Running head: CHILDREN'S MATERIALS SCIENCE PRACTICES

# Children's Materials Science Practices Before, During, and After a Design-Based Science Learning Experience

Kristen Wendell

Tufts University

Qualifying Paper #2

May 14, 2009

## Table of Contents

Literature Review ........ 8
Materials Science: The Study and Application of Material and Object Properties ........ 8
Existing Approaches to Instruction for Introductory Materials Science ........ 11
Theoretical Frameworks for Enabling Science Learning ........ 13
This Study's Materials Science Curriculum ........ 17
Existing Methods for Assessing Children's Materials Science Learning ........ 20
Previous Findings on Children's Materials Science Learning ........ 23
This Study's Measures for Assessing Materials Science Learning ........ 30
Method ........ 33
Setting ........ 33
Participants ........ 34
Data Sources ........ 36
Data Reduction and Analysis ........ 42
Results ........ 47
Design Task Approaches in Clinical Interviews ........ 47
Changes from Preinstruction to Postinstruction ........ 53
Connecting Pre/Post Changes to Instructional Experiences ........ 56
Deepening the Interview-Instruction Connection: Examples from Individual Students ........ 58
Discussion ........ 68
The Children's Approaches to Materials Selection Design Tasks ........ 69
How the Children's Approaches Changed Over Time ........ 72
Connections Between Interview Changes and Instructional Experiences ........ 73
Limitations of This Work ........ 77
Questions for Further Investigation ........ 79
References ........ 81
Appendix A. National and State Standards of Learning Related to the Properties of Objects and Materials ........ 88
Appendix B. Overview of the *Design a Model House* Curriculum Module ........ 89
Appendix C. Scoring Rubric for Interview Responses ........ 92
Appendix D. Scoring Rubric for Student Workbook Pages ........ 97
Appendix E. List of All Descriptors Generated by Students ........ 101
Appendix F. Mean Scores on Selected Workbook Segments ........ 106

## List of Figures

*Figure 1.* Sample of workbook pages. ........ 41
*Figure 2.* Sample scoring of workbook page 3-3. ........ 46
*Figure 3.* Histograms of the students' scores on the stepstool task, habitat task, and total interview ........ 54
*Figure 4.* Histograms of the students' scores in each of the interview's science-practice categories ........ 54

*Figure 5.* Students' gains on the interview design tasks plotted against their scores on the corresponding workbook exercises ........ 57
*Figure 6.* Workbook pages by Macon, Stella, Julia, and Chinelle. ........ 67

## List of Tables

Table 1. Key Practices and Content Knowledge of Expert and Introductory Materials Science . 10
Table 2. Profiles of the 9 Student Participants ........ 36
Table 3. Construct Analysis with Examples of Introductory Materials Science Practices........ 37
Table 4. Summary of Topics and Practices Assessed by Clinical Interviews ........ 40
Table 5. Generalized Version of the Scoring Rubric for Interview Responses ........ 43
Table 6. Material Selection Rationales Offered by All Students ........ 49
Table 7. Material Property Tests Proposed by the Overall Group of Students........ 50
Table 8. Average Pre-Instruction, Post-Instruction, and Gain Scores on Design Task Interviews ........ 53
Table 9. Frequency Table for Key Material Selection Practices........ 55
Table 10. Frequency Table for Key Material Property Testing Practices ........ 56
Table 11. Interview Gains and Workbook Scores by Individual Student........ 59

The study of matter and atomic molecular theory constitutes a major conceptual field within the physical sciences[1] (American Association for the Advancement of Science [AAAS], 1993; National Research Council [NRC], 1996; Smith, Wiser, Anderson, & Krajcik, 2006). One of the big ideas within this field is that the objects and materials of the physical world can be characterized through measurement, classification, and description according to their properties (Smith et al., 2006). For professional scientists and engineers, the practice of specifying objects and materials by measurable physical properties has led to the creation of an entire discipline called *materials science* (Callister, 2007). Expert materials scientists have deep knowledge about the structure and composition of both natural and synthetic materials, and they are skilled at conducting tests to determine properties of materials, selecting appropriate materials for particular tasks, using material properties to justify selections, and describing materials' invariant, intensive properties to other scientists.

Because the characterization of objects and materials is an important part of scientists' study of matter, it follows that materials science also has an important role to play in children's science learning. Fluency with the ideas and practices of materials science may provide building blocks for the construction of a particulate theory of matter (NRC, 1996; Smith et al., 2006).

Despite having reason to believe that children benefit from holding a scientific perspective on materials and objects, our knowledge of how to help children develop materials science proficiency is limited. Most research on children's understandings and abilities related to materials has to do with classification schemes, differentiation of matter and nonmatter, and

[1] Helping students develop a particulate perspective on the nature of matter is a major goal of primary and secondary science education (NRC, 1996; Smith et al., 2006). A rich understanding of the composition and properties of matter is fundamental to further study in many disciplines, including biochemistry and medicine, structural engineering, and environmental science. Scientists' knowledge about the nature of matter has enabled numerous important technological advances, including laser surgery, skyscrapers, space travel, pharmaceuticals, and renewable energy devices (Constable & Somerville, 2003).

differentiation of objects from their material composition. There is little research on children's abilities at testing and selecting materials for particular tasks.

Instead, many studies take the form of clinical interviews in which children sort objects or material samples into groups of similar samples. By having the children explain their sorting strategy, researchers are able to determine which characteristics of materials and objects children have learned to identify (e.g., Dickinson, 1987; Inhelder & Piaget, 1964; Krnel, Glazar, & Watson, 1998). In some studies, researchers probe for specific modes of classification, such as sorting by constituent material, state of matter, intensive properties like density or texture, or extensive properties like weight or shape (e.g., Krnel, Glazar, & Watson, 2003). In other studies, researchers focus on children's differentiation of object from material kind (e.g., Dickinson, 1987; Smith, Carey, & Wiser, 1985).

Although these studies have generated much knowledge about how children think about objects and materials, they have not told us whether or how children apply that thinking to select objects and materials for practical tasks. We do not know if the ability to classify materials enables children to take action with those materials to design useful everyday devices. For example, if a child can sort materials by their hardness, does that mean she can select an appropriate material for a table top?

