table as a trial, since if there were only one row generated the group would not have the chance to employ the control of variables strategy. For this reason, we look at how groups use the control of variables strategy as a part of the entire corpus of students and as part of the group that used the table scaffolding.

We can also transform COV Trials into a dummy variable, using 1 for any students who used the strategy (no matter the value in COV Trials) and 0 for those students who did not. We call this dummy variable COV. Overall, 137 (55.69%) of the 246 groups employed a control of variables strategy in the first iteration of the simulation. There were 216 groups that used the table scaffolds in the first iteration of the simulation to generate at least one row of data. When we narrow our focus to only those groups that generated at least two rows in the table during the first iteration of the simulation, there were 115 groups remaining. Of these 115 groups, 103 groups (89.56%) employed a control of variables strategy.

Other research has found that students using simulations generally fall into one of three groups: intentional, unsystematic, and exhaustive (McElhaney & Linn, 2011). Our groups also seem to fall into one of three similar categories: those that did not use the table scaffolding but did use the simulation and other scaffolding (N = 30), those that only generated one row of data in the table scaffolding (N = 101), and those that generated two or more rows in the table scaffolding and generally employed a control of variables strategy (N = 115). While students in this study did not seem to be "exhaustive" in conducting trials, student groups that ran zero or one trial could be considered unsystematic since they were not using the model to test more than one option, and groups that ran two or more trials could be considered intentional, especially when they are systematic (high percent of controlled variable trials) in their exploration of the simulation. All groups used at least some of the provided scaffolding and the simulation to design their ovens. Many of these groups used the table scaffolding, and even though they were not instructed to control variables when running trials, they did. The significant factor in these groupings is the number of rows students generated using the table scaffolding. However, the number of trials a student runs may not tell the whole story for the process of learning while using the simulation.

Effect on learning

Using pretest and posttest scores we aimed to understand the effect of actions with the simulation on learning. We used the variable COV Trials as well as the dummy variable COV. We also developed dummy variables based on the three groups found in the previous section. First, we aimed to understand the role of the number of rows of data a student generated using the table scaffolding on learning. We found that the number of rows generated in iteration 1 of the simulation is a significant predictor of individual posttest scores, when controlling for pretest scores and curriculum group ($b = 0.10$, $t(546) = 2.68$, $p < 0.01$).

Next, we wished to understand the impact of controlling variables on learning. We found that the number of COV Trials run in iteration 1, however, is not quite a significant predictor of posttest score, when controlling for group and pretest score ($b = 0.06$, $t(546) = 1.63$, $p = 0.10$). In addition, using the dummy variable COV does not significantly predict posttest scores when controlling for pretest scores and group ($b = 0.005$, $t(546) = 0.13$, $p = 0.90$). Together, these results indicate that perhaps the control of variables strategy, while a good practice in science, is not as helpful for developing an understanding of the scientific principles at play through a simulation. However, this variable is flawed, especially with consideration for this specific data set. The COV variables (dummy and count) require that students ran 2 or more trials, since controlling variables requires that the student is controlling with respect to a previous trial. Because of this, the variable is confounding. In a data set with a larger number of students who ran two or more trials (we have this data from high school classes), this would not be as problematic.

More experimentation using the model is beneficial for developing a better understanding of the scientific concepts. We split the students based on their actions during the simulation step (did not generate any rows in table, generated one row, generated two or more rows). We found that generating two or more rows in the table during iteration 1 significantly predicts posttest scores, when controlling for pretest score and working group ($b = 0.12$, $t(546) = 3.11$, $p < 0.01$), though generating no rows or 1 row were not significant predictors.

We also developed a variable, Percent Systematic, that is the percentage of the total rows a group generated that used the control of variables strategy. A histogram of this variable is shown in figure 3.15. This variable has the ability to show more nuance in how students were employing the control of variables strategy, but was also not predictive in determining posttest scores, when controlling for pretest and group id ($b = 0.05$, $t(508) = 1.32$, $p = 0.188$).

There were also two short-response scaffolding questions. We generated a variable based on the number of questions students answered (0, 1, or 2). This was predictive of posttest score, when controlling for pretest score and group id ($b = 0.10$, $t(546) = 2.56$, $p = 0.011$). This indicates that students who attempted to answer more questions during this step of the unit made greater learning gains from pretest to posttest.

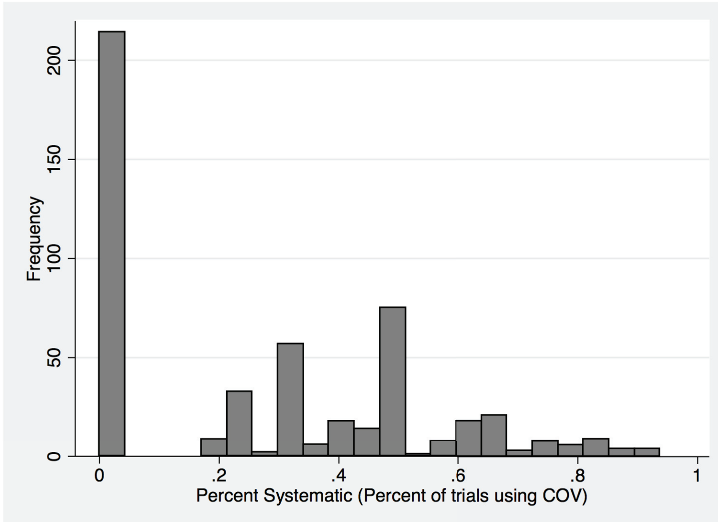


Figure 3.15: Histogram depicting the frequency of Percent Systematic variable (Mean: 0.275)

Overall, evidence suggests that students should be encouraged to experiment with the model and guided to produce at least two rows of data in the table to improve learning outcomes and use the short response questions. Perhaps changing more than one variable at a time in this type of environment indicates that students are spending more time thinking about possible outcomes. However, asking students to conduct a certain number of trials before moving on may actually decrease success or motivation (McElhaney, et al., 2011), so any guidance given to students about the length and type of interactions they have with the simulation should be developed carefully so as to maintain student feelings of autonomy.

Limitations

While we have found simulations to be beneficial for student learning in previous work (McBride, Vitale, Applebaum & Linn, 2016), it is important the note that not all student learning is due to interactions with the simulation. Student interactions with the simulations generally take about two class periods, while the entire curriculum takes 10-15 class periods. There are many other learning opportunities, and the curriculum is designed to take advantage of students' outside knowledge. While there is likely some difference between students who generated one row versus those who generated two or more rows, it is difficult to understand the differences between using a control of variables strategy and generating multiple rows of data in the table because the control of variables strategy, as we have defined it, confounds the number of trials variable.

Another limitation to using a control of variables strategy as a factor in our analysis is that students may not be intentionally using the strategy while interacting with the simulation. The design of the simulation makes it easy for students to change only one thing between trials, making it seem like they are using a control of variables strategy. However, this may be because students want to do less work and so only change one variable at a time. Also, since we look across all rows generated for the control of variables strategy, students who generate many rows of data may seem like they are using the control of variables strategy, even though they were not.

Discussion

The results of this study give insights for further research and development of simulation scaffolding. This data suggests that the table scaffolding assists students in using a control of variables strategy, even though this strategy was not discussed in the curriculum or by the teachers. However, the control of variables strategy may be too limiting and not as beneficial for student learning as a more exploratory strategy. Students should ideally be guided to generate at least two rows in the table. The table scaffolding provides practical benefits to students, in that it allows them to save the data from trials they have already run instead of relying on their memories to make design decisions (Reiser, 2004). While directing students to run a certain number of trials may not lead to greater learning gains (McElhaney, et al., 2011), developing guidance that directs students to explore the model further or use other scaffolds before moving on in the project may still be useful. Future data may also be more enlightening about the effect of the control of variables strategy on learning. If a higher percentage of students in a data set have run two or more trials, we can evaluate the impact of the control of variables strategy on learning gains more effectively since it will not be impacted so significantly by students who run zero or one trial.

Conclusions

This chapter touches on a number of critical issues surrounding the use of computer models as educational tools. Namely, how students use computer models within a science context. Together, these studies point to the need for further scaffolding within curriculum that uses computer models. For example, it is still necessary to encourage students to use evidence from the trials they run using the model while making the designs for their physical ovens. Tightening the links between the physical and virtual is a next step for the solar ovens model.

The study using the global climate change curriculum shed light on the numerous ways students use virtual models. The data collected for that study showed that dividing students into categories or groups based on their "actions" might be more difficult than we believe it to be, since students may exhibit different types of action sequences (e.g., watching, exhaustive testing, goofing off) at different points during their use of the model. Therefore, developing scaffolding that encourages students to use the model in ways that will encourage learning is more useful than developing guidance based on student action sequences.

The first study in this chapter showed that students who spend time watching the model made greater learning gains overall. Using this finding to develop scaffolding for watching the model in the solar ovens virtual model, we were then able to study even further how students' trials impacted their learning.

Students do not necessarily need to control variables while using models to make careful decisions, but instead should be conducting multiple relevant trials. This holds true for both science and engineering contexts. When using models to consider or learn about a complex scientific phenomenon, like climate change, it is important that students make careful choices and then attend to what is happening in the model. Without attention, many of the nuances of the complex interactions may be lost. When using computer models to make engineering design decisions based on scientific phenomena, it is equally as important to carefully consider what is happening in the model, as well as the overall outcomes.

In this work, we built up scaffolds to allow students to attend to the model while saving the data from the trials they ran. This was a positive change, but students must still be encouraged to use models to engage in exploration. In the second study of this chapter, many students still ran only one trial. Perhaps developing a more complex model with more options that are realistic to the choices students will have to be making during design would encourage further exploration in the model.

Chapter 4: Student Opinions about the Affordances of Physical and Virtual Models

Introduction

When designing inquiry curriculum, we often use interactive virtual models to allow students to investigate how variables in a system may be related. This study compares student perceptions of virtual and physical models. Models improve student learning by abstracting away unnecessary features and making invisible phenomena more visible (Snir, Smith, & Grosslight, 1993; Wilensky & Reisman, 2006). Virtual models are now quite ubiquitous, due to their inclusion in the Next Generations Science Standards. However, their ubiquity is still relatively new, and while there is a wealth of research on supporting student learning of science concepts through the use of virtual models, not much work has been done to explore how students understand the practices involved in utilizing these virtual models. We elicit students' ideas about the relative affordances of physical and virtual models before and after a curriculum in which they use an interactive virtual model to design a solar oven, which they then build and test in the physical classroom environment.

In the Next Generation Science Standards, using models is a skill that spans grade levels and topics and is one of the science and engineering practices (NGSS Lead States, 2013). Many studies have found students learn concepts, inquiry skills, and scientific practices at the same level or at a higher level through the use of virtual laboratories (versus physical laboratories) (e.g., De Jong, Linn, & Zacharia, 2013; Brinson, 2015). Virtual models allow students to develop their own knowledge by asking scientific questions, answering those questions using evidence, developing explanations, and connecting explanations to scientific knowledge (Olsen & Loucks-Horsley, 2000). However, students may develop additional questions relating to the practice of using models (e.g., why do scientists use models? What can models tell us? What are the limitations of models?).

We use the knowledge integration framework to develop the curriculum about solar ovens, because the framework focuses on building coherent understanding (Linn & Eylon, 2011). The knowledge integration framework has proven useful for design of instruction featuring dynamic visualizations (Ryoo & Linn, 2012) and engineering design (Chiu et al., 2013; McElhaney & Linn, 2011). The framework emphasizes linking of ideas by eliciting all the ideas students think are important and engaging them in testing and refining their ideas. We provide students with multiple sources of ideas (physical and virtual models) while they are developing their ideas about design and how energy transformation works. Helping students to distinguish which sources of evidence are relevant and supportive (or not) of their ideas is a particular challenge for instruction. In order to develop relevant instruction, we need to know how they naturally think about the relative affordances of each type of model.

We use virtual models in our curriculum to help students understand the interplay between science concepts (like energy transformation) and engineering design decisions.

Supporting students in making connections between these two areas is important, given that research on this topic has shown that many projects do not live up to the claim that engineering projects improve student achievement in mathematics and science (Teacher Advisory Council, 2009). While hands-on projects may generate more student interest and engagement (Hmelo et al., 2000; Cantrell et al., 2006) than typical science curricula, they often fall short on developing science concepts. Studies show that a blended approach, combining physical and virtual, can be more effective than either method alone (Jaakkola, Nurmi, & Lehtinen, 2010; Jaakkola, Nurmi, & Veermans, 2011; Olympiou & Zacharia, 2012).

We found that students can come to understand the benefits virtual models hold over physical models after using both in the curriculum. However, students maintain that physical models are more accurate than virtual models. We also found that virtual models can be connected with "the internet" in students' minds, and may seem inherently untrustworthy. Students have likely had prior instruction about responsible use of information from the internet that has led them to develop skepticism about online sources. We should build on that healthy skepticism about online sources, including our virtual model, to help students distinguish between useful features of virtual models as well as limitations.

Work has been done on how opinions of technology impact adoption and use of the technology (Ertmer, 1999; Ertmer, 2005). Ideas about what the technology can or cannot do impacts how people use the system, and therefore what they are able to learn while using the system (Jackson, et al., 2009). In addition, students may be unfamiliar with the ways they should use models in order to gain information about complex causal relationships, and thus develop non-normative ideas (Perkins & Grotzer, 2005).
This work examines what students find to be useful features of physical and virtual models, and explores how this impacts learning. Having an understanding of student ideas around the practice of using scientific models will help us develop guidance in our curriculum for students to further consider relative affordances of physical and virtual models, which ties in to the evidence they use in making design decisions.

Methods

Materials and data source

Five teachers from three different schools participated in this study, along with their students (*N*=640). All students were in the 6th grade, and all schools are in the suburbs of a large U.S. city serving mainly middle SES communities. Teachers A (*N*=137) and B (*N*=80) teach at school A (42% free and reduced lunch eligible, 13% English language learners), teachers C (*N*=190) and D (*N*=78) teach at school B (31% free and reduced lunch eligible, 8% English language learners), and teacher E (*N*=155) teachers at school C (32% free and reduced lunch eligible, 5% English language learners). Students completed pre- and posttests individually; during the curriculum, students worked in pairs or triads.

Curriculum

This study was implemented in a curriculum unit entitled *Solar Ovens*. The goal of the unit was to familiarize students with the way energy transforms from solar radiation to heat using a hands-on project and interactive models, covering the modeling aspect of the Science and Engineering Practices of the NGSS, as well as the standards associated with energy (NGSS Lead States, 2013). Students engaged with the curriculum in WISE (Web-based Inquiry Science Environment), utilizing a variety of instructional and assessment tools (Linn & Eylon, 2011).

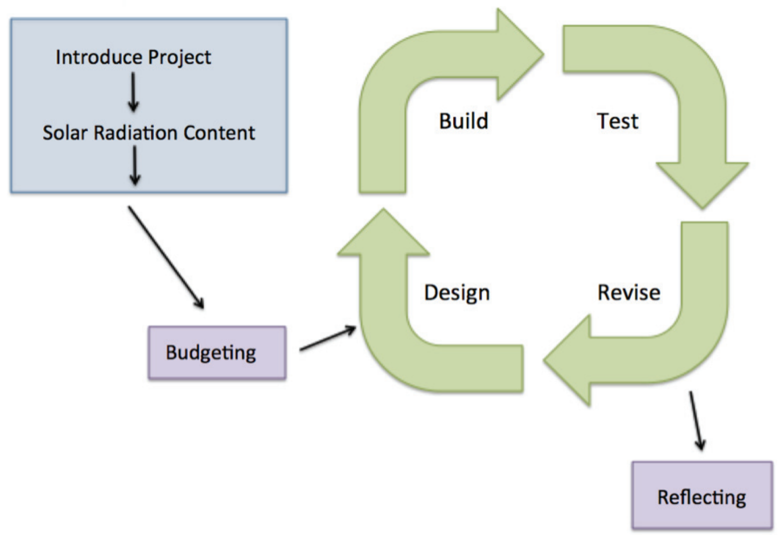
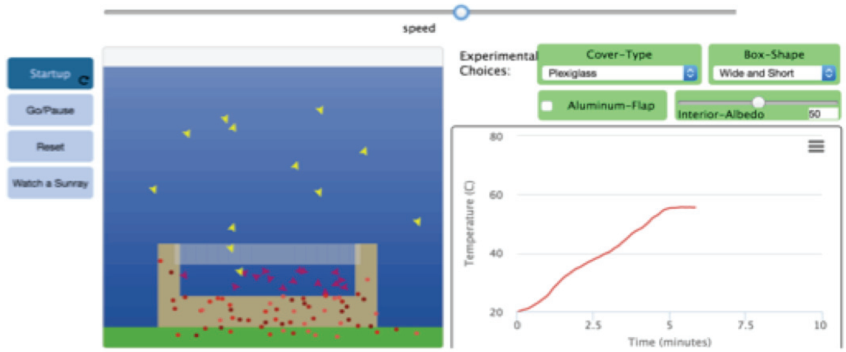


Figure 4.1: The outline of the solar ovens curriculum. Students go through the "design, build, test" cycle two times.