In this paper, I propose that children may develop materials science abilities within a science learning environment that is focused on the *use* of materials and objects, rather than only the *classification* and *study* of them. Such a learning environment would include design-based science instruction, an approach supported by the theories of distributed cognition (Hutchins, 1995) and situated cognition (Brown, Collins, & Duguid, 1989; Lave & Wenger, 1991) and consistent with a situative, social constructivist perspective on learning (Driver, Asoko, Leach, Mortimer, & Scott,1994; Greeno, 1998).

In light of the limited research on children's abilities related to materials selection and testing, and in response to theories that align design-based learning with improved understanding and ability in science practices, the aim of this study is to investigate how children's approaches to materials science tasks change after design-based science instruction on material and object properties. An additional goal is to explore how the changes in their approaches might be associated with the science practices they exhibit during classroom instruction. Further, this study will examine whether children's approaches to materials science change even for tasks that require reasoning about properties not addressed through explicit instruction. In this study, those properties include the physically complex but common everyday properties of strength and thermal insulation.

Two formal research questions guided this study:

a) How do children approach materials selection design tasks in a clinical interview setting, before and after they have participated in a design-based science curriculum module on material and object properties?

b) How do the pre/post changes in children's interview performance compare to their engagement with and performance on materials science practices throughout the curriculum module?

Three aspects of this study distinguish it from previous work. First, its primary measurement tool is a clinical interview assessment structured around design tasks rather than traditional classification tasks. Second, instead of relying only on pre/post measures, it uses students' written work from throughout the instructional intervention to track the source of performance changes. Third, its instructional intervention is a design-based science curriculum module (Fortus, Krajcik, Dershimer, Marx, & Mamlok-Naaman, 2005; Kolodner, 2006; Roth, 1996), which situates instruction within an engineering design challenge intended to foster exploration of specific science content and practices.

I expect that by exposing students to the tools and discourse of materials science, the curriculum enacted for this study will begin to enculturate students into the practices of selecting, testing, and describing materials. Therefore, I hypothesize that after instruction, students will more often discuss intensive physical properties as they describe their material selection choices, and they will more often propose material tests that involve quantified measures and fair comparisons. I also predict that the selections, tests, and descriptions that students generate during postinstruction interviews will be associated with the writing and drawing they do in their science workbooks throughout instruction itself.

To answer the above research questions, this paper is organized into four main sections. In the first section, I provide more detail about the discipline of materials science, and I review other research related to children's materials science learning. In this section I also explain the instructional intervention used in this study. In the method section, I present this study's research design by describing its participants, data sources, and data analysis procedures. Then in the third section, I present the results of that analysis. I conclude the paper by comparing the results to previous work, summarizing the contributions and limitations of this study, and suggesting future research directions.

# Literature Review

*Materials Science: The Study and Application of Material and Object Properties*

*Materials science at the expert level.* What does the discipline of materials science consist of, at its most sophisticated level? What comprises a scientific perspective on materials? Expert materials scientists are skilled at particular scientific practices and knowledgeable about particular scientific content. The primary practical skills of expert materials scientists include a) analyzing existing materials for their constituent parts, structure, and physical properties, b) identifying the properties needed for particular tasks and devices, c) selecting materials with those properties, d) uniquely specifying materials by intensive properties that are invariant to the amount of material, and e) creating new materials with desired properties (Ashby, Shercliff, & Cebon, 2007; Dowling, 1998). The primary content knowledge of materials scientists deals with a) how particular properties relate to particular technological challenges, b) how materials are composed and structured, c) how macroscopic properties *generally* relate to materials' structure and composition, and d) how *specific* macroscopic properties are determined by a material's structure and composition (Allen & Thomas, 1999; Schaffer, Saxena, Antolovich, Sanders, & Warner, 1999). Some of the properties considered most important and studied most frequently by materials scientists are the mechanical qualities of compressive strength, elastic modulus, density, and the thermal qualities of heat capacity, thermal conductivity, and expansion coefficient (Ashby, Shercliff, & Cebon, 2007). These properties are crucial to the success and safety of many physical technologies. Buildings, vehicles, conduits, manufacturing machines, and computers are all subject to mechanical and thermal stress (Ashby & Jones, 2005; Ashby, Shercliff, & Cebon, 2007).

*Materials science at the introductory level.* In expert scientific practice, materials science is an important aspect of the study of the nature of matter. Hence, the exploration of materials and their properties is an important part of children's beginning studies of matter. The understanding that "material kinds have characteristic properties that can be measured and explained" is, in fact, a major component within the study of matter and atomic molecular theory (Smith et al., 2006, p.14). In this paper, I will use the term *introductory materials science* to refer to children's exploration of this idea about objects and materials.

Beyond the fundamental idea that objects and materials can be characterized by their properties, what does introductory materials science entail? What does materials science look like for primary grade students, and what materials science practices and content should students learn as part of their initial study of the nature of matter? Educational benchmark documents reveal high expectations for children's abilities related to the science of materials. Both the *Massachusetts Curriculum Frameworks* (2001) and the *National Science Education Standards* (NRC, 1996) state that by upper elementary school, students should be able to describe and measure the observable properties of materials and objects, choose suitable materials for a particular task, and justify their choice of materials (see Appendix A). The AAAS *Benchmarks for Science Literacy* (1993) mention in particular the properties of strength, hardness, flexibility, water resistance, fire resistance, ease of heat conduction, buoyancy, and magnetism. Smith et al. (2006) agree with the *Standards* that students should be able to measure, classify, and describe materials according to their properties, which include directly observable properties like color and hardness as well as less obvious properties like density, flammability, and conductivity.

This set of learning goals can be considered a pared down list of the practices and content knowledge of expert materials scientists. To summarize, introductory materials science involves the practices of a) testing materials for particular properties, b) identifying the material properties important to a given task, c) selecting materials with those properties, d) describing materials specifically via intensive and extensive properties, and e) proposing combinations of materials that would accomplish a task. Additionally, introductory materials science involves content knowledge about a) the difference between an object and its constituent material and b) the macroscopic properties that are important for very common mechanical tasks. Table 1 compares introductory materials science practices and content with the list for expert materials science. It is important to note that in describing introductory materials science, I am summarizing just one area of study within the much larger field of understandings about matter. Regarding the nature of matter, there are many additional important practices and concepts to learn.

Table 1
*Key Practices and Content Knowledge of Expert and Introductory Materials Science*

| | Expert Materials Science | Introductory Materials Science |
|---|---|---|
| Key Scientific Practices | a) Analyzing existing materials for their constituent parts and physical properties | a) Testing materials for particular properties |
| | b) Identifying the properties needed for particular tasks and devices | b) Identifying the material properties important to a given task |
| | c) Selecting materials with those properties | c) Selecting materials with those properties |
| | d) Uniquely specifying materials by intensive properties that are invariant to the amount of material | d) Describing materials specifically via intensive and extensive properties |
| | e) Creating new materials with desired properties | e) Proposing combinations of materials that would accomplish a task |
| Key Scientific Content Knowledge | a) How particular properties relate to particular technological challenges | a) A few of the properties important for very common mechanical tasks |
| | b) How materials are composed and structured | b) How objects are different from their constituent materials |
| | c) How macroscopic properties *generally* | |

| | | |
|---|---|---|
| | relate to materials' structure and composition<br>d) How *specific* macroscopic properties are particularly determined by a material's structure and composition | |

*Existing Approaches to Instruction for Introductory Materials Science*

I have specified the practices and content knowledge that comprise materials science for primary grade students. Next I will summarize four kinds of learning experiences that have been developed to foster children's learning related to materials science. The first kind includes classroom activities that give children experience measuring, comparing, classifying, and representing information about material samples, mainly at the macroscopic level. This information may be based on intensive criteria, which do not vary with the size of an object or amount or a substance, or extensive criteria, which do vary with size or amount (Schwartz, 1996). Examples of learning experiences within this category include the *Science and Technology for Children* (*STC)* primary units on "Solids and Liquids" and "Rocks and Minerals" (National Science Resources Center [NSRC], 2005), the *Full Option Science System* (*FOSS*) primary units on "Wood and Paper" and "Earth Materials" (Lawrence Hall of Science [LHS], 2005), and several of the learning performances for K-5 students proposed by Smith et al. (2006). Related to these examples, but at a more sophisticated middle-school level, are instructional activities in which students determine particular properties of material samples and then go on to account for those properties in scientific, quantitative ways. Smith's (2007) work with eighth graders on the specific properties of weight, mass, and volume, and Smith et al.'s (Smith, Maclin, Grosslight, & Davis, 1997) curriculum units that have students build models of density, mass, and volume, are examples of these kinds of experiences. Some teaching studies extend these experiences to include the evaluation of computer-based models of material phenomena (Snir, Smith, & Raz, 2003; Smith et al., 1997).

The second category of instructional activities for children's materials science includes curricula that foster reflection on how measurement devices work and on why different devices can result in different measures of material properties. The *STC* units on "Comparing and Measuring" and "Balancing and Weighing" (NSRC, 2005), and the *FOSS* unit on "Measurement" (LHS, 2005), all provide for this type of reflection. Smith et al. (2006) suggest that learning performances for third through fifth grade students should involve critical thinking about appropriate measurement units and devices.

In the third group of introductory materials science curricula are experiences that enable children to witness transformations of material samples and then require them to explain those transformations. The *STC* "Changes" unit (NSRC, 2005), the *Insights* "Changes of State" unit (Education Development Center [EDC], 2003), and the *FOSS* "Mixtures and Solutions" unit (LHS, 2005) all include instructional activities that focus on observing and explaining material transformations. Teaching studies by Acher and Arcà have given young children opportunities to study transformations and model them through drawing, gestures, and role play (Acher & Arcà, 2006; Acher, Arcà, & Sanmartí, 2007). In middle school, these kinds of activities typically also intend for students to formulate a general conceptualization of matter that can account for state changes. Lee et al. (Lee, Eichinger, Anderson, & Berkheimer, 1993) and Johnson (2000) require middle schoolers to do this kind of work, as do the classic *Introductory Physical Science* curriculum (Haber-Schaim, 1999), the *STC* "Properties of Matter" middle school unit (NSRC, 2006), and the *FOSS* "Chemical Interactions" middle school unit (LHS, 2005).

The hallmark of the fourth category of introductory materials science curricula is scientific argumentation about material identification. Curricula in this category provide children with experiences in making and evaluating claims about the identity of an unknown substance. Smith et al. (2006) propose that this kind of activity be used as a learning performance for K-5 students. Formal curriculum units that deal with claims and evidence about material identities include the *STC* "Chemical Tests" and "Food Chemistry" units (NSRC, 2005), the *Insights* "The Mysterious Powder" unit (EDC, 2003), and the *FOSS* "Earth Materials" unit (LHS, 2005).

None of the above four categories focuses on providing opportunities to use material samples in the design, construction, and testing of artifacts that serve authentic purposes. The research literature related to introductory materials science seldom makes references to these kinds of design-focused learning activities. However, a few published science units include lessons that fit into this category. They include the *STC* "Floating and Sinking" unit in which upper elementary students create clay boats (NSRC, 2005), the *Insights* "Structures with Trusses" unit for sixth graders (EDC, 2003), and the *FOSS* "Solids and Liquids" unit (LHS, 2005), which includes one tower-building activity for young students. These three units represent a small fraction of the introductory materials science instruction enacted in primary classrooms; most instruction and research endeavors have instead focused on students' determining and explaining material properties, witnessing and explaining material phenomena like changes of state, and conducting scientific argumentation about material properties and changes. Little work has been done on creating materials science learning environments that are centered on the *use* of materials and objects, rather than the *classification* and *study* of them.

### *Theoretical Frameworks for Enabling Science Learning*

In response to this analysis of existing work, I propose an alternative approach to fostering children's learning of introductory materials science. However, before specifying my approach, it is important to acknowledge the range of theoretical frameworks that have informed research-based approaches to children's science learning. The different frameworks can be characterized by how they define *learning* in science.

*Cognitive perspective.* A number of frameworks fall into a broad category that can be called the individual constructivist or cognitive perspective on learning. Researchers who work within these frameworks often base their thinking in the work of Piaget (1929; 1930), who studied humans' personal construction of knowledge about natural phenomena via their everyday interaction with the world. Broadly, from the cognitive or individual constructivist viewpoint, science learning is the harnessing of prior knowledge to acquire more advanced, complex scientific understandings (Smith, diSessa, & Roschelle, 1993). This learning happens when an individual's personal interactions with physical entities and events provide necessary opportunities for deep cognitive conflict followed by reflection (Posner, Strike, Hewson, & Gertzog, 1982). When this general perspective on learning is applied to the science classroom, science learning requires practical activities that challenge an individual's existing state of understanding.