Students followed the design, build, test cycle with two iterations (Figure 4.1). During each design phase, students use the interactive model to test different features on a virtual solar oven. This model is shown in Figure 4.2. Often when students build a physical model they neglect the scientific basis for their decisions (Crismond, 2001), instead focusing on details of construction that may be superficial. Prior work has shown the computer model developed for this unit helps students make connections between science concepts and decisions made during the design process (McBride, Vitale, Applebaum & Linn, 2016).



#	Box Shape	Cover Type	Flap?	Int Albedo	Temp at 0m	Temp at 1m	Temp at 5m	Temp at 10m
1	Wide and S	None	No	50	20	24	30.2	31.5
2	Skinny and	None	No	50	20	26.2	29.7	31.1
3	Wide and S	Plexiglass	No	50	20	25.6	55.4	53.4

Figure 4.2: The interactive model built to help students during the design process of the curriculum.

Assessment

First, we aim to find ideas students hold about the benefits and drawbacks of physical and virtual models. We do this by assessing student responses to a pre-/posttest question called *David's Claim*. This question (Figure 4.3) asks students to help a fictional student, David, decide whether the box he will use for his solar oven should be tall and skinny or short and wide. Students are told that David thinks the tall and skinny box will heat up faster because the window on the top is smaller and will let less energy leave the box. Students are then asked whether David is correct or incorrect, and to explain their answer using evidence from the interactive model (where they can only manipulate box shape).

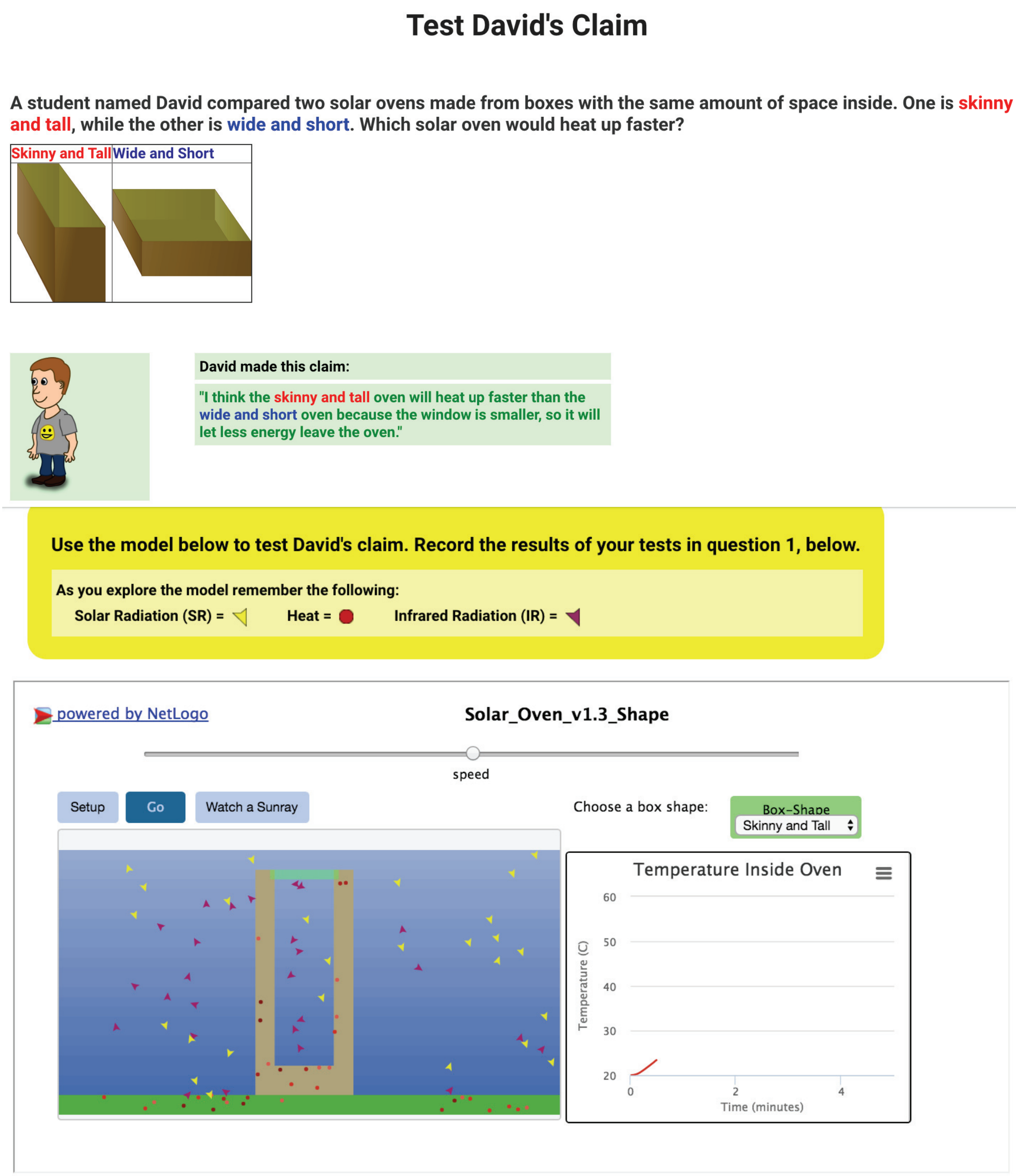


Figure 4.3: Screenshot of the *David's Claim* pretest and posttest item and accompanying interactive model

To understand student thoughts about affordances of physical and virtual models, we analyze a follow-up question, asking students whether they would rather use a physical or virtual model to help David (item: *Opinion*). While more students preferred to use the physical model at both the pretest and the posttest, there was a shift toward students preferring to use the virtual model or to use both models at the posttest (Table 4.1).

Table 4.1: Students who would prefer to use a physical or virtual model (or both) to help explain a science and design problem to another fictional student in the *Opinion* item

	Pretest	Posttest
Physical	333	317
Virtual	176	193
Both	6	15
No Answer (Unclear Answer)	31	21

We found that students had ideas about benefits (or drawbacks) of physical and virtual models that fell in the 10 categories, shown in Table 4.2. After dropping students who did not answer the *David's Claim* question on both the pretest and posttest, we were left with 558 students. Under this coding scheme, students could have ideas falling into multiple categories, but the vast majority of students only fell into one category. A breakdown of responses is shown in Table 4.4.

Table 4.2: 10 categories for student ideas about the relative affordances of physical and virtual models, with description and variable name.

Variable Name	Description
V_accurate	Virtual models are more accurate or valid than physical models
P_accurate	Physical models are more accurate or valid than virtual models
V_visible	Virtual models have features that make them easier to use for explaining and learning how energy works (e.g., graphs, depiction of energy using symbols)
P_visible	Physical models are better because you can see them and see what's happening from any angle
V_fun	Virtual models are more fun than physical models
P_fun	physical models are more fun than virtual models
V_fasteasy	Virtual models are faster and easier to use than physical models, and do not require materials
P_fasteasy	Physical models are faster and easier to use than virtual models
P_experience	Physical models give a better experience, and help students learn and focus better
V_limitations	Virtual models have limitations and cannot show all the options that exist in real life

Table 4.3: Examples of student responses in each coded category for the *Opinion* item

Category	Examples
V_accurate	• I would rather use the computer model because it will give an accurate answer. • I would rather do a computer model because the weather will change in real life. It would be hard to find the perfect day to test the solar oven. Also, the clouds can block the solar radiation in the middle of the process.
P_accurate	• I would like to do a real experiment because he might say it's the internet never trust the internet!"." • I would rather do it in real life because you can get more accurate information; a computer model doesn't actually do the experiment.
V_visible	• I think the computer model will help David more because it shows the heat that is getting trapped in the box. And what type of heat is coming from the sun to the box; then what type of energy is going back to the sun. • I will rather do it in a computer model than with a real solar oven because on the computer model it gives you a graph and it shows you which one will heat up faster.
P_visible	• I would rather do an experiment with a real solar oven to help David because doing it on computer isn't really going to help him understand the energy in both of their ovens and how his claim was wrong. Also, doing an experiment with a real solar oven could help him because he could see how the energy could be transferred into the box. • I would rather use a real solar model so i can see it more clearly.
V_fun	• I would rather use a computer model than do an experiment with a real solar oven because its less dangerous and little bit more fun.
P_fun	• i would do both but in a real experiment it will more funner and cooler cause we can really do it on are own
V_fasteasy	• Computer model because it's more easier not to build and using your time shortly. • I would do a computer model because it's an easier way to get data than doing the actual thing.
P_fasteasy	• I would so a experiment with a real solar oven because it would be easier for me to do. • I would do an experiment with a real solar oven because I think I would find it easier than to do it on a computer, and also so you can actually tell whether the experiment works or whether it

	doesn't.
P_experience	• I would rather use a real solar oven to help David. I would use a real solar oven because in my opinion, it is easier to learn with a real life model. This is because you can actually see it happen instead of watching a fake model in the computer. It is easier for me to learn with the real life model because I can understand it more. • I would enjoy doing an experiment more because I'm not to fond of the computers that we use here and I prefer to use my hands and do projects were you build stuff.
V_limitations	• An expirement because people's enviorments have different temperature. • I would rather use a real solar oven because it would be much more accurate than a computer model since a computer model can't always show the current weather and what the outcome would be on a cloudy day.

Table 4.4: Number of ideas students present in each category at pretest and posttest on the *Opinion* item

	Idea	Pretest	Posttest
Virtual	V_visible	96	102
	V_fasteasy	68	70
	V_accurate	31	40
	V_limitations	6	22
	V_fun	4	1
Physical	P_accurate	153	171
	P_visible	84	75
	P_experience	64	52
	P_fun	35	32
	P_fasteasy	11	10

Next, we group these categories into "science ideas" and "non-science ideas". The science ideas include V_accurate, P_accurate, and V_visible, and V_limitations. All others fall into the "non-science ideas" group, which include non-normative ideas as well as practical or irrelevant ideas (e.g., the model is fun). A breakdown of how many ideas were in each category can be found in Table 4.5.

Table 4.5: Breakdown of student "science" and "non-science" ideas on *Opinion* item

	Pretest	Posttest
Avg. # Science Ideas per student (std. dev)	0.52 (0.54)	0.62 (0.56)
Avg. # Non-Science Ideas per student (std. dev.)	0.48 (0.55)	0.44 (0.56)

Table 4.6: Knowledge Integration rubric for an item asking students to explain what would happen to the temperature inside a car, if the car was left in the sun on a cold day

Score	Level	Examples
0	No Answer	
1	Off Task	*I don't know.*
2	Irrelevant/Incorrect	*The inside air and the outside air are the exact same temperature because the sun is not enough to heat the inside if the car.*
3	Partial Normative isolated ideas without a valid link	*The solar radiation would go through the metal and would stay in the car when the outside air wouldn't be able to get inside.*
4	Basic Elaborate a scientifically valid link	*it would be warmer than the outside air because if the car hasn't been driven for a week and its been in the sun the whole time the car will absorbe the heat and scence there is know way the heat can get out of the car the heat will just keep building up.*
5	Complex Elaborate two or more scientifically valid links	*The sun produces solar radiation which heats up the car and the infrared radiation gets trapped in the car which leads to the temperature rising.*

To assess student progress, we used written responses from pre- and posttests. The pre- and post-test assessments measure student ability to link concepts about energy and design. Each item on the pre- and posttest was scored using a knowledge integration rubric that evaluates the number of normative links students make for scores between 1 (low) and 5 (high). A sample rubric can be found in Table 4.6. The pre- and posttests were made up of 7 items each, with the same items appearing on both tests.

Results

By far the most common responses to the *Opinion* item were that students felt the physical model was more accurate than the virtual model and students found value in the features of the virtual model (e.g., the graph, the visualization of the phenomenon). From pre- to posttest, students' belief in the greater accuracy of the physical model compared to the virtual model increased, as did student appreciation for the visibility offered by the virtual model. Student reasoning about the limitations of the virtual model also increased drastically from pretest to posttest. In contrast, student statements about the benefits of the visibility of the physical model decreased, as did student statements that the experience of building a solar oven would help them to answer the question. This data is shown in Table 4.4, with relevant examples in Table 4.3.

In relation to learning, students with a higher number of science ideas on the *Opinion* item at posttest also made significant gains from pretest to posttest on the *David's*

Claim item ($p<0.001$, $\beta=0.237$). This fits with the knowledge integration framework and goals of the curriculum, since we designed the curriculum to help students add new ideas to their repertoire (increasing the number of science ideas) and to distinguish among their ideas to develop a coherent understanding.

In addition to this, we explored relationships between students' perceptions of models and learning. Students who wrote that the physical model was more fun at the posttest were significantly less likely to make learning gains from pretest to posttest ($p<0.01$, $\beta=-0.103$). Students who identified limitations of the computer model at the posttest were significantly more likely to make learning gains ($p<0.001$, $\beta=0.154$). We look at these two groups of students because they represent two levels of interpretation (limitations: deep, fun: shallow) of the relative affordances of the models.

We ran a regression model to predict a posttest score on *David's Claim,* controlling for the pretest score and using the ideas defined earlier as science-relevant as covariates (V_visible_post, P_accurate_post, V_accurate_post, and V_limits_post), and find that each variable except V_accurate_post is significantly predictive of student learning (V_visible_post: $p<0.001$, $\beta=0.223$; P_accurate_post: $p<0.001$, $\beta=0.178$; V_accurate_post: $p<0.214$, $\beta=0.050$; V_limits_post: $p<0.01$, $\beta=0.101$). Citing the accuracy of the virtual model may be less important for student learning if the student relies too much on the model and does not critically evaluate the findings and whether they fit with results from the physical model.

Conclusions

This work takes advantage of a curriculum featuring both physical and virtual models of solar ovens to explore benefits and drawbacks. Moreover, the virtual model acts as a representation of the physical model students will build, providing students with an opportunity to explore the relative affordances of each type of model.

This study reveals limitations in student understanding of virtual models that deserve attention to increase effectiveness of instruction featuring models. Focusing on increasing student consideration about *relative* benefits of physical and virtual models could help students appreciate virtual models and have lasting impact on their use of models. Having a greater understanding of student ideas around the practice of using scientific models will help us to develop guidance in curricula for students to further consider relative affordances of physical and virtual models, which ties in to the evidence they use in making design decisions. This style of blended laboratory, combining physical and virtual models, is useful for helping students develop both practices and conceptual understanding and is likely to be practiced in both educational and work environments. Thus, encouraging students to consider and understand the relative affordances of various types of models and representations is important.

This work points to many improvements that must be made within science and engineering projects for the goals of the NGSS to come to fruition. Further focus within curriculum on the benefits and limitations of computer models, and further scaffolding for productive use of computer models towards learning goals will help develop student proficiency in a number of science and engineering practices, not just "using models". This focus will also help students in planning and carrying out investigations, analyzing and interpreting data, designing solutions, and engaging in argument from evidence.

Chapter 5: Promoting Knowledge Integration using Automated Assessment of Student Ideas

This chapter investigates a typical phenomenon in science education. Students develop a set of complex and often incompatible ideas. This is also consistent with research on knowledge integration, showing that students hold many disparate and, often, conflicting ideas (Linn & Eylon, 2011). We explore the ideas students write about light propagation in the context of an online and hands-on unit and investigate a way to automatically assign guidance on student writing in order to improve reasoning. We focus on the consistency of students' ideas about how light "attracts", showing that students are inconsistent in their ideas about how light interacts with different types of objects (e.g., shiny or dark-colored).