Each framework within the cognitive/individual constructivist perspective offers a slightly different viewpoint on how that state of understanding *changes.* The conceptual change framework (Carey, 1985) is one of these individual constructivist models of learning. There are different kinds of conceptual change: a) the differentiation of concepts such as weight from density, b) the coalescence of concepts such as plant and animal into living thing, c) the reanalysis of relationships, and d) shifts of concepts from the core to the periphery of a system of understanding (Carey & Spelke, 1996). The mechanisms of conceptual change can include mappings across existing systems of understanding as well acquisition of new concepts to produce knowledge restructuring (Carey & Spelke, 1996).

Differing from the conceptual change framework, but still in line with the individual constructivist perspective, is the knowledge enrichment framework (Spelke, 1991). According to this view, knowledge about the world is gradually built up over the course of a life; new concepts are added to existing concepts. This is seen as a limited kind of conceptual change (Carey & Spelke, 1996).

Other frameworks within the cognitive perspective include knowledge in pieces or facets of knowledge (diSessa, 1983; Minstrell, 1989), and misconceptions replacement (McCloskey,1983; Novak & Musonda, 1991). Within these frameworks, learning is accomplished through direct instruction that identifies students' misconceptions, points out where and why those conceptions are invalid, and exposes students to the scientifically normative conceptions.

*Behaviorist perspective.* In contrast to the cognitive perspective, the frameworks within the behaviorist perspective pose learning as primarily the acquisition of skills. For behaviorist researchers, who often have their intellectual roots in the work of Skinner (1958), learning consists of acquiring skills for doing science. This is typically accomplished by mastering a set of simple skills and then combining them together into more complex behavioral abilities. Learning experiences are organized from simpler to more complex tasks and provide distinct opportunities for structured practice with each skill (Bloom, 1976). Assessment of learning simply requires testing that a student can carry out those specific skills (Bloom, 1976).

*Sociocultural perspective*. Yet another broad category of frameworks that can inform research on science learning is the sociocultural perspective, which focuses on the interactions, discourse, and cultural tools of a group of science learners (Cole & Griffin, 1980; Jordan &

Henderson, 1995; Lemke, 1990; Rogoff, 1991). Often drawing upon the work of Vygotsky (1962; 1978), researchers in the sociocultural tradition see social interaction as necessary for learning; it is not merely supportive of or associated with learning. Likewise, they see knowledge as fundamentally social in origin and situated in interactions among people, rather than in the cognitive structures of individuals (Jordan & Henderson, 1995). In the extreme sociocultural view, science learning consists of moving more toward the core of the practicing community (Lave & Wenger, 1991).

*This study's perspective on learning.* For this study, I have adopted a theoretical framework that draws heavily from the sociocultural perspective but can also be seen as representing somewhat of a middle ground between the individual constructivist and the behaviorist perspectives. The framework I am using is the *social constructivist* view (Driver et al., 1994), where learning science involves both individual and social processes – both personal construction of knowledge about the world and social interactions supported by the tools and practices of the scientific community. This framework has also been termed the *situative* perspective on learning, which conceives of all learning as situated in activity (Greeno 1998; 2006). The situative, social constructivist perspective involves both cognitive and behaviorist aspects because it includes "both conceptual understanding and skill acquisition as valuable aspects of students' participation" (Greeno, 1998, p. 14).

Coming from this situative, social constructivist framework, I see science learning as comprised of both social enculturation into *practices* and personal construction of *ideas*. Social interaction exposes learners to the cultural tools of science, and individual reflection enables learners to apply those tools to personal meaning-making about natural phenomena. Becoming a more effective participant in a scientific community involves both constructing meanings for concepts and conducting scientific practices. The curriculum and assessment tools described in this paper attempt to relate the individual aspects of scientific understanding to its social aspects.

*This Study's Materials Science Curriculum*

My adoption of a situative, social constructivist framework for science learning has consequences for how I understand and present the introductory materials science curriculum used in this study. I suggest that students can effectively learn about object and material properties while socially engaged in solving problems via actual materials science practices. Primary level instruction about the properties of materials and objects should have students write with, talk with, and physically use the cultural tools (symbolic resources) of materials science.

Engineering design is one kind of authentic activity that requires the use of both materials science practices and materials science content knowledge[2]. This study's curriculum module, *Design a Model House: The Properties of Objects and Materials,* situates materials science learning within the authentic social activity of engineering. In the opening lesson of the module, students learn that their engineering design challenge is to create a miniature model house that is stable, soundproof, and thermally insulated. Over the next six lessons, students conduct a series of engineering tests to identify the materials that will enable their house to meet these requirements. They use digital sound and temperature sensors for several of their tests and LEGO™ construction kit elements for the frames of their houses. As students test materials and begin to construct their houses, they are asked to make scientific arguments about the best materials for each portion of the house. As they make choices, they are encouraged to consider

[2] At the introductory level, those practices include a) testing materials for particular properties, b) identifying the material properties important to a given task, c) selecting materials with those properties, d) describing materials specifically with intensive and extensive properties, and e) proposing combinations of materials that would accomplish a task. The content knowledge concerns a) the difference between an object and its constituent material and b) the macroscopic properties that are important for very common mechanical tasks.

the material and object properties of stability, strength, soundproofing, insulation, and reflectivity. In the module's two concluding lessons, students employ their new understanding of object and material properties to complete the design and building of their miniature model houses. The module is intended for third-grade (8- to 9-year-old) students and requires approximately 12 hours of instructional time for its nine lessons (see Appendix B).

In at least three ways, the *Design a Model House* curriculum module differs from existing instructional approaches that deal with introductory materials science. First, the module enables students to study materials science in the context of completing a design challenge (the creation of a miniature house to satisfy certain criteria and constrains), rather than in the context of exploring assorted material samples. Second, this module enables students to practice testing material properties with the end goal of selecting the best material for a task, rather than with the end goal of classifying materials. Finally, this study's curriculum module gives young students the chance to engage in sense-making about two complex material properties: *strength* and *thermal insulation*. These are important engineering properties, but due to their complexity they are often avoided in materials science instruction for children.

This curriculum's developers, including myself, took the stance that although the properties of strength and thermal insulation are conceptually challenging, children's everyday experiences do provide some resources for exploring them. Furthermore, I conjecture that the inclusion of these properties might actually support children's learning about the practices of materials science. My rationale for this conjecture is that these properties are more consistent with authentic materials science practice. Materials scientists and engineers often engage in the practices of testing and selecting for complex, non-perceptual properties. They attend especially to the *strength* and *thermal conductivity* of materials because these properties often determine a material's value as part of a mechanical system, for which load and thermal stressors are constant concerns. Thus, design tasks that deal with strength and insulation are more authentic to scientific practice, and thus more consistent with my perspective on learning, than are tasks that deal only with simple, perceptually accessible properties like color and weight.

The primary instructional goal of this study's curriculum module is to enable students to carry out introductory materials science. The curriculum developers reduced the list of introductory materials science practices and content (Table 1) to the three main learning objectives. These objectives state that students should be able to describe objects and materials by their extensive and intensive properties, identify the properties that are most important for accomplishing a specific design task, and conduct tests to select materials and objects that exhibit the properties desired for a specific design task. In summary, the curriculum was designed with a focus on children's practices of *describing*, *selecting*, and *testing* material and object properties.

Particular resources within the *Design a Model House* curriculum are intended to expose students to the practices of materials science and the cultural tools used to conduct those practices. These resources include:

a) physical tests demonstrated by the teacher and conducted by the students (compressive test of strength, tap test for stability, insulation test with heat stimulus and temperature sensor, sound absorption test with decibel sensor),
b) graphical representations in the students' science workbooks (charts for recording test results, designated areas for sketching plans to use materials),
c) scientific language spoken by the teacher and written in the workbook (material, property, test, pass, fail, result, select, strong, stable, insulating, absorbing), and
d) physical processes of using materials to build and re-build components of the model houses, as demonstrated by the teacher and conducted by the students.

I consider these resources to be the main *affordances* of this curriculum (Greeno, 1998). They enable children to pick up the discourse, practices, and representations – the cultural tools – of materials science. I acknowledge that the curriculum does not offer many affordances for making deep changes to one’s concepts of material phenomena. With this acknowledgment, I am not suggesting that conceptual change about material phenomena is impossible. Instead I am suggesting that the affordances of this particular curriculum lead mainly to changes in *practice*, rather than to changes in *concepts*. In other words, this particular curriculum does not provide many opportunities for deep cognitive conflict followed by reflection, thought experiments, and new ways of thinking about the puzzling observations, which many scholars say are needed for conceptual change to occur (Driver et al., 1994; Posner et al. 1982; Smith, 2007). I do concur with the many researchers from the cognitive tradition who hold that change in one’s concepts about particular physical phenomena is an important part of learning science. But because the *Design a Model House* curriculum does not afford many opportunities for deep conceptual change to occur, it is not at the heart of this study’s investigation into science learning.

*Existing Methods for Assessing Children’s Materials Science Learning*

Having explained my theoretical framework for science learning and provided an overview of this study’s learning experiences, I now turn to the challenge of assessing materials science learning. Current assessment tools related to introductory materials science can be roughly grouped into five categories: a) sorting material samples into groups, b) describing the composition of samples, c) describing samples’ properties and explaining why they have those properties, d) predicting whether a sample’s material kind will be preserved under transformations, and e) representing mental models of weight and density.

*Sorting samples.* When researchers ask students to sort material samples into groups, they typically follow up by asking the student to explain why he or she placed certain samples together (Krnel et al., 2003). In some cases they tell the students ahead of time how to sort the samples, such as by the "same kind of stuff" (Dickinson, 1987), or by a particular property like size, weight, or "heaviness for size" (Carey, 1992). These assessment tasks give researchers information about the characteristics to which children naturally attend, about their knowledge of particular properties, and about their ability to keep one variable consistently in mind as they classify. However, these tasks are limited by their specificity to the particular samples chosen and by the possible capriciousness of children's sorting methods. The fact that a child chooses to sort by color, for example, cannot be used as evidence that the child is *not* able to sort by weight.

*Describing composition.* In the second category of assessment tasks, researchers want to understand children's ideas about material composition, so they present samples to children and ask what they are "made of" (Smith, Carey, Wiser, 1985). Some researchers do this to elicit children's vocabulary for material kinds. Other researchers, curious if children hold a particulate view of the nature of matter, are looking for children to mention molecules, atoms, or other small particles (Nakhleh & Samarapungavan, 1999; Renstrom, Andersson, & Marton, 1990). One weakness of these tasks is that young children often interpret "made of" as meaning "constructed with" rather than as meaning "constituted of" (Dickinson, 1987).

*Explaining properties.* The third category of materials science assessment tasks includes describing material samples' properties and then explaining them. Because researchers typically do not offer measurement devices or tools for these tasks, they are investigating the student's ability to account for macroscopic perceptions in terms of pre-existing knowledge about microscopic or otherwise non-perceptible structure and processes (Nakhleh & Samarapungavan, 1999, Lee et al., 1993; Lewis & Linn, 1994). Explaining properties is an important ability for students to develop and for educators to assess, but these tasks may not elicit students’ full set of abilities because students might not see the questions as authentically problematic. Expert material scientists strive to explain material properties when they need to understand a surprising behavior or predict how a material will perform under new conditions. In contrast, students might not see any need to know why a sample has the properties that it has, and thus their responses to these assessment tasks may be underdeveloped compared to their actual abilities.

*Explaining transformations.* A large number of assessment tasks have been developed and conducted to study children’s thinking about the preservation of material kind under various transformations, including dividing, powderizing, melting, heating, and dissolving. For example, Smith, Carey, and Wiser (1985) present children sequentially with a paper cup, rubber balloon, wooden airplane, and metal spoon. For each object, they ask what the object is made of, cut it into smaller pieces, and ask if the smaller pieces are still made of the same stuff. Dickinson (1987) expands on this by showing powderized versions of the original objects, and Au (1994) goes one step farther by showing powdered substances, including sugar and corn starch, and asking if they change to a different substance when made invisible by dissolving in water. These kinds of assessment tasks are good at indicating whether children use the same material kind labels after changes have been made to physical samples. However, these tasks are not able to tell us whether children think that a transformed sample has the same properties and structure as in its pre-transformation condition. Children’s material kind labels may imply different sets of properties than do scientist’s material kind labels. For example, when a scientist labels something “plastic,” she is implying that it is a poor conductor, that it is synthetically produced, and so on, but when a child labels something “plastic,” she may be implying only that the object has the same surface appearance as a milk jug she knows is called plastic.