We identify a complex sequence that some students articulate. This sequence involves arguing that it is necessary for a material to first "attract" light so that it can reflect the light. This non-normative idea may hinder future learning about energy in the context of solar ovens. These ideas cannot be investigated or falsified using the interactive model discussed in earlier chapters, so a new approach is needed. We explore a method for automatically scoring student written responses using natural language processing and machine learning. Using this method, we can score for both the presence of the "attract" idea and the completeness of the response. We study how we can use this information to provide automatic guidance to the student based on their written response. This research can inform refinements to the curriculum unit on solar ovens.

Introduction

Writing is an important activity in science because it can reveal confusions and help students to develop a more complete understanding of complex processes. The writing process can help students to express their ideas, generate new ideas, link their ideas together, and distinguish among their ideas.

Written response questions require students to generate their own ideas while developing responses. Multiple-choice questions are commonly used in assessment, but only require students to recognize the correct answer from a list. Questions that require students to write their own responses are better for assessing what students know, since it is more difficult for students to simply guess the correct answer when they must generate the answer themselves. Research has also found benefits for student learning when students generate their own explanations instead of reading text (DeWinstanley & Bjork, 2004; Richland, Bjork, Finley & Linn, 2005).

An important aspect of any curriculum is to guide writing so that students are prompted to integrate their ideas. Standards now call for coherent explanations of scientific phenomena (NGSS Lead States, 2013; National Research Council, 1996), and written explanations are useful in classrooms to give teachers a deeper understanding of their students' ideas and where they come from.

We examine student non-normative and inconsistent ideas in a hands-on engineering project. Hands-on projects may generate more student interest and engagement (Hmelo et al., 2000; Cantrell et al., 2006) than typical science curricula, but oftentimes they do not improve student achievement in mathematics and science (Teacher Advisory Council, 2009) because they do not help students develop an understanding of science concepts. This work focuses on particular assessment items that connect solar ovens (the output of the hands-on project) to the science phenomenon that explain how it heats up, with the goal of strengthening science learning and assessment in hands-on projects. We focus on understanding student non-normative ideas around a particular idea, that energy can "attracted" to shiny or dark objects. Then we show how we can develop classifiers using natural language processing and machine learning to score this particular idea, as well as the cohesiveness of the response.

The knowledge integration framework emphasizes linking of ideas by eliciting all the ideas students think are important and engaging them in testing and refining their ideas (Linn & Eylon, 2011). This framework has proven useful for design of instruction featuring dynamic visualizations (Ryoo & Linn, 2012) and engineering design (Chiu et al., 2013; McElhaney & Linn, 2011). The knowledge integration framework also offers theoretical explanations on the importance of student writing; students can hold multiple conflicting views, especially considering their classroom and everyday knowledge (Linn & Hsi, 2000; Linn, Lee, Tinker, Husic & Chiu, 2006). Instructional designers cannot capture all the possibilities for student ideas or misconceptions in multiple-choice questions. The knowledge integration framework (Linn & Eylon, 2011) takes into account the wide variety of ideas students hold about a phenomenon and supports linking of ideas by first eliciting all the ideas students think are important, then engaging them in exploring their ideas. The instruction used in this research was developed using the knowledge integration framework, and aims to help students integrate their ideas about energy.

We score student written responses using knowledge integration rubrics that evaluate how well students are able to integrate a variety of ideas. These rubrics have proven useful in understanding how well students understand a given topic (Linn & Hsi, 2000). However, they present a challenge for automatic scoring since students use a wide variety of ideas in responses. For example, in response to a question about how reflective shields for car windows work, one student wrote "...since it becomes the same color as the sun when it's shinny surface reflects it". A better answer might have been something like " ...it will reflect the sun and keep the car cool. The not shiny side would absorb the heat and make the car hot". While both students demonstrated that they understood that reflection was taking place and that something from the sun (light) was being reflected there were also other ideas present. Since short response questions can collect a wide variety of student ideas that are very nuanced, they can be difficult to assess automatically. Scoring responses by hand requires a great deal of time and effort. Often, building a model to automatically score a particular question requires approximately one thousand scored student responses, more than any single teacher could generate in school year by themselves. Building models to detect particular non-normative ideas that may hinder future learning may be useful for helping students to revisit those ideas.

One such idea is that of "attraction": that light is attracted to shiny things (like tin foil) and heat can be attracted to dark-colored things (like black construction paper). This idea is both common enough in student writing and may hinder student understanding of the actual phenomenon that it is worth addressing. If students believe that light can be attracted, or actively pulled in, by a material, they may not ever develop an understanding of the actual processes involved: reflection and absorption. However, using this specific wording may indicate a number of possible understandings. For some students, using the word "attract" may be an easy way to avoid describing the actual process, especially if they feel unsure about it. For others, this may be a vocabulary issue wherein they do not understand that "attract" is not a word used to describe physical processes. Others may even believe that the foil or black paper are drawing the energy to it without understanding the implications this would have on the greater system. Before developing automated scoring and guidance for students with this idea, we first must find out where this particular idea may be coming from.

In this paper, we first examine student ideas about energy and how think about reflection and absorption, and when and how they use the term "attract". Since these ideas are often brought up in written responses, we then discuss the development of an classifier that uses natural language processing and machine learning to automatically score student responses based on a rubric set by researchers. In this case, we score for the presence of the non-normative idea of "attraction", as examined in the first half of the paper. Indeed, the ideas we investigate here are difficult for students to investigate and falsify while using the interactive model, so automatically assessing their writing is likely the best way to detect the idea and provide students further resources to think about.

The design of guidance for students could be the topic of another entire chapter. However, this work aims to inform automated guidance for students by building on research that has already been done in this area (e.g., Tansomboon, Gerard, Vitale, & Linn, 2017; Gerard, Ryoo, McElhaney, Liu, Rafferty, & Linn, 2016). These papers use knowledge integration to guide students in considering making revisions to their writing or in revising their ideas about scientific mechanisms.

Curriculum

This work was implemented in a curriculum unit entitled *Solar Ovens and Solar Radiation* (referred to as *Solar Ovens*). The goal of the unit is to familiarize students with the way energy transforms from solar radiation to heat using a hands-on project and interactive models, covering the modeling aspect of the Science and Engineering Practices of the NGSS, as well as the standards associated with energy (NGSS Lead States, 2013). Students engaged with the curriculum in WISE (Web-based Inquiry Science Environment), utilizing a variety of instructional and assessment tools (Linn & Eylon, 2011).
Students follow the design, build, test cycle with two iterations in the curriculum. An outline of the curriculum is shown in Figure 5.1. Throughout the curriculum, students use interactive simulations to design their solar ovens (Figure 5.2). These simulations may provide students with a framework for thinking about how, when, and where reflection and absorption happen in their ovens. Students were allowed to use only a certain set of materials in addition to a cardboard box (Figure 5.3). Throughout the project, students responded to short answer questions about the choices they made in their design, how those choices played out, and how their design used and transformed solar radiation. The curriculum took between 2-3 weeks (10-15 hours) and students worked together in groups

of 2-3.

While the computer models allow students to investigate how energy and light transform to heat up a solar oven, they do not allow students to falsify the idea that light is attracted to shiny or dark objects. The random nature of light coming from the sun is modeled, but it is an underlying feature of the computer model that students may not attend to while they are investigating the model. Indeed, to make sense of what is happening in the model they do not need to attend to the underlying rules the model operates upon. The model may even reinforce some students' ideas that light is attracted to shiny or dark objects since the light in the model goes to those objects (as it goes to all objects within the modeling environment).

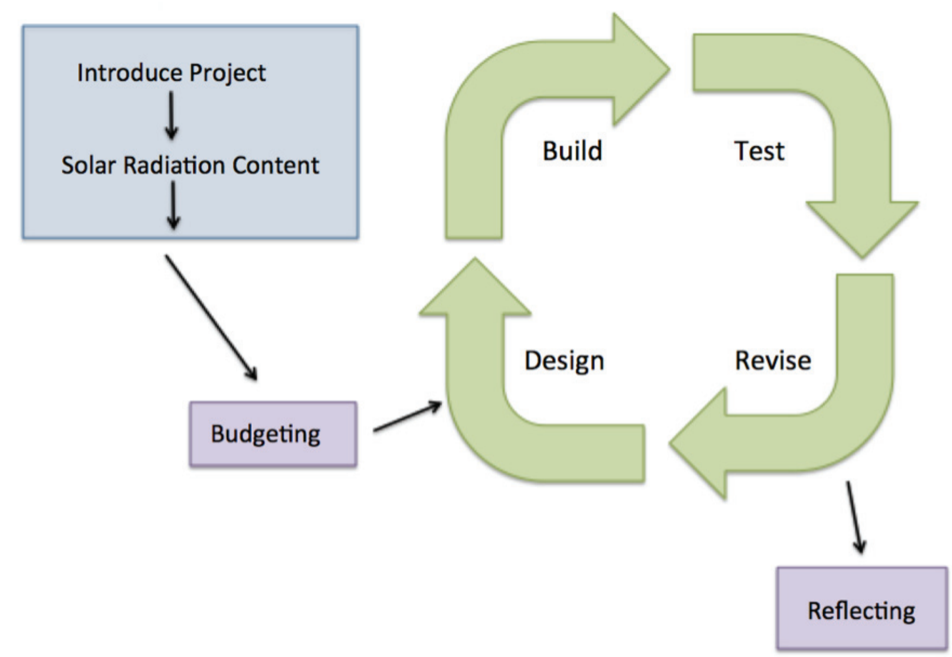
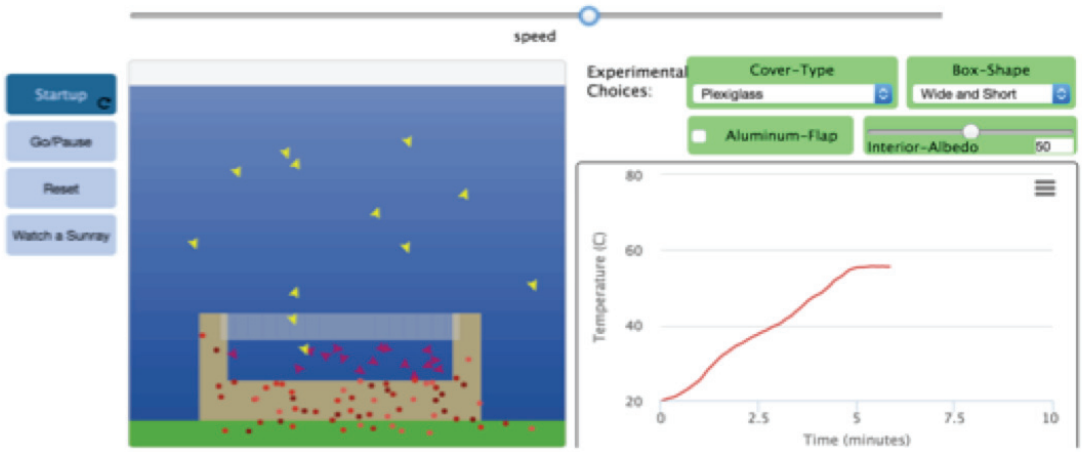


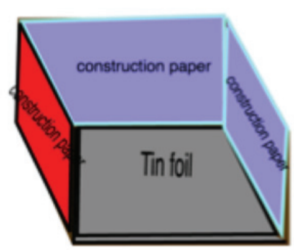
Figure 5.1: The outline of the solar ovens curriculum. Students go through the "design, build, test" cycle two times.



#	Box Shape	Cover Type	Flap?	Int Albedo	Temp at 0m	Temp at 1m	Temp at 5m	Temp at 10m
1	Wide and S	None	No	50	20	24	30.2	31.5
2	Skinny and	None	No	50	20	26.2	29.7	31.1
3	Wide and S	Plexiglass	No	50	20	25.6	55.4	53.4

Figure 5.2: Interactive simulation used by students to design their solar ovens

Description:
The construction paper and will absorb the solar radiation and the tin foil will both absorb and reflect the solar energy.
enlarge



1 sheet of construction paper (8.5 inch x 11 inch)	$2
1 sheet of tin foil (12 inch x 12 inch)	$7
1 sheet of plastic wrap (12 inch x 12 inch)	$6
1 sheet of plexi glass (12 inch x 12 inch)	$10
Tape (3 feet)	$3

Materials & Costs

1 sheet of construction paper (8.5 inch x 11 inch)	$2
12 inch x 12 inch Sheet of tin foil	$7
12 inch x 12 inch Sheet of plastic wrap	$6
You can RENT: 12 inch x 12 inch sheet of Plexiglas (thick plastic)	$10
3 feet of tape (Duct, masking, clear)	$3

Figure 5.3: Right: Example of a student design with a description of the oven; Left: Materials available and budget from Solar Ovens curriculum

Student Ideas About "Attracting" Light

Students often have a variety of ideas about scientific phenomena, some of which are inconsistent with the others (Linn & Eylon, 2011). A goal of many science curricula is to help students organize their ideas and develop a coherent understanding of the phenomena being studied. The Next Generation Science Standards recommend that students study similar content across grade levels and in multiple contexts (e.g., solar energy is a theme throughout much of 6th grade science; NGSS Lead States, 2013).

To develop this coherent understanding, students must first grapple with the ideas they already have (White & Gunstone, 1992), then distinguish among their ideas to decide which are relevant and correct. However, this can be a challenging and time-consuming process. In addition, there are a variety of factors that impact how students conceptualize and describe processes (Minstrell, 1982; Zee & Minstrell, 1997). When students have non-normative ideas about a process, there can be a learning opportunity as students develop skills and experience in falsifying their original ideas. However, students may be unaware of the implications of their vocabulary choices in the descriptions they write, or may be unable to falsify non-normative ideas using the tools at hand.

In the first half of this chapter, we examine the non-normative and contradictory ideas students have about energy transformation in a hands-on project. We find that students have a variety of ideas about how reflection, attraction, and absorption work in a solar oven, and that many students are inconsistent in their ideas about how energy interacts with the materials in a solar oven. This may indicate that they are actively distinguishing their ideas through the process of writing, but may also mean that further supports are required to help students reason about energy, especially where potentially new vocabulary are involved. This is especially important since, as discussed earlier, the computer model may not provide the resources for students to falsify the ideas they have about attraction, and it may even reinforce those ideas if students do not attend to the model carefully.

We used the knowledge integration framework to create the curriculum about solar ovens, because the framework focuses on building coherent understanding (Linn & Eylon, 2011). The framework offers instructional design principles to enhance connections between design decisions and scientific principles. The framework emphasizes linking of ideas by eliciting all the ideas students think are important and engaging them in testing and refining their ideas. This framework is consistent with our findings that students have non-normative ideas about how energy is transformed to heat a solar oven. We find that students are inconsistent in the word choices they make when writing about energy transformation, which signifies that they have not yet engaged in distinguishing among their ideas about attraction, reflection, absorption, and how those processes take place.

Methods

Materials and data source

One 6th grade teacher and their students (*N*=150) participated in this study. The school is in the suburbs of a large U.S. city serving a mainly middle SES community (32% free and reduced lunch eligible, 5% English language learners). We analyze data at the group level; there were 56 groups. Students used the curriculum as described earlier in this paper.

Assessment

To assess student ideas about how energy transforms, we use a series of four questions embedded within the curriculum that refer back to a fictional student and his description of how his solar oven heats up. The fictional student, Andrew, says that his oven heats up because the tin foil flap attracts the sunlight. Students then answer these questions:

1. Do you agree with Andrew? (Multiple-choice: Yes/No)
2. Why or why not? (Short response, refers to question 1)
3. Two students made pictures of what Andrew said. What's the difference between these two pictures? Which one do you think is more correct? (Short response, two pictures are shown in Figure 2)
4. How would you rewrite Andrew's explanation to be more specific?: Andrew's explanation was: "My oven works because the tin foil flap attracts the sunlight" (Short response)

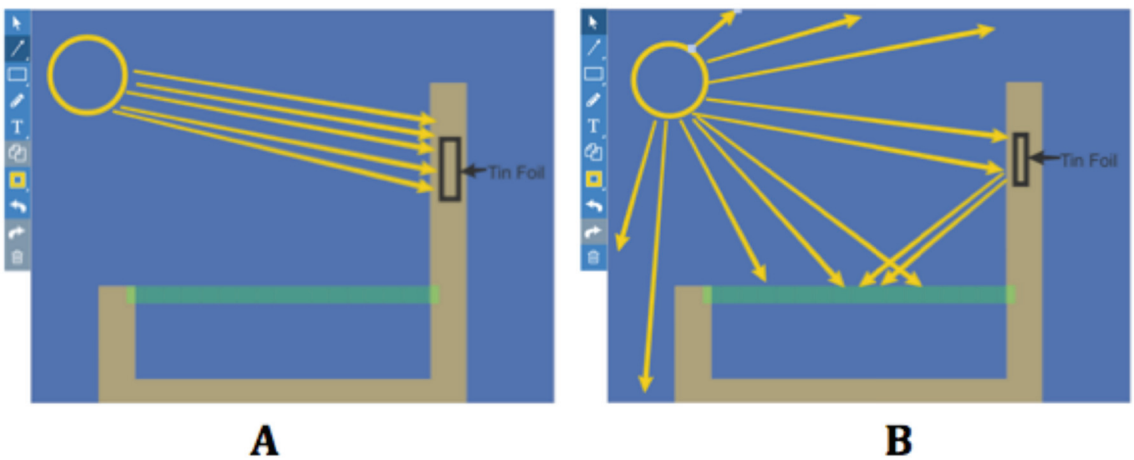


Figure 5.4: Two images, A & B, given as examples of fictional student work in item 3 of our analysis

Throughout this series of questions, students should correct Andrew's statement to be about *reflection* instead of *attraction*. In question 3, students should also have noticed that light from the sun comes from all directions instead of hitting earth in a single beam.

We categorize the responses as referring to a multi-step process while using the word "attract" or not. For example, "The aluminum with attract the sun, then it will reflect onto the food". In this evaluation, we also examine whether students present inconsistencies in the ways they describe the reflection process throughout the short response questions.

Results

We analyze data from groups completing at least two out of three of the short response questions (55/56 groups). Of these 55 groups, 35 (64%) have inconsistencies in their descriptions of reflection across the three short response questions. We define inconsistencies to be differences in words used or the process described across the three short response questions. For example, if a student wrote about tin foil "attracting" in one short response question, but in another question wrote about tin foil "reflecting", this would be considered an inconsistency. We only consider inconsistencies as they pertain to the attract/reflect process.

These inconsistencies provide evidence that students are in the process of developing their ideas about how reflection, absorption, and attraction work. When groups write different ideas or use different words when asked to describe the same process in slightly different ways, as in our three short response questions, they likely have not completely distinguished their ideas or developed a completely coherent explanation. Our curriculum aims to help students distinguish among their ideas about energy transformation and light propagation. We study inconsistent responses as students progress through the unit, rather that only looking at their ideas at the posttest (end of the project). We illustrate the process students describe in a case study.

Overall, 21 (38%) out of the 55 groups describe a two-part process at some point across the three short response questions. This two-part process usually involves attraction followed by something else, for example, "*The tin foil attracts the sunlight, and then bounces off into the box making it heat up.*" Responses from these groups also give insight into student difficulties with the concepts of reflection. While attracting may be a more concrete process since students may have encountered it in the context of magnets where they can see objects moving, the reflection of light is invisible. Students may then try to reason about how reflecting works, thinking that the light must first be attracted in order for other processes to happen.

Students may write about attraction because they are unsure of how light reaches the tin foil in the first place. In question 3, only 20 of the 55 groups (36%) correctly noted that the sun emits light in all directions (as shown in picture B; picture A shows all the light hitting the tin foil). However, 5 of those 20 groups had also used attraction to explain a two-part process. This indicates that students may hold contradictory ideas about how light gets to the tin foil: 1) that the tin foil attracts light, and 2) that light is emitted in all directions and some randomly hits the tin foil. That some of these ideas are elicited when considering pictures shows the value of probing these ideas with multiple perspectives.

Case study: inconsistent ideas

This case was chosen because it represents two common non-normative ideas students have. This group has inconsistent ideas about the way light gets to the oven, but also has inconsistent use of the verbs "attract" and "reflect". The group's responses are as follows:

1. *Is Andrew's explanation correct:* No
2. *Why or why not?:* The aluminum does not attract light. The aluminum only reflects the light that hits it.
3. *Two students made pictures of what Andrew said. What's the difference between these two pictures? Which one do you think is more correct?:* Letter B is correct because the light doesn't target the tin foil. It does not target anything it is just scattered by the sun and goes through the atmosphere and then some of it hits the tin foil and it bounces off into the box.
4. *How would you rewrite Andrew's explanation to be more specific?: Andrew's explanation was: "My oven works because the tin foil flap attracts the sunlight":* My

oven may attract the sun light and then it will bounce off of the tin foil and into the box.

First, while this group recognizes in questions 1 and 2 that Andrew should have written "reflect" instead of "attract", they themselves use "attract" in question four when asked to rewrite Andrew's explanation. This is an inconsistency in how the group recognizes and uses these two verbs. Second, the group writes in question 3 that the light is going to be scattered by the sun, implying that they may understand the random nature of how light reaches objects on earth. However, then in question 4 they write about attraction, and describe a two-part process of attracting and then bouncing off. This indicates that the group may not fully understand what the word "attract" means, since it is inconsistent with their ideas about how light works from question 3.

Significance

This study documents a specific non-normative student idea about energy – that matter must attract energy before it is reflected or absorbed. Understanding the non-normative ideas students have about energy facilitates development of improved curricular materials. While many of the groups in our dataset seem to understand the mechanism at a basic level, they do not fully understand what each of the words "attract", "reflect", and "absorb" mean. In addition, they may also not understand how light gets to objects before it can be reflected or absorbed.

Students' non-normative ideas about attraction are not falsified in the curriculum – the curricular materials do not explicitly show why attraction is incorrect. Therefore, while students gained normative ideas from the curriculum, they also likely held on to their non-normative ideas. This study suggests that middle school curricula on light and heat should specifically address non-normative ideas of attraction. One approach may be to modify simulations that show reflection and absorption of light to also depict the non-normative idea of attraction, so students can see how this non-normative mechanism is inconsistent with our physical world. However, since students bring up this idea in written responses during many different contexts, developing automated guidance to detect this idea, and others like it, is also important.

Model Development for Automated Guidance

The interactive computer model (Figure 5.2) used in this curriculum allows students to investigate how solar radiation transforms into heat to warm a solar oven. While this interactive model has proven useful for student learning of the scientific mechanism of energy transformation, it does not attend to all instructional needs for students. The idea investigated in the first half of this chapter, of light attracting to shiny or dark objects, is one such idea that students cannot investigate properly using the computer model. Thus, it is important to detect this idea in student writing and provide students with resources to consider whether "attracting" is actually how light reaches any object. The next section of this chapter investigates how we have developed automated scoring for student writing to detect the idea of "attracting" as well as other ideas students write about. With this method, we are able to give all students some guidance on making improvements to their ideas, instead of providing guidance to only those students with the non-normative idea of light being “attracted”.

Materials and data source

The dataset used to develop this automated scoring comes from over 1000 students who used a project from the Web-based Inquiry Science Environment (WISE) on Solar Ovens and Solar Radiation (Solar Ovens). These students come from 4 different schools and 7 different teachers over the course of two school years. In each of these classrooms, students interacted with the curriculum as described in the curriculum section of this paper. We develop classifiers to score short response items embedded within the curriculum. The items embedded within the curriculum are most important because they are places where we can provide guidance to students based on what they write. This guidance can direct students to reexamine a model, add new ideas, or consider whether their ideas fit together.

One of the embedded short response questions that is useful for examining student thinking asked students to describe how their solar oven works. Students answered this question two times during the project, once after each design iteration. In the activity, students were also asked to talk to another group in the class to get ideas about how to improve their oven in a collaboration activity. Students wrote about a new addition or change they wanted make to their oven and how the oven would work with the change. Since these questions all deal with the oven design and interpretation, we combined them together into one corpus of student responses. We refer to this data as “embedded” within the curriculum in future sections of this paper.

Evaluation of student responses

When developing scoring rubrics for the embedded questions, we found that a typical knowledge integration rubric, which scores for the number of connections between ideas, was too complex for the types of responses present. We instead developed a coding scheme based on a simplified knowledge integration rubric. This coding scheme is based on the quality of the explanation and how specific the response is about where a material is located on the oven. For example, the response shown in figure 3 says “The construction paper and will absorb the solar radiation and the tin foil will both absorb and reflect the solar energy". This response does not specify where the black paper is used. Also present in each of these coding schemes is the presence of the non-normative idea of “attraction”.

Version 1 (V1)	Version 2 (V2)	Version 3 (V3)	Version 4 (V4)
1. No Ideas	1. No Ideas	1. No Ideas	1. Needs revision
2. No exp. & specific	2. No exp.	2. No/poor exp.	2. Good exp.
3. Poor exp. & specific	3. Poor exp.	3. Good exp.	3. “attract”
4. Good exp. & specific	4. Good exp.	4. “attract”	
5. No exp. & not specific	5. “attract”		
6. Poor exp. & not specific			
7. Good exp. & not specific			
8. Non-normative: “attract”			

Figure 5.5: Four coding schemes for embedded data, based on quality of explanation and specificity of material placement on the oven

Not many students wrote responses that specify where the materials are used, so we do not have a large corpus of data with specific responses. This makes it difficult for us to build a model that can identify a variety of specific responses using an algorithm based on the data we already have. We chose to code at the most complex level first, since we would then be able to collapse categories from the complex coding scheme to create simpler coding schemes without completely recoding the data. Our four coding schemes are shown in figure 5.5. Version 1 (V1) has 8 groups, with one of the groups for students who did not write any relevant ideas (1) and one for students who exhibit a particular non-normative idea (8: that materials can attract light instead of absorbing it). In version 2 (V2) we collapsed responses so that groups are based only on the quality of the explanation, without using specificity of the material location. In version 3 (V3), we further collapsed the groups with no explanation and those with a poor explanation (e.g., “I will use black paper because it will help my oven heat up"). A breakdown of the number of responses in each group for each coding scheme is shown in table 5.1. There were 1089 responses in total, with 759 of those coming from question 1 and 330 coming from question 2. Of these embedded responses, there were 79 responses that included the non-normative idea of “attract”.

Table 5.1: Breakdown of embedded data across codes and coding schemes

Code	V1 Count	V2 Count	V3 Count	V4 Count
1	111	111	111	700
2	196	327	589	310
3	165	262	310	79
4	212	310	79	N/A
5	131	79	N/A	N/A
6	97	N/A	N/A	N/A
7	98	N/A	N/A	N/A
8	79	N/A	N/A	N/A

Development of classifiers

We developed a custom algorithm to automatically score each question in the pre/posttest. The classifier takes the text response as input and classifies the response into a score category (1-8, based on the categories presented in Table 2). We first experimented with various classifier options to determine how each worked with our data. We then created and tuned an ensemble classifier using the strongest classifiers.

Spell checkers are commonly used in the pre-processing of text data. However, we found that using a spell checker did not help our classification accuracy, and in many cases decreased classification accuracy. This was not surprising, since many relevant words (e.g., absorb, plastic wrap) are spelled incorrectly in the responses ("obsorb", "plastic rap"). However, students sometimes spell these words so incorrectly that a spell checker mis-corrects the words to other (irrelevant) words. Sometimes students also mistakenly use homophones for the word they actually wish to write, as in the case of "wrap" versus "rap". A spell-checker would not correct these misused words. Since sixth grade students are still developing their vocabulary, spelling skills, and typing skills, it can be difficult to deal with these errors in spelling.

In pre-processing our data, we used lemmatization, or removing in endings and returning the lemma, or set of morphologically related forms. We filtered out punctuation and common stopwords from the text, and normalized the text to be completely lower case.

After creating a classification pipeline using feature unions, we selected specific features that were relevant for each question. Feature unions make it possible to extract several features at once. For example, in the "attract" category, we used word features like "attract", "attrakt", "attraction", and "atract" to identify the non-normative idea students have about energy being attracted to shiny or dark-colored objects. We then used stemming for each feature to find the word in any form in which a student might use it. We also used the length of the answer as a feature in each algorithm. This improved our classifiers, but simply using length may not be the best feature, since some students demonstrate good understanding of concepts by writing very concise and correct responses while others employ a "kitchen sink" method of writing everything that might be relevant.

Classification results

We used 70% of the data for each question to develop our classifiers, and the remaining 30% of the data to test how accurately each classifier could score responses into knowledge integration categories. This split (as opposed to using 80% for development and 20% for testing) increased the probability that all categories would occur in both the development and test sets. Results of the classifier are shown in table 5.2.

With the embedded question classifier, we looked at using each coding scheme. Then we used the most successful coding scheme in terms of the % accuracy and looked at separating the data by question. Perhaps not surprisingly, the coding schemes with the highest accuracy were the least complex (V3 and V4). Accuracy increased after separating the two questions, but only for group with a larger amount of data (question 1 data only). Often with this classifier, the last code (code 3 in V4) that looks for the non-normative idea of "attract" is quite accurate. This can be seen in the confusion matrix in table 5.3, which shows the performance of the classification algorithm. Each column shows the number of responses predicted to be in each group, while the rows show the number of responses that were actually in each group (based on scored data). The diagonal shows the number of responses that were accurately classified, while any off-diagonal values show where and how the algorithm mistakenly classified responses. Responses scored as a 3 in the actual data were often correctly classified by the model.

Table 5.2. Results (% accuracy, Cohen's kappa, and 10-fold cross validation) for embedded response classifier using multiple versions of coding schemes and question types

Coding Scheme	**Questions**	**% Accuracy**	**Cohen's Kappa**	**10 fold CV Average**
V1 (8 groups)	1 & 2 (N = 1089)	43.4	0.339	0.449
V2 (5 groups)	1 & 2	61.2	0.481	0.633
V3 (4 groups)	1 & 2	77.1	0.589	0.723
V4 (3 groups)	1 & 2	80.4	0.568	0.813
V4 (3 groups)	1 only (N = 759)	86.6	0.674	0.829
V4 (3 groups)	2 only (N = 330)	63.6	0.386	0.712

Table 5.3. Cross-tabulation using coding scheme v4 using questions 1 & 2

Predicted	1	2	3	All
True				
1	199	13	0	212
2	49	43	0	92
3	1	1	21	23
All	249	57	21	327

Conclusions

Use of automated scoring has the potential to improve online curriculum for various types of platforms, such as MOOCs or platforms like WISE that are run in K-12 environments. Using automated scoring to provide guidance for students would be beneficial for ensuring that students have multiple chances to re-consider and distinguish among any disparate ideas they may have.

Using feature selection, we can also expand this work to other types of questions and curricula in which students often have specific non-normative ideas that may be problematic for their future learning. For example, in curricula on climate change, students often think that the ozone hole contributes significantly to global warming (Andersson & Wallin, 2000; Koulaidis & Christidou, 1999). We can target the mention of “ozone" using automated scoring algorithms and direct students back to sections of the project to explore that idea further. In the case of ozone, there are already interactive computer models in existence that allow students to investigate the phenomenon, making the task of guiding students to falsify their idea less challenging. The task of guiding to students to falsify their own ideas about light anthropomorphically being attracted to shiny or dark objects needs careful consideration and design of guidance within an online environment.

The design of automated guidance is another important area for consideration. In the future we will develop guidance for students by building on prior research around providing useful and effective guidance for students (e.g., Tansomboon, Gerard, Vitale, & Linn, 2017; Gerard, Ryoo, McElhaney, Liu, Rafferty, & Linn, 2016). These papers use knowledge integration to guide students in considering making revisions to their writing or in revising their ideas about scientific mechanisms.

The assessment items used in the first half of this chapter to understand what students mean when they describe light as being "attracted" to shiny or dark objects may also be useful in developing guidance. Since these items asked students to critique another student's response, students were generally quite engaged in the task. We may have success by using a similar critique activity targeted at helping students reflect on how light could or could not be anthropomorphically attracted to an object, especially when paired with an activity to add ideas about reflection to students' repositories. This could involve the development of new features within the existing model to illustrate the impact of the angle of reflection or of a completely new visualization to illustrate how reflection takes place.