*Modeling weight and density.* A fifth category of materials science assessment tasks stems from the research on students’ concepts of weight and density. The tasks in this group focus on eliciting students’ mental models of these concepts. Researchers indirectly elicit these models by prompting students to create pairs and orders of objects based on their weight and density (Carey, 1992; Smith et al., 1997), and by asking questions about the weight and density of sets of objects, such as an aluminum block and an equal-sized steel block, or a tiny piece of Styrofoam and a large piece (Carey, 1992; Smith, 2007). More directly, researchers ask students to draw visual models of density, in both clinical interview settings and on written tests (Smith et al., 1997). Written tasks of calculating volume, weight, mass, and density also help researchers understand students’ models of these ideas (Smith et al., 1997; Smith, 2007). While these tasks and others like them provide much information about students’ concepts of weight and density, they cannot tell us whether or how students *use* weight and density measurements to make decisions about the appropriateness of one material over another for a particular task.

### *Previous Findings on Children’s Materials Science Learning*

Although the assessment tools discussed above are not able to tell us about children’s ability to select and test materials for particular tasks, they have identified numerous conceptual and practical challenges involved in developing a scientific perspective on the properties of materials and objects. Here I describe some of those challenges, with a focus on the concepts and practices that are especially relevant to the present study, including the properties of strength and thermal insulation and the phenomena of heat and temperature.

*The challenge of the particulate model.* Primary grade children are able to classify, describe, and explain objects and materials with rich and varied sets of ideas. Generally, however, children's ideas about objects and materials tend to focus on perceptual, macroscopic qualities without reference to particles as an explanatory model (Krnel, Glazar, & Watson, 2003; Stavy, 1991). Although understandings of matter vary greatly among children of even the same age (Wiser & Smith, 2008), many studies have shown how difficult it is, at any age, to develop an explanatory particulate model of matter (e.g., Lee et al., 1993; Novick & Nussbaum, 1974; Snir, Smith, & Raz, 2003; Wiser & Smith, 2008.)

*Challenges related to explaining properties.* When asked to explain the qualities of objects – to provide a mechanistic cause for a macroscopic property – primary grade children tend to invoke other macroscopic characteristics. For example, Nakhleh and Samarapungavan (1999) prompted children to give explanations for two observable material properties by asking, "Why does the wooden toothpick hold its shape but the water flows?" and "You can bend the copper but the toothpick breaks. Why do you think that happens?" They chose to investigate children's explanations of fluidity and malleability because they view these properties as easily observed and readily familiar to children. They found, however, that the children's explanations remained primarily at the macroscopic level and were often scientifically inaccurate. Most of the children's explanations for the difference between the water and toothpick (what these researchers call fluidity) and between the wire and toothpick (what these researchers call malleability) took the form of other intensive and extensive properties.

To explain why toothpicks hold their shape, the children said things like, "the wood is harder than water," and "the wood is heavier than water." To explain why copper wires bend but do not break, the children suggested that "wire is harder than wood," that "wood is stronger than metal," and that "wire is denser than wood." The children were more likely to explain samples' fluidity and malleability by discussing their hardness, weight, strength, or density, than by referring to their material composition or particulate structure (Nakhleh and Samarapungavan, 1999).

*Challenges related to strength.* In the present study, the assessment tools ask children to choose materials and objects with the properties of strength and insulation. More precisely identified as the engineering properties of *compressive yield strength* and *thermal conductivity*[3] (Schaffer et al., 1999), these properties are determined by a number of microscopic factors, including molecular composition, free electron count, crystal structure, porosity, and pore shape (Schaffer et al., 1999).

A small amount of research has been conducted on children's understanding of strength, including one study that used a paper-and-pencil test item to investigate upper elementary students' ideas about what makes a tower stable[4] (Gustafson, Rowell, & Rose, 1999). Many of the students' suggestions for designing a stable tower (a tower that would not "tip over") related to the weight of the tower's materials. Some students suggested adding weight to the tower's base, support posts, or top, and other students proposed increasing the weight of the entire tower. These ideas about structural stability may stem from an implicit belief that weight implies sturdiness. Students seem to think that a tower's tendency to tip over is dependent on some overall "sturdiness factor," which they believe to be determined by either the tower's overall weight or by the weight of some of its components (Gustafson et al., 1999).

This inclination to associate weight with design improvement may be related to children's general proclivity to attend to the property of weight. When asked to describe material samples (and given the opportunity to hold them in their hands), children show a preference for

[3] A material's insulation ability is closely related to the *inverse* of its thermal conductivity (Schaffer et al., 1999).
[4] Although structural stability is distinct from structural strength, the strength of the materials and joints in a structure can affect its stability.

referring spontaneously to weight (Krnel et al, 1998; Smith, Carey, & Wiser, 1985). When children in preschool or the early primary grades mention a sample's weight, however, they are typically referring to *felt* weight and are using an undifferentiated weight/density concept (Smith, 2007; Smith, Snir, & Grosslight, 1992).

In an investigation related to the stable tower study, researchers posed another paper-and-pencil item to elicit children's ideas about testing structural strength (Gustafson, Rowell, & Guilbert, 2000). Below a sketch of two bridges, the item asked students to predict which bridge would be the strongest and explain how they could find out if their prediction was correct. Many students misinterpreted the item. Rather than proposing a method to test their prediction, they simply explained their choice for strongest bridge. Their reasons included scientifically correct statements like, "diagonal elements are more supportive than vertical elements," as well as incorrect statements like, "that one bridge is larger than the other one."

When students did follow the instructions to propose tests of their predictions, they mostly responded in one of three ways. Some students wrote simply that the bridge should be built and tested. Other students described a specific action, such as adding weights or shaking the bridges, but did not discuss how to compare one bridge to the other. A third group of students suggested comparing the two bridges, but they did not specify that the conditions should be held constant for both bridges. Because few responses were both explicit about equivalent conditions and specific about a method for quantifying the bridge's strength (e.g., counting how many weights could be added before collapse), Gustafson et al. conclude that students need explicit instruction on strategies for comparing the strength of multiple products.

*Challenges related to thermal insulation.* Extensive research has been done on children's thinking about heat, temperature, and the related properties of thermal conductivity and insulation. Most assessments designed to explore students' understandings of heat and temperature include at least one item requiring students to differentiate between the ambient temperature and the perceived temperature of objects within that environment (e.g., Clark, 2006; Lewis & Linn, 1994; Paik, Cho, & Go, 2007). In this kind of item, typically the student is shown a set of objects that are made of different materials (often wood, metal, and glass) but exposed for the same extended amount of time to the same thermal environment, such as an oven, hot plate, or refrigerator. The student is then asked whether the objects will be at the same temperature or different temperatures. A scientifically normative response would indicate that perceived temperature is dependent on thermal conductivity. It would explain that all the objects will be at the same temperature, but they will feel different to your hand because they conduct heat at different rates. This is not the response made by most children. Instead, children's responses typically indicate that they think the temperature of a substance depends on its constituent material. They do not differentiate between a substance's speed of heat transfer and its temperature (Paik et al., 2007). They also think an object's capacity to increase in temperature depends on its size or thickness (Erickson, 1979; Clark, 2006).