In the first half of the chapter we discovered that students were quite inconsistent with their use of the verb "attract" to describe what was happening to light in their solar ovens. This may be due, not to an engrained conceptual issue, but to students not having enough information about the process to develop the correct understanding of how it functions within a system. The idea of reflection as opposed to "attraction" may seem simple, but providing students with the resources to so they can add new ideas about how reflection works may be key to helping them later distinguish between their ideas about "attracting" and reflecting. By providing these resources, we will also have a resource to direct students back to if automated guidance on student writing finds the "attract" idea.

In developing guidance for this particular idea, we will likely need to use a combination of instructional tools and design-based research methods. In the knowledge integration framework it is important for students to grapple with all their ideas in order to develop a coherent understanding of a process. By providing students with a variety of tools to add new ideas about the process of light propagation, as well as tools and spaces to allow students to distinguish among their prior and newly added ideas, students may be able to develop a coherent understanding of how light works in a solar oven.

Chapter 6: Learning Design Through Science vs. Science Through Design

This chapter investigates two ways of framing design projects and their impacts on learning. The study explores the benefits of learning science concepts before or during a design project. Based on the NGSS science and engineering practices, in an *engineering* condition, students learn the necessary science concepts during a design project. In a *science* condition, students learn the science concepts first, then apply them during a design project. The study explores the benefits of each approach to inform instructional design. We use the knowledge integration framework to develop curriculum and assessment items, including an interactive computer model of a solar oven. Using three types of pre/posttest assessment items, we found students in both conditions gained insights on science and engineering design items; students in the *engineering* condition outperformed the *science* condition on a science-design integration item and conducted more trials during the design process while using an interactive computer model.

Introduction

Engineering projects are becoming more common in K-12 schools, but while it is often claimed that engineering projects improve student achievement in mathematics and science, research on this topic has shown that many projects do not live up to the claim (Teacher Advisory Council, 2009). While engineering projects may generate more student interest and engagement (Hmelo et al., 2000; Cantrell et al., 2006) than typical science curricula, they often fall short on developing science concepts. Ideally, undertaking a science project should be motivating, while also helping students to understand the interplay between science concepts (like energy transformation) and engineering design decisions. However, the framing of goals can impact what aspects of the project are emphasized. In projects aligned with science goals, students learn the science concepts and then do a design project to apply those concepts (*science* condition). In projects with goals more aligned with engineering, students learn the science concepts during the process of completing a design project (*engineering* condition).

Often, in science the goal is to develop knowledge, while in engineering the goal is to develop a solution (Lewis, 2006; Purzer, et al., 2015). We use this distinction in designing the two conditions in this study. In addition, we draw on the Next Generation Science Standards (NGSS) science and engineering practices, specifically the practice of "constructing explanations (for science) and designing solutions (for engineering)" (NGSS Lead States, 2013) to inform our conditions. This study compares versions of a solar ovens unit that loosely use one or the other goal frames and present a focus on either constructing explanations or designing solutions, while keeping the overall content of the curriculum the same.

We use the knowledge integration framework (Linn & Eylon, 2011) to guide the development of the curriculum and this study. This framework focuses on building a coherent understanding of concepts, and has proven useful for design of instruction featuring dynamic visualizations (Ryoo & Linn, 2012) and engineering design (Chiu et al., 2013; McElhaney & Linn, 2011). The framework emphasizes linking of ideas by eliciting all the ideas students think are important and engaging them in testing and refining their ideas. When students build a physical artifact, as in many engineering projects, they can only test a few of their ideas due to time and material constraints. Features in this curriculum, like using interactive computer models, allow students to explore many more ideas, thereby facilitating knowledge integration.

Though engineering projects are potentially motivating, when students build a physical model they often neglect the scientific basis for their decisions (Crismond, 2001), instead focusing on aesthetic and otherwise superficial details of construction. Tools like interactive computer models can help students connect science principles and design decisions by making mechanisms such as energy transformation visible (Snir, Smith, & Grosslight, 1993; Wilensky & Reisman, 2006). The combination of computer models and hands-on activities in design activities allows students to test many designs while also visualizing how energy transformation takes place in their designs.

In addition to providing science content knowledge, design projects utilizing computer models provide students with an opportunity to explore authentic practices of scientists and engineers. The NGSS envision that instruction would combine practices including modeling, data, analysis, computational thinking, and design to enable students to integrate their scientific and engineering ideas (NGSS Lead States, 2013). The solar ovens curriculum used in this research familiarizes students with the way energy transforms from solar radiation to heat (MS-PS3-3) by using a hands-on project and interactive models, emphasizing the modeling aspect of the science and engineering practices of the NGSS as well as the standards associated with energy (NGSS Lead States, 2013). This curriculum draws on all eight of the science and engineering practices in the NGSS, focusing on using models, developing solutions, and engaging in argument from evidence.

A project framed as an engineering design project from the beginning may offer students meaningful opportunities for science learning, especially when they must consider trade-offs in their designs (Purzer et al., 2015). This type of consideration of design trade-offs may be especially useful in helping students to integrate their science ideas with their design decisions. Design projects have been found, in some cases, to positively impact students' scientific reasoning (Silk et al., 2009). However, these students may not learn complex science concepts if their focus is on incidental aspects of design. Hands-on projects that directly follow a related science unit may allow students more time to focus on understanding the complex scientific phenomena they are being asked to apply, while still motivating them to learn the concepts in order to apply them to their design. However, the separation of the science content from the design project may seem disjointed to students and lead to lower motivation in learning the concepts.

We use knowledge integration assessment items (Linn & Eylon, 2011; Liu et al., 2008) at pretest and posttest targeted at three specific areas to better understand how each of our conditions impacts learning. These items measure science concept integration, engineering design practices, and the integration of science and engineering design practices. While there has been much work done to advance engineering education at the K-12 level (e.g., National Research Council, 2009; Bybee, 2011), there has not been as much work done to develop valid items for assessing engineering practices.

The two conditions in this research are meant to understand two common ways hands-on activities are framed in the classroom. By understanding the benefits of each method of framing, we hope to develop a curriculum that helps students to integrate their ideas about science concepts and engineering design better. While teachers may have their own preferred way to conduct hands-on projects in their classrooms, this work is meant to help strengthen student learning in both the science and engineering domains no matter the framing of the classroom project.

Methods

Participants and procedures

One teacher and her 153 students participated in this study. Out of these students, 139 students completed a pretest, (some part of) the curriculum, and a posttest. The pretest was conducted one day before beginning the unit, and the posttest was conducted one day after finishing the unit. Both the pretest and posttest were administered to students individually. Pairs, or in some cases triads, of students were assigned to collaborative workgroups by their teacher to work on curriculum. Workgroups were randomly assigned to a condition (*science* or *engineering*) by the software. All students received the same curricular content, but activity focus and order varied by condition.

Curricular materials

This study was implemented in a curriculum module entitled *Solar Ovens* in the Web-based Inquiry Science Environment (WISE), which utilizes a variety of instructional and assessment tools (Linn & Eylon, 2011). The goal of the unit was to familiarize students with the way energy transforms from solar radiation to heat through a hands-on project and interactive models, covering the modeling aspect of the Science and Engineering Practices of the NGSS, as well as the standards associated with energy, specifically standards related to the transfer of thermal energy (NGSS Lead States, 2013).

The solar ovens curriculum within WISE has been designed and refined with the collaboration of multiple expert teachers and researchers to help students test and refine their ideas about energy transformation. The curriculum seeks to help students utilize their ideas about how radiation works in various contexts, like in the atmosphere and inside solar ovens.

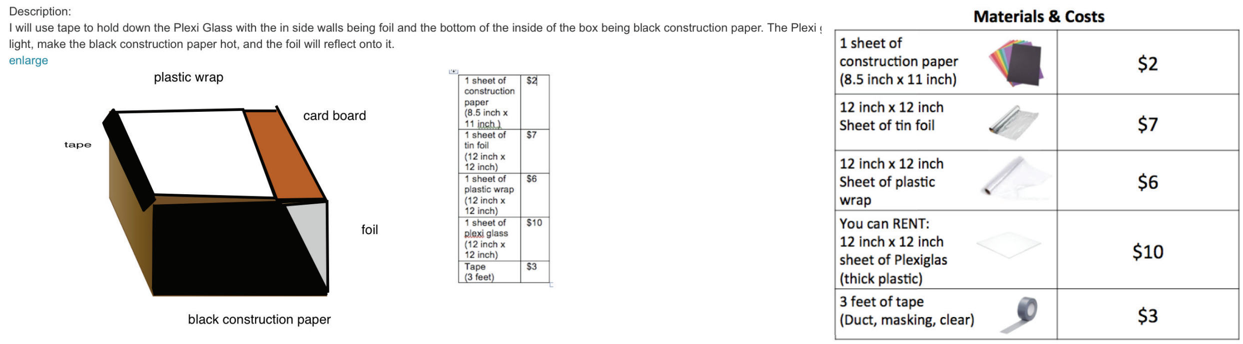
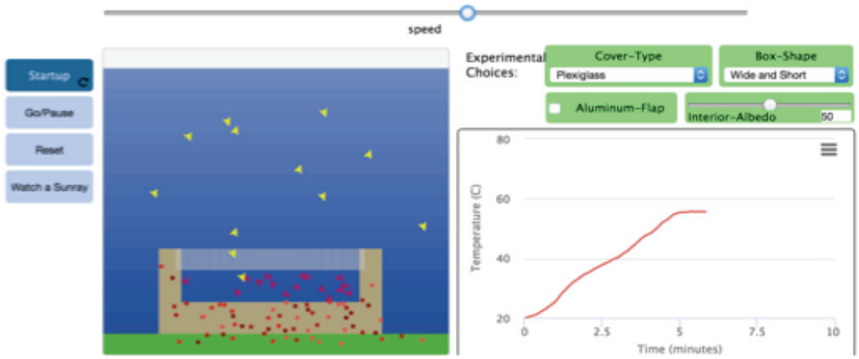


Figure 6.1: Student budget (left) and example of student design (right); students were given $20 for their first design iteration and $13 to add to their oven for the second iteration

Students in both conditions followed a modified "design, build, test" approach. An important feature of this unit is a budget activity in which students make decisions about and justify the materials they choose to use for building (Figure 6.1). During the design phases, students also draw pictures of their ovens and explain how energy transfer will occur. Students also use an interactive model of a solar oven, designed using NetLogo (Wilensky, 1999), to test features in the solar oven and understand how solar radiation transforms into infrared energy (Figure 6.2). Students generate trials using the model by allowing the model to run for 5 simulated minutes without changing the input variables. When students test their physical prototypes they also test them under a lamp for 5 minutes. After each trial is generated, it is automatically added to a table, allowing students to track the trials they ran and the results of those trials. The computer model has been previously tested to understand how students use it at different points during the curriculum and how it impacts learning. Our earlier findings indicate that the computer model aids students in integrating their science and design ideas, and that students interacting with the model earlier during the curriculum (during the planning phase) benefit more than students who interact with the model later in the curriculum (the reflecting phase) (McBride, et al., 2016). After designing, students build physical solar ovens, which are tested under lamps with a common set of requirements, so results are comparable between trials and groups.



#	Box Shape	Cover Type	Flap?	Int Albedo	Temp at 0m	Temp at 1m	Temp at 5m	Temp at 10m
1	Wide and S	None	No	50	20	24	30.2	31.5
2	Skinny and	None	No	50	20	26.2	29.7	31.1
3	Wide and S	Plexiglass	No	50	20	25.6	55.4	53.4

Figure 6.2: The interactive model used by students for design solar ovens and understanding energy transformation, with an automatically generated table below

Condition differences

Conditions did not differ in content, only in the order the content was presented and in the framing of questions or activities. In the *engineering* condition students were introduced to the design project in the first step, then were prompted to learn or consider science concepts during the design process. In the *science* condition, students learned all the science concepts at the beginning of the project in a module about the atmosphere, and were then introduced to the design project as a way to apply what they had just learned. Students in each condition used a concept-mapping tool to map energy flow. In the *engineering* condition, students mapped energy flow in their solar ovens, while in the *science* condition students mapped energy flow in the atmosphere. These differences are outlined in Tables 6.1 and 6.2 and depicted in Figures 6.3 and 6.4.

Table 6.1: (left) shows the main steps in the curriculum for the *engineering* condition, including the number of steps.
Table 6.2: (right) shows main steps in the curriculum for the *science* condition, including number of steps.

Engineering Condition (26 steps)	
Activity (# Steps)	**Details**
Design & Science Concepts (18)	- Introduction to project - Solar radiation - Solar oven model - Concept map of energy transformation (in solar oven) - Reflectivity - Preliminary design - Why do you need a cover? - Greenhouse gas model - Budget and final design
Build (1)	- Build physical solar oven
Test (3)	- Test physical solar oven
Redesign (2)	- Collaborative critique activity - Use solar ovens model to redesign oven, write updated description
Connect (2)	- Make connections between solar oven and greenhouse gases

Science Condition (25 steps)	
Activity (# Steps)	**Details**
Solar Radiation and the Atmosphere (7)	- Solar radiation - Reflectivity (of earth) - Concept map of energy transformation
Solar Radiation and Greenhouse Gases (3)	- Greenhouse gas model - Update concept map of energy transformation to include greenhouse gases
Design & Build & Connect (10)	- Introduction to project - Model a solar oven - Make connections between solar oven and greenhouse gases - Budget and design - Build physical solar oven
Test (3)	- Test physical solar oven
Redesign (2)	- Collaborative critique activity - Use solar ovens model to redesign oven, write updated description

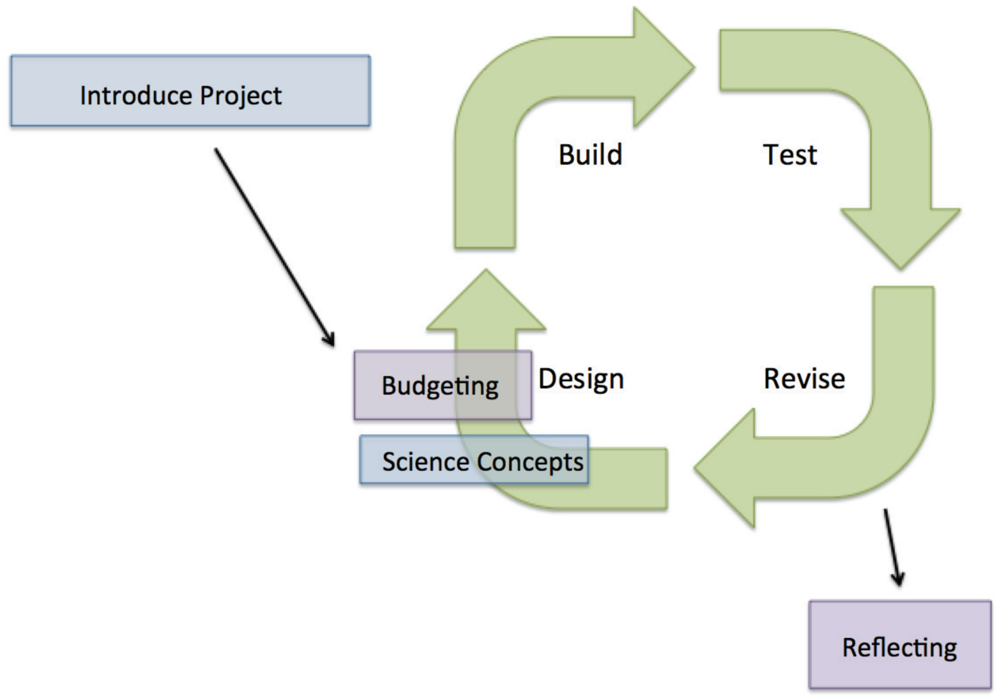


Figure 6.3: Pictorial description of *engineering* condition curriculum

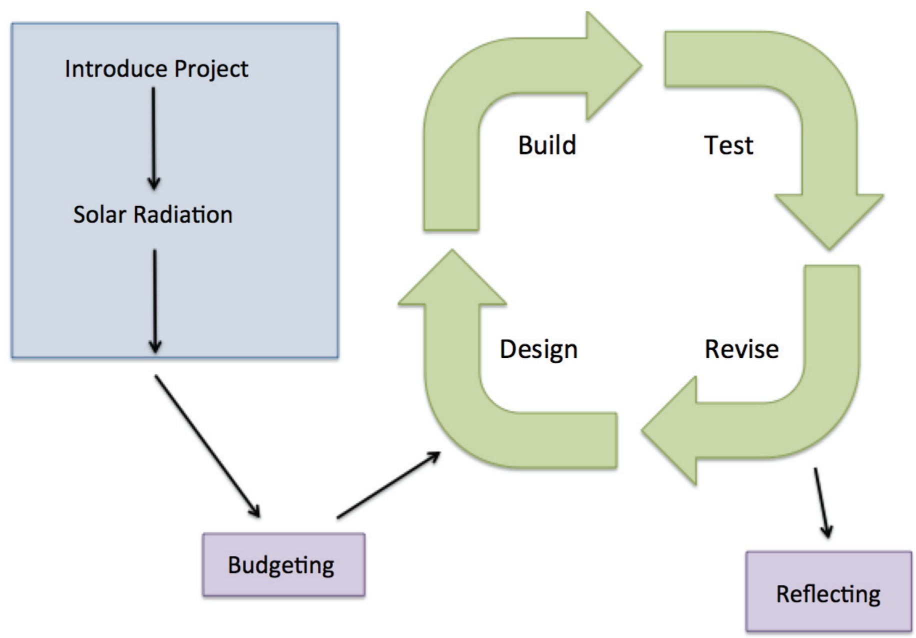


Figure 6.4: Pictorial description of *science* condition curriculum

Test materials

The pre- and posttest assessments we used consisted of 9 assessment items. These items fell into three areas: science concepts, engineering practices, and the integration of science and engineering. All items use short response format, and are scored using knowledge integration rubrics. Of these 9 items, 5 items measure integration of science concepts, 3 items measure integration of engineering design ideas and practices, and 1 item measures the integration of design practices with science concepts.

One of the science concept items, *Car* prompted students to explain what would happen to a car left in the sun during a cold day. In an engineering item, *Budget,* students were asked to describe how two fictional students would build solar ovens using two different lists of materials and then to describe the tradeoffs made in each design. In the science-engineering integration item, *Model,* students were asked to use a basic solar oven model (like that shown in Figure 6.2, but with only a box shape drop-down option) to help a fictional student determine whether a tall, skinny box or a short, wide box would heat up faster. This item is shown in Figure 6.5. The pretest and posttest were composed of the same items.

While the science and engineering integration items measure how well students link their ideas about design or about science concepts, we were particularly interested in the performance of students in each condition on the integration item, since a goal of this curriculum is to help students use their science ideas to justify their design decisions. This integration item has been tested with over 1000 students in prior work.

This integration item makes use of both science content and engineering practices relating to modeling. The rubric, in Table 6.3, shows this. The science content students should be writing about is the mechanism of energy transformation. There are four specific performance expectations in the NGSS (NGSS Lead States, 2013) on engineering design. Two of these standards are related to modeling, which is also listed as one of the science and engineering practices in the NGSS. These performance expectations ask that students can "Analyze data from tests to determine similarities and differences among several design solutions to identify the best characteristics of each that can be combined into a new solution…" and can "develop a model to generate data for iterative testing and modification of a proposed object…such that an optimal design can be achieved".

These performance expectations indicate that it is important for students to use data from a model in order to make decisions about the design of a product or object. We also developed our rubric to analyze how students are using data from the model to respond to this item, brining in the engineering practices. Since students must use both the engineering practices of using data from models and their understanding of how energy transforms within a solar oven to score highly in this item, we consider this item an integration item.

The science items used in this analysis prompt students to write about the mechanism of energy transformation in various situations, for example inside a car. Students are also asked to explain how reflection works to either aid or hinder heating, how the color of objects impacts their ability to heat up, and how parts of the solar oven relate to the atmosphere or to greenhouses in the science items.

The engineering items ask students to describe how they would build a solar oven using a different set of (commonly found) materials than they are given in the unit. They are then prompted to write about the tradeoffs between two ovens and what each oven would be good at. In another engineering item, students are prompted to write about what they would need to know in order to develop a structure in which their community could grow tomatoes in the winter. These items are each scored according to students' ideas about the engineering design, and taking into account the engineering performance expectations. The remaining two engineering performance expectations refer to analyzing tradeoffs ("Evaluate the competing design solutions using a systematic process…"; NGSS lead states, 2013) and defining "criteria and constraints of a design problem". The engineering items used in this study were developed with these engineering performance expectations in mind.

We also use the automatically generated table from students' interactions with the interactive computer model (Figure 6.2) to analyze how many trials students ran during the design process. In addition, we use three other measures of students' interactions with the interactive computer model. We use the amount of time students spent on the project step that included the computer model, the number of clicks students made in the computer model, and the average number of clicks made per hour (time spent divided by number of clicks). All of these measures come from analysis of student log files. A click refers to any action made by a student while using the mouse, for example, clicking on the graph to resize it or making a choice for the virtual solar oven (e.g., cover type).

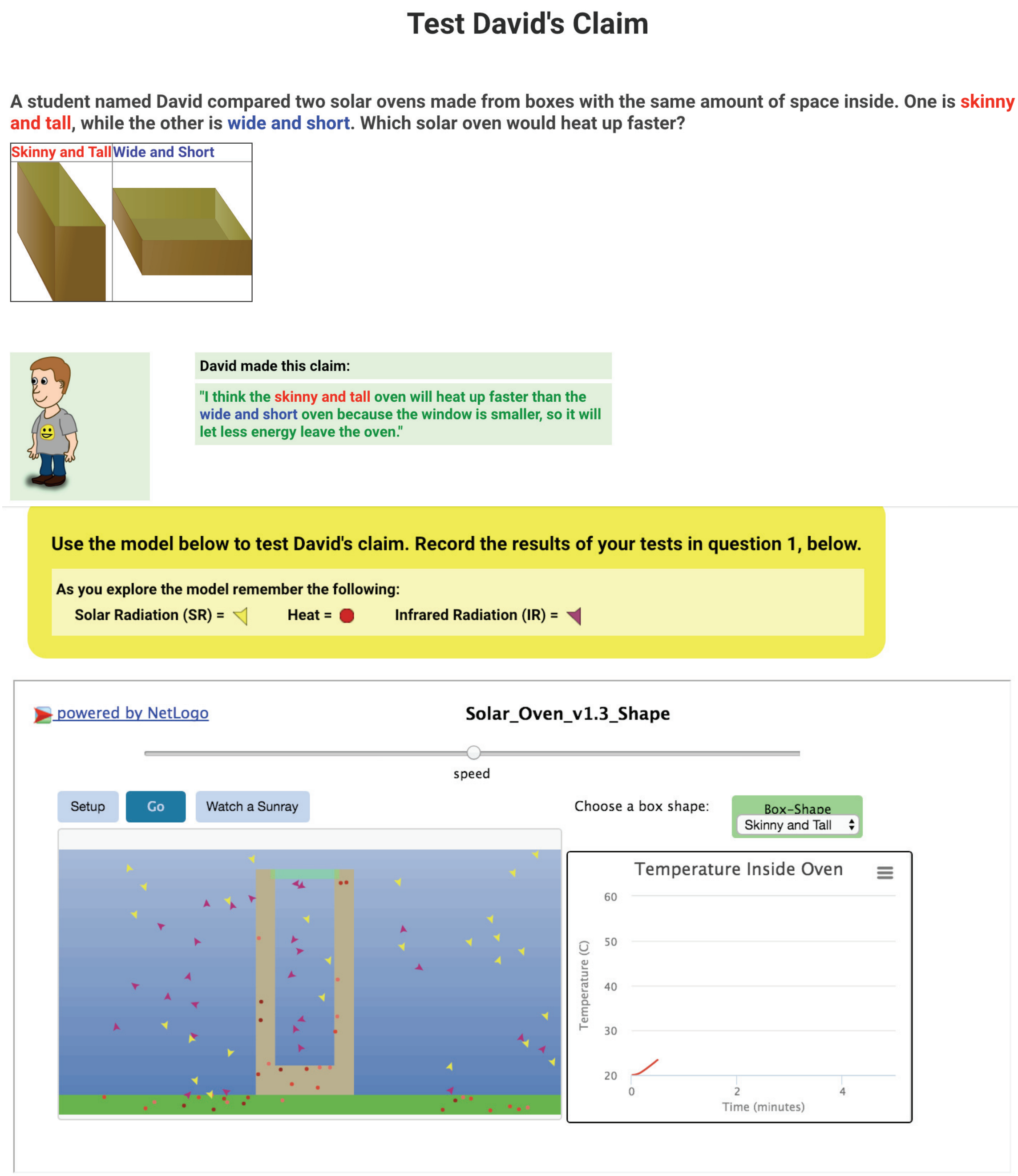


Figure 6.5: *Model* item used on the pretest and posttest (integration item)

Analysis

To measure knowledge integration, the items were scored using knowledge integration rubrics to assess links between multiple normative science ideas (Linn & Eylon, 2011; Liu et al, 2008). The knowledge integration rubric for *Model*, the science/design integration item, shows how links are scored (Table 6.3). Multiple researchers develop the rubrics for each item; initial scoring of data is also done by at least two researchers, with high inter-rater reliability ($\kappa > 0.8$).

Since this research investigates the differences between framing as a whole project (more similar to engineering) or as an application of concepts (more similar to science), our analysis looks at whether there are differences between conditions on the science, engineering, or integration items. However, unless otherwise specified, we examine the corpus of all 9 items.

To analyze the differences between conditions based on how students used the interactive computer model during the design phase of the project, we used a count of the number of trials run by each group. Each trial is added to an automatically generated table after students allow the model to run for 5 simulated minutes (takes about 30 seconds to 5 minutes in real time). We do not count trials that were not allowed to run for shorter than this time period because, since they were not added to the automatically generated table, students did not have a record of them and were therefore not able to look back at these trials while making their decisions. This analysis is done at the workgroup level.

Table 6.3: Sample knowledge integration scoring rubric for the *Model* pre/post open response item

Score	Level	Examples
1	**Off Task**	*I don't know.*
2	**Irrelevant/Incorrect**	*David is correct because I chose the skinny and tall one and the heat went up really fast.*
3	**Partial** Normative isolated ideas without a valid link	*David's claim is not correct because in the model it show solar radiation stayed trapped inside the wide and short one making heat easily trapped inside.*
4	**Basic** Elaborate a scientifically valid link	*David's claim is incorrect because the skinny box got to 33.8 in 2 minutes and the wide box got to 44.7 in 2 minutes. The wider box could keep a lot of energy because of the space and the skinny box doesn't have a lot of space. So, this means David was wrong.*
5	**Complex** Elaborate two or more scientifically valid links	*David's claim is incorrect because the more area for radiation to come the more radiation can get trapped and turn into heat.there is less of the when you have a skiny box.*

Results & Discussion

A t-test of pooled pre- and posttest data across conditions revealed a significant effect of testing session [$t(304) = -6.44$, $p < 0.0001$], demonstrating that across both conditions students made gains from pre- to posttest (Figure 6.6). Most students also completed the unit within the time allocated by their teacher.

When pooling all pre/post items together, there were no overall differences between the *science* and *engineering* conditions. When considering the groups of science and engineering assessment items, there were also non-significant differences between conditions. Students in the *science* condition made slightly greater gains on the science assessment items between pretest and posttest, and likewise, students in the *engineering* condition made slightly greater gains on the engineering assessment items between pretest and posttest; neither of these differences were significant.

When considering the integration assessment item, there was a significant impact of condition. Using a regression model, students in the *engineering* condition scored higher on the posttest integration item, when controlling for pretest score ($\beta = 0.18$, $p < 0.01$). This is shown in Figure 6.7. This suggests that the *engineering* condition curriculum helped students to use the model in a way that led them to integrate engineering practices with science knowledge.

Looking at the number of trials students run in the interactive computer model during the design phase of the project, we also see an advantage for the *engineering* condition. Groups in the *engineering* condition ran significantly more trials than those in the *science* condition ($\beta = 0.33$, $p < 0.02$). Figures 6.8 and 6.9 show data on the variance between the conditions in terms of the number of trials run. In the *engineering* condition, more of the groups used the model to run trials, and a larger proportion of groups ran more than one trial. In the *science* condition, many groups did not even allow the model to run for a full trial, and of groups that did run any trials, a majority of them only ran one trial. From classroom observations, we understand that many of these students who ran zero or one trial, in either condition, already had ideas about how they would build a solar oven that they were quite tied to. Students who ran one trial using the model often spoke about confirming that the oven they planned to build would work by running that single trial. However, all ovens that could be tested within the modeling environment would “work” to a degree, so this may not be the best test of functionality. Students who ran zero trials were sometimes confusing about how the model worked, even though the teacher gave a tutorial and was roaming the classroom talking to students about the trials they were running in the model. Based on previous findings that students do not necessarily need to run controlled trials to learn science concepts, but running multiple trials is important (McBride, et al., 2017), these students who run zero or one trial are important to pay attention to when revising curriculum.

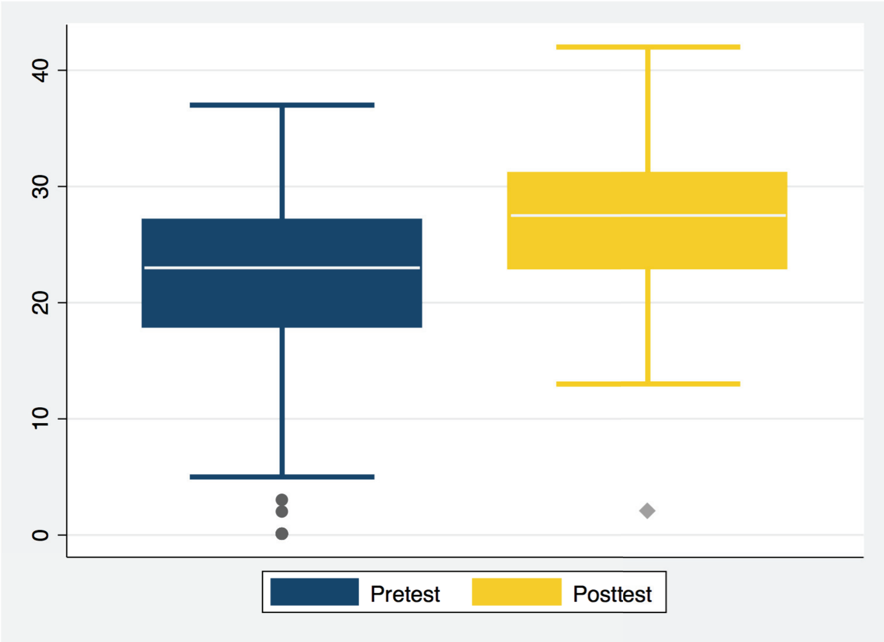


Figure 6.6: Differences between pretest and posttest scores for *Solar Ovens*

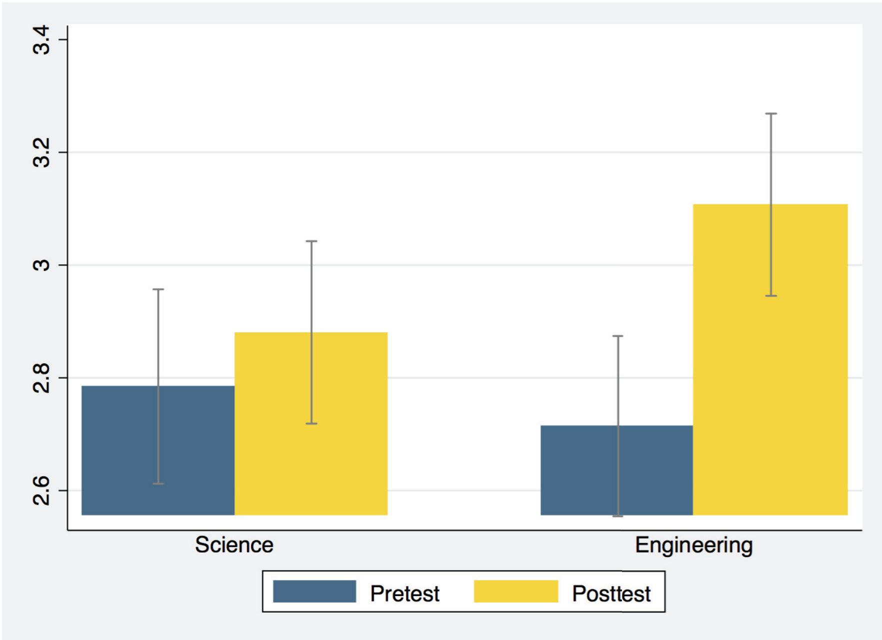


Figure 6.7: Differences between conditions (*Science* and *Engineering*) on the integration item at pretest and posttest

There was not a significant difference between conditions when examining the amount of time students spent using the computer model. On average, students spent about 20 minutes using the computer model, with students in the *engineering* condition spending slightly longer on average than students in the *science* condition.