One of the major conceptual challenges related to thermal phenomena is the distinction between the concepts of temperature and heat. Many preschool and early primary grade children conceive of heat as a gas or vapor (Paik et al., 2007), and they tend to conflate heat with its effects, such as burning and melting (Shayer & Wylam, 1981). Most older children conceive of heat as perceived hotness, which means that their concept of heat remains undifferentiated from their concept of temperature (Wiser & Amin, 2001). In contrast, the distinct scientific understandings of heat and temperature are that heat is exchanged energy, and temperature is average molecular kinetic energy (Wiser & Amin, 2001).

A second major conceptual challenge is the equivalence between insulation that maintains cold temperatures and insulation that maintains hot temperatures. Children instead believe that materials that can insulate "hotness" are fundamentally different from materials that can insulate "coldness" (Lewis & Linn, 1994; Paik et al., 2007). They tend to think that materials that feel cool to the touch at room temperature (such as metal and glass) are good at keeping objects cold, while materials that feel warm to the touch (such as wool and cotton) are good at keeping things hot (Lewis & Linn, 1994). For example, children would choose a wool blanket as the wrapper for a mug of hot soup, but they would deny the blanket's usefulness for keeping a soda cold and instead choose aluminum foil as a soda wrapper. Because children associate the term "insulation" with warmth, and because they have heard that "conductors" are the opposites of insulators, children tend to label warm-feeling materials as insulators and cool-feeling materials as conductors (Lewis & Linn, 1994).

When children are asked to explain their choices of materials for keeping objects hot or cold, three frequently-held intuitive beliefs often surface (Lewis & Linn, 1994). The first common belief, based on the experience of touching metal at room temperature and perceiving it to be cold, is that all metals are cold and able to attract, hold, and absorb cold by conducting heat more slowly than insulators. This idea leads to the choice of aluminum foil to keep a cold soda cold. The second frequently observed intuitive belief, perhaps based on the observation that wooden cooking utensils seldom feel hot, is that insulators work by conducting heat away to the environment very quickly. This belief is connected to the notion that good insulators are materials that give a strong sensation of emanating warmth. The third commonly-observed belief about thermal conductivity, based on experience using clothing to maintain body temperature, is the notion that textile-based insulators, like wool, are insulators because they generate their own heat. This belief leads children to deny that wool could be used to keep cold things cold; they think it would instead make the cold things warm.

*Thoughts on including strength and insulation in a third-grade curriculum.* Given that the concepts of thermal conductivity and strength present such substantial challenges for children, it is important to explain the rationale for their inclusion in this study. First, if we task children with testing and selecting only for simple perceptual properties like weight and color, we cannot make strong claims about the depth of their proficiency at materials science practices. There are many correct commonsense ways to test and select for simple properties. A student with a deeper understanding of material testing and selecting would be able to test and select for complex properties as well as for simple properties (even if she is not able to explain the mechanism for the complex properties). Thus, by asking children to test and select for complex properties, we should be able to better differentiate among levels of proficiency at these practices. Second, children's competence at sorting by size, shape, color, and texture suggest that they will also be able to test and select for these simple, perceptually-accessible properties. A more novel contribution can be made by asking students to test and select for non-perceptually accessible properties.

One conclusion that could be drawn from previous research is that we should not confront young children with tasks involving complicated, non-perceptual properties like strength and insulation. In other words, if we want children to learn the science practices of material testing and selection, we should have children test and select for simpler perceptually-based properties like weight and hardness. However, there is an alternative way to think about the research described above. Although most children find non-perceptual properties like strength and insulation to be scientifically puzzling, they also readily recognize them as important in daily life. Children sense the usefulness of knowledge about strength and insulation, even if they struggle to explain strength and insulation scientifically. Thus, these conceptually complicated properties are not necessarily alienating to children.

Because the *Design a Model House* curriculum affords resources for students to engage in the practice of materials science, I envision that the students will be able to make more sense of strength and insulation than previous research would predict. I do not think students will undergo deep conceptual change about what makes a material strong or insulating, but rather I expect them to be able to apply their new science practices to do some reasoning about a) when strength and insulation are important (i.e., when they should factor into materials selection) and b) how to test one material's strength and insulation ability against another's. Deep conceptual change about the explanations for material phenomena is not likely with this study's curriculum module. Nonetheless, because the science practices in this curriculum are conducted within a context of reasoning about materials, I do anticipate that some students will express new ideas about material and object properties.

### *This Study's Measures for Assessing Materials Science Learning*

This study's learning measures differ from those used in previous research because they are based on design tasks. More precisely, this study's measures are based on *materials selection* design tasks. These are tasks that probe students' ability to determine important material properties, predict which materials exhibit those properties and explain why, and propose tests for those properties. From a situative perspective, assessment as well as learning should be embedded in authentic practice. Thus, this study's approach to assessment is to record how students approach materials science design tasks, which involve the selection and testing of materials for a real purpose.

The pre-post assessment used by this study is a clinical interview that consists of two materials selection design tasks. In the first task, students choose the material for a collapse-proof stepstool, and in the second task, students choose the material for a warm pet habitat. The *Method* section describes these tasks in greater detail.

To explore how pre/post changes on the interview tasks might be connected to classroom instruction, this study also makes use of the students’ science workbooks as embedded data collection tools. These workbooks allow the practices exhibited during instructional time to be compared with the practices exhibited during assessment time.

To reiterate, this study was designed to investigate two main research questions:

a) How do children approach material selection design tasks in a clinical interview setting, before and after they have participated in a design-based science curriculum module on material and object properties?

b) How do the pre/post changes in children’s interview performance compare to their engagement with and performance on materials science practices throughout the curriculum module?

I expect that this study’s design task interviews will be easier than more traditional classification-based interviews in at least four ways. First, the contexts for the design tasks were chosen for their familiarity. Most children have experience with stepstools and animal habitats, and they are familiar with the problems of reaching higher places and of providing special living spaces for pets. Second, the material samples presented for each design task were also chosen for their familiarity, so that the children’s reasoning need not be complicated by having to guess the composition of the samples. Wood, plastic, foam, and metal are common in everyday artifacts.