However, there was a significant difference between conditions when looking at the number of clicks, or actions, students made while using the model ($\beta = 0.34$, $p < 0.01$), with students in the *engineering* condition making 25 more clicks than students in the *science* condition (mean for *engineering* condition: 56, mean for *science* condition: 31). Here, clicks mean any action taken in the modeling environment by the student, for example changing the cover type using the dropdown menu or adjusting the speed of the simulation. Since students in the *engineering* condition ran more trials, we would generally expect them to also have made more clicks. When combining the measures of time and clicks to be the number of clicks per hour (calculated: clicks divided by time), we find no significant difference between conditions. This measure is important to check because in some cases, students may make rapid clicks on an interactive model without allowing the model to run and reveal the results or patterns to students. We found that there is generally a linear relationship between the amount of time spent and number of clicks in the model. We also found few outliers, meaning that most students were using the model appropriately.

Students used the model to run more trials in the *engineering* condition, even though students in both conditions generally spent the same amount of time using the model. This may mean that students in the *engineering* condition used the model more effectively to test their ideas. This is likely because students are introduced to the model very early in the project, so they are using the model to add and test new ideas about their design. Students in the *science* condition may have already been considering their design throughout the project, but before they were able to test their ideas using the model. This may have caused students to become attached to certain choices they made before they had a chance to use the model to test design options.

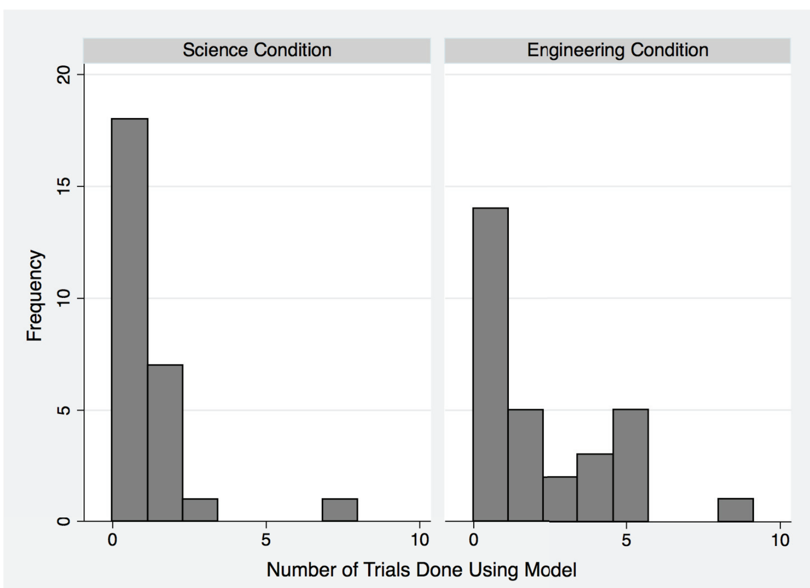


Figure 6.8: Histograms showing the number of trials done using the interactive computer model in each condition;

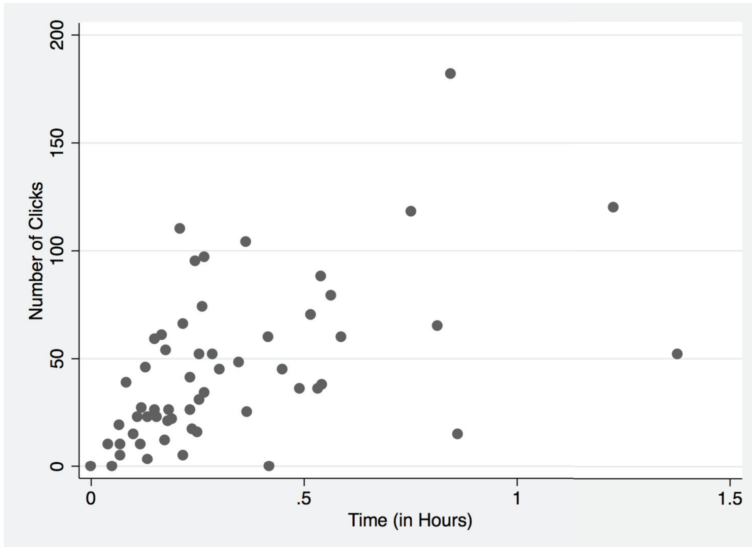


Figure 6.9: scatterplot showing the number of clicks and the time spent using the model for each group of students

Conclusions & Implications

This study compared instruction designed to take advantage of an *engineering* framework with instruction designed to use typical *science* practices. Students in the *engineering* condition were more successful in integrating their science ideas with their oven design than were students in the *science* condition. Students in the *engineering* condition may have used the design of their oven as an artifact for testing their science ideas. The students in the *engineering* condition conducted more trials than students in the science condition giving them more opportunities to test their science ideas. In the *science* condition the students may have seen designing the solar oven as separate from learning the science concepts.

Students in the *science* condition spent more of the curriculum solely focused on learning science concepts, therefore it makes sense that students in this condition would do slightly better at integrating their science ideas on the science integration items. Similarly, students in the *engineering* condition spent a longer time considering the trade-offs of their designs, and also performed better on items that measured engineering practices, like analyzing designs for trade-offs.

In addition, students in the *engineering* condition may have seen the ideas they tested in the interactive computer model as more open to questioning, which may have encouraged students to test more ideas (Sandoval & Morrison, 2003). Students in the *science* condition seemed to be more attached to their original ideas, testing fewer ideas in the computer model. The *engineering* condition seemed to open students up to more possibilities in their design, while the *science* condition in some ways gave students a more limited idea of the possibilities for their designs. This may have been because students were introduced to the model earlier in the curriculum in the *engineering* curriculum. This would serve to make the modeling environment a space where students were both developing new ideas and testing them. This differs from the *science* condition, where students may have developed their ideas earlier in the unit, during the science concepts instruction, then had to wait to test those ideas until they reached the modeling environment.

Since adding and testing new ideas is a proven feature or curricula that improves student learning (Linn & Eylon, 2011), it is important to emphasize this in all design projects, including those following a format more similar to the *science* condition.

These results indicate that there are benefits for each type of framing. This is important to recognize in aligning the design of the curriculum with teachers' learning goals for their students. In this work we recognize that there may be outside factors that impact teachers' choices in how to frame a hands-on project. However, the results show that there are impacts in what students take away from different framings of the same hands-on project. To improve the *science* condition, curriculum designers or teachers may have to work to integrate the addition and testing of ideas earlier during the curriculum to overcome students' fixation on certain ideas during the design process. To improve the *engineering* condition, science concepts must be emphasized.

This work included only minimal differences between conditions; ordering and question framing on only some curricular activities. In spite of this, it still generated a useful and statistically significant finding. It would be helpful to separate the conditions even further in order to understand how to frame design projects. However, separating the conditions further may be very challenging for one teacher to orchestrate (since students are randomly assigned within class periods).

In addition, understanding how the framing of hands-on projects impacts learning outcomes also relies on valid and reliable measures for learning. While this has been studied and many psychometrically valid items have been developed in science contexts, this is not yet the case for engineering design in K-12 settings. This study illustrates several directions for such items, but more are needed to adequately address both performance expectations relating to engineering and the science and engineering practices in the NGSS (NGSS lead states, 2013). This work would benefit from further research into measuring engineering and design practices in K-12 settings and the development of useful items that are not reliant on specific scientific content.

Chapter 7: Conclusions

This dissertation research investigates ways to integrate engineering practices and science concepts (like energy transformation) in classroom settings. I investigate ways to integrate the Next Generation Science Standards (NGSS) performance expectations in these areas and the science and engineering practices while expanding the use of the knowledge integration theory in engineering projects. Prior research has shown that hands-on or engineering-type of projects can be motivating for students (Hmelo et al., 2000; Cantrell et al., 2006), but that students may not develop both scientific understanding and engineering practices from participating in these types of projects (Teacher Advisory Council, 2009). However, using the knowledge integration framework can help when developing tools and activities for hands on projects. For example, the dynamic model developed to use in this project aids students in making sense of energy transformation, while also allowing students to explore engineering design practices.

The findings from this dissertation research suggest promising ways of thinking about the integration of science content and engineering practices and suggest guidelines for designing hands-on engineering projects that enable students to integrate their ideas with practices. However, these findings also show the difficulties in measuring the relatively new area of "engineering practices" and in helping students to develop these practices.

Summary

The research studies presented in this dissertation investigated the following questions:

1. How do students use interactive computer models to integrate science and design during engineering projects?
2. What sources do students use as evidence for design decisions? How can we support students in making decisions based on scientific concepts or evidence?
3. What design principles guide student use of interactive tools (e.g., project report spaces, photos, notebooks, and automated guidance) to support integrated understanding?
4. In classroom instruction, what is an optimal balance of science concept development and engineering design activities to promote integrated understanding?

In this dissertation, I presented five empirical chapters that investigate ways of supporting students in integrating their ideas about energy transformation with ideas about engineering design. The goal of these studies, collectively, is to further research in developing K-12 engineering projects that also integrate science content in a meaningful way, while building on the knowledge integration theory and expanding work on the theory within engineering.

In chapter 2, I investigated ways to combine a dynamic computer model with a hands-on design project to help students integrate science disciplinary knowledge with engineering practices. Findings from this research led to the development of further scaffolds built into the computer modeling activity to support students in making choices based on the trials they ran using the model (e.g., an automatically generated table and reflection questions). Based on this research, the modeling activity was introduced to all students during the initial design phase. We found added benefits for knowledge integration of science disciplinary ideas when students interacted with the model earlier in the curriculum. The model helped students to add more ideas to their repertoire, giving them more ideas to reflect upon later in the unit.

Chapter 3 investigated the function of supports for students using computer models, and explored how students interacted with the model. Findings from this research led to further questions about how best to help students make sense of computer models in both a practical way and a way that allows them to explore the scientific phenomena shown in the model. Specifically, this study showed that when students have access to an automatically generated table to keep track of their trials and results from the model, students may be freer to explore the science concepts shown in the model. Since they do not have to remember the results of each trial, they are also able to compare across trials and make more meaningful decisions based on their comparisons. However, students may not wish to spontaneously generate more than one trial, and must be encouraged to run more than one trial. They also need guidance to consider what the model is able to show and what the limitations of the model are. The virtual model designed for this curriculum does not encompass all possibilities for the design of a solar oven, so students must be able to use the model to make some design decisions, but not all. For example, the virtual model does not currently allow students to manipulate the angle of the reflective flap on their solar oven. However, this is an important aspect to consider in the design of the solar oven. It is important for students to develop the skill of reasoning about what models are and how they can inform decisions, along with other disciplinary ideas and life experiences.

Issues of how students interpret affordances of physical and virtual models are discussed in chapter 4. These ideas about affordances inform development of curricular materials that accompany any modeling activities. The practice of modeling is certainly an important scientific and engineering practice, and is called out across the Next Generation Science Standards. However, making sure students are able to make distinctions about what physically and virtually modeled phenomena are useful for is also important. We must also build on research about opinions of and acceptance of technology (e.g., Ma, Andersson, & Streith, 2005) to understand student perceptions of computer models. For students who have always known a world in which technology is ubiquitous, perceptions of technology may be quite different than for previous generations. Understanding these perceptions will allow us to develop better curriculum around critical thinking and using computer models and data.

Chapter 5 investigated student ideas about light propagation, focusing on the common idea that shiny or dark objects “attract” light to them. I first collected data about the ideas students have on this non-normative idea, then presented a method to automatically score student written responses for the presence of this idea. This automatic scoring algorithm would allow for the development of automated guidance that could then encourage students with this non-normative idea to reconsider what they have written.

This chapter also discusses how instruction could help students to integrate their ideas about light propagation by using a variety of instructional tools to help students add new ideas and then distinguish among all their ideas.

Chapter 6 investigated two ways to frame the curriculum. Since the goals of this curriculum are to integrate both science content ideas and engineering design ideas, I investigated two different frameworks for presenting the curriculum – science-centered or engineering-centered. Findings from this study indicate that students learn content differently based on the framing of the project. Students who experienced the project framed with an engineering focus make greater learning gains in engineering practices, while students who experienced the project framed as a science project made greater gains in science content.

Together, these chapters show specific findings about the structure of hands-on projects that aim to teach both science content and engineering design. Using computer models has proven useful in helping students integrate science and engineering. However, these models require scaffolding to help students integrate their ideas and there should be careful consideration about when during the curriculum they are introduced. In addition, it is critical to encourage students to integrate their science ideas relevant to the hands-on engineering project. Including science content in a meaningful way is critical for the success of hands-on projects. Instructional designers need to anticipate student dilemmas and design knowledge integration activities to help students develop a firm foundation for the engineering design decisions they make. In addition, it is important to consider the goals for learning when considering how to structure the curriculum for hands-on projects integrating science and engineering.

Changes in the availability of technology have lead to exciting changes in online instructional tools. This research explores some of the issues with new technologies like virtual models and automated guidance. There are also further questions brought up about the best way to design instructional methods that address complex non-normative ideas. Future research on this topic may utilize the automated guidance discussed in chapter 5, building on other research about how to develop that guidance (e.g., Tansomboon, Gerard, Vitale, & Linn, 2017; Gerard, Ryoo, McElhaney, Liu, Rafferty, & Linn, 2016).

Changes in computer modeling environments and the ability to connect elements like tables of results with visualizations and graphical results from trials have made connecting science concepts and engineering design practices easier. However, careful consideration is still necessary when constructing scaffolds and uses for the computer models. Just because we can build a model of something does not mean it will be useful for student learning.

Designing Engineering Projects

Through the research presented in this dissertation, we have developed insights about the features that are important in engineering projects. These features may differ from hands-on projects, as they include specific constraints and tools that are important in developing engineering practices.

Several features from research on solar ovens characterize effective designs for engineering projects. These features include:

- Structure of the project: using iterative, engineering cycles
- Design constraints: budgeting, time, materials, requirements
- Using a virtual model to test ideas
- Using peer critique to improve design

The structure of this project was built around the common engineering design cycle of "design, build, test". We include reflecting as a fourth step in this iterative cycle to emphasis to students that they should use results and evidence from their tests to make design modifications in the next iteration. This cycle works for the structuring of engineering projects at multiple levels of granularity. Overall, the project is structured to guide students through the design, build, test cycle. However, within each design section of the project, students are also working on iteratively refining their ideas about what is important in their designs by using virtual modeling environments and other resources. Within a design session, students may not be building a physical solar oven, but they are building virtual models and testing them. Because of the iterative nature of this engineering cycle, and the emphasis on students refining their design ideas throughout the project by using different evidence sources (e.g., the results from a physical or virtual test), this engineering design cycle is extremely compatible with the knowledge integration framework.

Other engineering design cycles may also be useful in an education context, namely those like rapid prototyping, wherein students would prototype smaller pieces of a project to test that the components function separately before putting them together. Rapid prototyping may be more important to implement in projects that have moving parts or are made up of more components than a solar oven. For example, rapid prototyping may be a useful engineering design cycle to implement in a project about building self-propelled (rubber band) cars.

Another aspect of engineering projects that we found to be important in emphasizing the practices of working engineers is that of providing constraints. In the solar ovens project this is done by providing a set of materials that students must use, as well as by providing a budget for materials that students must work within. In our project, each material available has an associated price. During the design phase of the project, students must select their materials from those available. The budget they must work within makes sure that students cannot buy all the available materials, and must therefore make choices about what they believe is most important for their design. These constraints force students to consider tradeoffs between the different ways they could potentially use their budget. Considering tradeoffs among different potential designs is an important aspect of engineering thinking, and can only be facilitated by providing strict constraints for students.

There may also be other ways to provide constraints in engineering projects. For example, providing constraints about what the final product must look like or be able to do are also useful in different ways. Providing constraints for the design itself, instead of in the materials students may use in design, would be more useful in a project with different goals. For example, if students are tasked with designing a city park, designing a specific set of needs the park must meet would be a useful set of constraints to place on the project.

Using virtual models to provide a testing environment for students has also proven to be an important aspect of engineering projects in this research. Practicing engineers often use virtual models to test alternative designs before spending time and resources building and testing those designs in the real world. By providing this resource to students, we allow students to test the many ideas they may have, but also to engage in a different type of designing, building, and testing. Making comparisons among the tests they run is an important engineering skill, which is also outlined in the NGSS engineering performance expectations (NGSS Lead States, 2013).

A final aspect of engineering projects we have implemented in this work is that of peer critique. Often, engineers will engage in design review. This type of review is meant to provide engineers with critical feedback on their design so that engineers can learn from their peers and the design can improve. In the solar ovens project, students must also be able to present their design to their peers and be able to sort through different design ideas their peers may provide to determine whether those ideas are useful or not. This aligns with the knowledge integration practice of distinguishing among ideas. The practice of peer critique is important in many fields besides engineering, which should further increase the importance of including this aspect of engineering projects in all educational engineering projects.

These four features of engineering projects have proven useful for our design of the solar ovens unit used in this research. Through design-based research, we have iteratively modified how each of these features occurs in the solar ovens curriculum. These features come from both engineering practice and the performance expectations and practices within standards (NGSS Lead States, 2013). By using these features in all engineering projects, we can make sure hands-on projects are more focused on the aspects of engineering thinking that are important to instill in students.

References

Andersson, B., & Wallin, A. (2000). Students' understanding of the greenhouse effect, the societal consequences of reducing CO2 emissions and the problem of ozone layer depletion. *Journal of Research in Science Teaching*, *37*(10), 1096-1111.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, *26*(3-4), 369-398.

Bybee, R. W. (2011). Scientific and engineering practices in K-12 classrooms. *Science Teacher*, *78*(9), 34-40.

Cantrell, P., Pekcan, G., Itani, A., & Velasquez-Bryant, N. (2006). The effects of engineering modules on student learning in middle school science classrooms. *Journal of Engineering Education*, *95*(4), 301.

Chang, H. Y., & Linn, M. C. (2013). Scaffolding learning from molecular visualizations. *Journal of Research in Science Teaching*, *50*(7), 858-886.

Chiu, J. L., & Linn, M. C. (2011). Knowledge integration and WISE engineering.*Journal of Pre-College Engineering Education Research (J-PEER)*, *1*(1), 2.

Chiu, J. L., Malcolm, P. T., Hecht, D., DeJaegher, C. J., Pan, E. A., Bradley, M., & Burghardt, M. D. (2013). WISEngineering: Supporting precollege engineering design and mathematical understanding. *Computers & Education*, *67*, 142-155.

Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching*, *30*(10), 1241-1257.

Coll, R. K., France, B., & Taylor, I. (2005). The role of models/and analogies in science education: implications from research. *International Journal of Science Education*, *27*(2), 183-198.

Conati, C., Fratamico, L., Kardan, S., & Roll, I. (2015, June). Comparing representations for learner models in interactive simulations. In *International Conference on Artificial Intelligence in Education* (pp. 74-83). Springer, Cham.

Crismond, D. (2001). Learning and using science ideas when doing investigate-and-redesign tasks: A study of naive, novice, and expert designers doing constrained and scaffolded design work. *Journal of Research in Science Teaching*, *38*(7), 791-820.

Crouch, C., Fagen, A. P., Callan, J. P., & Mazur, E. (2004). Classroom demonstrations: Learning tools or entertainment?. *American Journal of Physics*, *72*(6), 835-838.

De Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science*, *340*(6130), 305-308.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68*(2), 179-201.

DeWinstanley, P. A., & Bjork, E. L. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Memory & Cognition*, *32*(6), 945-955.

DiSessa, A. A. (2006). *A History of Conceptual Change Research: Threads and Fault Lines*. Cambridge University Press.

Donnelly, D. F., Linn, M. C., & Ludvigsen, S. (2014). Impacts and characteristics of computer-based science inquiry learning environments for precollege students. *Review of Educational Research*, *84*(4), 572-608.
Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*(3), 287-312.
Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education*, *94*(1), 103-120.
Elger, D. F., Beyerlein, S. W., & Budwig, R. S. (2000). Using design, build, and test projects to teach engineering. In *Frontiers in Education Conference, 2000. FIE 2000. 30th Annual* (Vol. 2, pp. F3C-9). IEEE.
Ertmer, P. A. (1999). Addressing first-and second-order barriers to change: Strategies for technology integration. *Educational Technology Research and Development*, *47*(4), 47-61.
Ertmer, P. A. (2005). Teacher pedagogical beliefs: The final frontier in our quest for technology integration?. *Educational Technology Research and Development*, *53*(4), 25-39.
Gerard, L. F., Ryoo, K., McElhaney, K. W., Liu, O. L., Rafferty, A. N., & Linn, M. C. (2016). Automated guidance for student inquiry. *Journal of Educational Psychology*, *108*(1), 60.
Gobert, J. D., Sao Pedro, M., Raziuddin, J., & Baker, R. S. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences*, *22*(4), 521-563.
Gobert, J. D., Kim, Y. J., Sao Pedro, M. A., Kennedy, M., & Betts, C. G. (2015). Using educational data mining to assess students' skills at designing and conducting experiments within a complex systems microworld. *Thinking Skills and Creativity*, *18*, 81-90.
Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. *Educational Psychologist*, *50*(1), 43-57.
Hmelo, C., & Day, R. (1999). Contextualized questioning to scaffold learning from simulations. *Computers & Education*, *32*(2), 151-164.
Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to learn about complex systems. *The Journal of the Learning Sciences*, *9*(3), 247-298.
Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn?. *Educational Psychology Review*, *16*(3), 235-266.
Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding complex systems: Some core challenges. *The Journal of the Learning Sciences*, *15*(1), 53-61.
Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *The Journal of the Learning Sciences*, *16*(3), 307-331.
Jackson, G. T., Graesser, A. C., & McNamara, D. S. (2009, July). What students expect may have more impact than what they know or feel. In *AIED* (pp. 73-80).
Johnson, C. C., Peters-Burton, E. E., & Moore, T. J. (Eds.). (2015). *STEM Road Map: A Framework for Integrated STEM Education*. Routledge.
Joseph, E. (2005). Engagement tracing: using response times to model student disengagement. *Artificial Intelligence in Education: Supporting learning through intelligent and socially informed technology*, *125*, 88.

Kali, Y. (2006). Collaborative knowledge building using the Design Principles Database. *International Journal of Computer-Supported Collaborative Learning*, *1*(2), 187-201.
Kali, Y., & Linn, M. C. (2008). Technology-enhanced support strategies for inquiry learning. *Handbook of Research on Educational Communications and Technology*, 145-161.
Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., ... & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design (tm) into practice. *The Journal of the Learning Sciences*, *12*(4), 495-547.
Koulaidis, V., & Christidou, V. (1999). Models of students' thinking concerning the greenhouse effect and teaching implications. *Science Education*, *83*(5), 559-576.
Krajcik, J.S. & Blumenfeld, P. (2006). Project-based learning. In Sawyer, R. K. (Ed.), *the Cambridge Handbook of the Learning Sciences*. New York: Cambridge.
Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., Fredricks, J., & Soloway, E. (1998). Inquiry in project-based science classrooms: Initial attempts by middle school students. *Journal of the Learning Sciences*, *7*(3-4), 313-350.
Krajcik, J., Blumenfeld, P., Marx, R., & Soloway, E. (1998). Instructional, curricular, and technological supports for inquiry in science classrooms.
Krajcik, J. S., Blumenfeld, P. C., Marx, R., & Soloway, E. (2000). Instructional, curricular, and technological supports for inquiry in science classrooms. In J. Minstrell & E. H. Van Zee (Eds.), *Inquiring into Inquiry Learning and Teaching in Science* (pp. 283–315). American Association for the Advancement of Science.
Krajcik, J. S., & Blumenfeld, P. (2006). Project based learning. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences*. New York: Cambridge University Press.
Lachapelle, C. P., Hertel, J. D., Shams, M. F., & Cunningham, C. M. (2013). *Evaluating the impact of Engineering is Elementary: Years 3 and 4 of implementation in Minneapolis and Hopkins Public Schools*. Boston, MA: Museum of Science.
Lewis, T. (2006). Design and inquiry: Bases for an accommodation between science and technology education in the curriculum.. *Journal of Research in Science Teaching*, *43*(3), 255-281.
Lin, X., & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching*, *36*(7), 837-858.
Linn, M. C. (2005). WISE design for lifelong learning—pivotal cases. *Cognition, Education, and Communication Technology*, 223-256.
Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration. *Science Education*, *87*(4), 517-538.
Linn, M. C., & Eylon, B. S. (2011). *Science learning and instruction: Taking Advantage of Technology to Promote Knowledge Integration*. Routledge.
Linn, M. C., Gerard, L., Ryoo, K., McElhaney, K., Liu, O. L., & Rafferty, A. N. (2014). Education technology. Computer-guided inquiry to improve science learning. *Science (New York, N.Y.)*, *344*(6180), 155–6. doi:10.1126/science.1245980
Linn, M. C., & Hsi, S. (2000). *Computers, Teachers, Peers: Science learning partners*. Routledge.

Linn, M. C., Lee, H. S., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and assessing knowledge integration in science. *Science*, *313*(5790), 1049-1050.
Linn, M. C., & Songer, N. B. (1991). Teaching thermodynamics to middle school students: What are appropriate cognitive demands. *Journal of Research in Science Teaching*, *28*(10), 885-918.
Liu, O. L., Lee, H. S., Hofstetter, C., & Linn, M. C. (2008). Assessing knowledge integration in science: Construct, measures, and evidence. *Educational Assessment*, *13*(1), 33-55.
Ma, W. W. K., Andersson, R., & Streith, K. O. (2005). Examining user acceptance of computer technology: An empirical study of student teachers. *Journal of Computer Assisted Learning*, *21*(6), 387-395.
McBride, E. A., Vitale, J. M., Applebaum, L., & C Linn, M. (2016). Use of interactive computer models to promote integration of science concepts through the engineering design process. Singapore: International Society of the Learning Sciences.
McBride, E.A., Vitale, J.M., Linn, M.C. (June 2017). Student Use of Scaffolded Inquiry Simulations in Middle School Science. Poster presented at the 10th International Conference on Educational Data Mining, Wuhan, China.
McElhaney, K.W., & Linn, M.C. (2011). Investigations of a complex, realistic task: Intentional, unsystematic, and exhaustive experimenters. *Journal of Research in Science Teaching*, *48*(7), 745-770.
McElhaney, K.W., Chang, H. Y., Chiu, J. L., & Linn, M.C. (2015). Evidence for effective uses of dynamic visualisations in science curriculum materials. *Studies in Science Education*, *51*(1), 49-85.
Minstrell, J. (1982). Explaining the "at rest" condition of an object. *The Physics Teacher*, *20*(1), 10-14.
National Research Council. (2009). *Engineering in K-12 education: Understanding the status and improving the prospects*. National Academies Press.
NGSS Lead States. (2013). *Next generation science standards: For states, by states*. National Academies Press.
Noorani, R. (2006). *Rapid prototyping: principles and applications*. John Wiley & Sons Incorporated.
Olson, S., & Loucks-Horsley, S. (Eds.). (2000). *Inquiry and the National Science Education Standards: A guide for teaching and learning*. National Academies Press.
Perkins, D., & Grotzer, T. (2005). Dimensions of Causal Understanding: the Role of Complex Causal Models in Students' Understanding of Science. *Studies in Science Education, 41*(1), 117-165.
Podolefsky, Moore, & Perkins. (2013). Implicit scaffolding in interactive simulations: Design strategies to support multiple educational goals. *Chemistry Education Research and Practice*, 14(3):257.
Purzer, Ş., Goldstein, M. H., Adams, R. S., Xie, C., & Nourian, S. (2015). An exploratory study of informed engineering design behaviors associated with scientific explanations. *International Journal of STEM Education*, *2*(1), 9.
Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences*, *13*(3), 273-304.
Richland, L. E., Bjork, R. A., Finley, J. R., & Linn, M. C. (2005). Linking cognitive science to education: Generation and interleaving effects. In *Proceedings of the twenty-seventh annual conference of the Cognitive Science Society* (pp. 1850-1855).

Rogers, C., & Portsmore, M. (2004). Bringing engineering to elementary school. *Journal of STEM Education: innovations and research*, *5*(3/4), 17.
Rye, J. A., & Rubba, P. A. (1998). An exploration of the concept map as an interview tool to facilitate the externalization of students' understandings about global atmospheric change. *Journal of Research in Science Teaching*, *35*(5), 521-546.
Ryoo, K., & Linn, M. C. (2012). Can dynamic visualizations improve middle school students' understanding of energy in photosynthesis?. *Journal of Research in Science Teaching*, *49*(2), 218-243.
Ryoo, K., & Linn, M. C. (2014). Designing guidance for interpreting dynamic visualizations: Generating versus reading explanations. *Journal of Research in Science Teaching*, *51*(2), 147–174. doi:10.1002/tea.21128
Sandoval, W. A., & Morrison, K. (2003). High school students' ideas about theories and theory change after a biological inquiry unit. *Journal of Research in Science Teaching*, *40*(4), 369-392.
Sao Pedro, M. A., de Baker, R. S., Gobert, J. D., Montalvo, O., & Nakama, A. (2013). Leveraging machine-learned detectors of systematic inquiry behavior to estimate and predict transfer of inquiry skill. *User Modeling and User-Adapted Interaction*, *23*(1), 1-39.
Sato, M. E. (2015). *Designing Critique for Knowledge Integration*. University of California, Berkeley.
Silk, E. M., Schunn, C. D., & Cary, M. S. (2009). The impact of an engineering design curriculum on science reasoning in an urban setting. *Journal of Science Education and Technology*, *18*(3), 209-223.
Slotta, J. D., & Linn, M. C. (2009). *WISE science: Web-based inquiry in the classroom*. Teachers College Press.
Smith III, J. P., Disessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, *3*(2), 115-163.
Snir, J., Smith, C., & Grosslight, L. (1993). Conceptually enhanced simulations: A computer tool for science teaching. *Journal of Science Education and Technology*, *2*(2), 373-388.
Svihla, V., & Linn, M. C. (2012). A design-based approach to fostering understanding of global climate change. *International Journal of Science Education*, *34*(5), 651-676.
Tang, X., Coffey, J. E., Elby, A., & Levin, D. M. (2010). The scientific method and scientific inquiry: Tensions in teaching and learning. *Science Education*, *94*(1), 29-47.
Tansomboon, C., Gerard, L. F., Vitale, J. M., & Linn, M. C. (2017). Designing Automated Guidance to Promote Productive Revision of Science Explanations. *International Journal of Artificial Intelligence in Education*, 1-29.
Teacher Advisory Council. (2009). *Engineering in K-12 Education: Understanding the Status and Improving the Prospects*. National Academies Press.
White, R., & Gunstone, R. (1992). Probing understanding. London and New York: The Falmer Press.
Wilensky, U. (1999). {NetLogo}.
Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep, or a firefly: Learning biology through constructing and testing computational theories—an embodied modeling approach. *Cognition and Instruction*, *24*(2), 171-209.
Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, *8*(1), 3-19.

Wilkerson-Jerde, M. H., & Wilensky, U. J. (2015). Patterns, probabilities, and people: Making sense of quantitative change in complex systems. *Journal of the Learning Sciences*, *24*(2), 204-251.
Zee, E. H. V., & Minstrell, J. (1997). Reflective discourse: Developing shared understandings in a physics classroom. *International Journal of Science Education*, *19*(2), 209-228.
Zhang, H. Z. (2010, June). Exploring drawing and critique to enhance learning from visualizations. In *Proceedings of the 9th International Conference of the Learning Sciences-Volume 2* (pp. 234-235). International Society of the Learning Sciences.