Third, by offering a limited set of only two materials for each design task (wood/plastic and foam/metal), the interviews do not require the children to generate spontaneously their own list of materials. Instead, the children can focus their efforts on evaluating the presented materials. Fourth, the design tasks posed in the interviews are aligned to the design task posed in classroom instruction. Strength and insulation are important to the miniature model houses, and throughout the curriculum module, students have opportunities to write and draw in their workbooks about selecting, testing, and describing materials for those two properties.

I also expect that there are some ways in which this study's design task interviews, as well as its design-based workbooks, will be more difficult than traditional classification tasks. Some of the difficulties stem from information that is purposely not provided in the interview's design scenarios. In classification interviews, children are often told which properties to use as sorting criteria, but in this study's interview, the children are asked to select materials without being told which properties are most important for design success. The children are also asked to propose tests for materials without being told which tools and measurement devices would be available for conducting those tests.

Other difficulties stem from the decision to include thermal insulation as a focus property. Because, as research has shown, the concepts of heat and temperature are difficult for children, including the property of insulation in the instructional and assessment tasks makes both the workbooks and interviews cognitively demanding. However, this study offers the chance to explore what happens when children are asked to consider this complex, non-perceptual property during an authentic design task that really does call for it to be considered.

A third source of difficulty in this study's interviews and workbooks is the use of material samples that have some misleading extensive properties. The obvious properties of tubes and sheets made of plastic, wood, foam, and metal, such as their color, weight, and surface roughness, are not highly relevant to their success as materials for model houses and stepstools and pet habitats. These obvious but irrelevant properties might distract children from the intensive properties that are more important to design success.

Despite these difficulties, I expect the *Design a Model House* curriculum to enculturate students into the practices of material selection and testing, which are part of the discourse and cultural tools of materials science. Thus, I hypothesize that in students' postinstruction interviews, they will more often refer to intensive physical properties as they explain their material selection choices, and they will more often propose material tests that involve quantified measures and fair comparisons. Finally, I expect that the interview's practical design scenarios will support performance of the science practices that are most emphasized in the science workbooks throughout the curriculum module.

## Method

### *Setting*

This study took place in one single K-8 school in an urban public school district near a major northeastern city. The school has a student population of 465, and the student body is 15% African American, 10% Asian, 40% Latino, 33% White, and 2% multiracial non-Latino. The community served by the school has a large percentage of families with limited economic means, and a substantial number of families have recently immigrated to the United States. Of the student body, 79% are eligible for free or reduced-price lunch, 60% learned English as a second language, and 27% are designated as having limited English proficiency.

The school's district has been placing increased emphasis on the teaching of both science and engineering in the primary grades, and they have made several engineering curricula available for primary grade teachers to use in their classrooms. However, the school's performance on annual state science assessments remains at low levels. In the spring of 2008, 16% of fifth graders earned a "proficient" or higher rating on the science achievement test, which also includes engineering and technology items. Among all schools in the state, the average percentage of proficient or higher ratings on the fifth-grade science test was 50%.

In the summer of 2007, five third-grade teachers (one of whom eventually took part in the study reported in this paper) voluntarily participated in a 30-hour training workshop on two design-based science modules, including the *Design a Model House* module. Two of the teachers were from the school where this study's data were collected; the other three teachers were from two other schools. In the workshop, a team of university researchers modeled the role of design-based science teacher, and the participating teachers worked through each module as if they were students. The teachers completed every writing prompt, drawing task, and design-and-build challenge that students are asked to complete in the modules, and they also discussed reflection questions at the end of each activity. Each teacher received a stipend, professional development points, and a full classroom set of LEGO engineering materials, including a LEGO Mindstorms NXT kit for every pair of students. The teachers agreed to implement the modules in their classrooms and allow researchers to observe instruction and review students' work.

*Participants*

This study was conducted in a third-grade classroom led by a teacher who volunteered to attend the summer training workshop and implement the *Design a Model House* science module. She had no prior experience teaching engineering or using LEGO in her classroom, but she had previously taught about material properties during a science unit on rocks and minerals. This teacher was the primary implementer of the curriculum. As a graduate researcher representing the university, I was present for half of the instructional sessions, during which I provided only technical assistance with the LEGO electronics and construction materials. I also conducted the pre- and postinstruction interviews. The teacher initiated each lesson by explaining the model house criteria to be investigated that day, and then students worked independently for five minutes to respond to a brainstorming prompt related to those criteria. Then, for the majority of the lesson, students worked in pairs to complete the day's model house investigation. The teacher concluded each lesson by guiding the students through summary questions in their workbooks.

Eighteen students participated in the curriculum module in this classroom over a period of four weeks. Among them, two boys and seven girls who returned permission forms were included in this study. These 9 participants represented a broad range of academic abilities. On written science tests, 4 of the 9 participants typically earned scores in the bottom quartile, 2 earned scores in the second quartile, 2 earned scores in the third quartiles, and 1 earned scores in the top quartile. Thus, the 9 participants represented a varied sample of the entire class's academic science ability as measured by traditional written assessments.

Table 2
*Profiles of the 9 Student Participants*

| Name | Sex | Typical Written Science Test Achievement | Currently Designated as English Language Learner? |
|---|---|---|---|
| Ava | Female | Middle (Third Quartile) | No |
| Brian | Male | Middle (Third Quartile) | No |
| Chinelle | Female | Low (Fourth Quartile) | No |
| Elisa | Female | Low (Fourth Quartile) | No |
| Julia | Female | Low (Fourth Quartile) | Yes |
| Katie | Female | Middle (Second Quartile) | No |
| Macon | Male | Middle (Second Quartile) | No |
| Stella | Female | Low (Fourth Quartile) | No |
| Viola | Female | High (First Quartile) | No |

### *Data Sources*

The data sources for this study were clinical interviews and student workbooks. The interviews were conducted at pre- and postinstruction, and the workbooks were used throughout the curriculum module. These data sources were used to measure three aspects of students' approaches to introductory materials science: a) describing objects and materials, b) testing and measuring objects and materials, and c) selecting objects and materials.

*Constructs measured by the data sources.* For this study, children's understanding of the macroscopic properties of materials and objects was treated as a latent construct (Wilson, 2005) that can be partially tapped by observing three distinct practices: describing, testing, and selecting objects and materials (see Table 3). These practices are drawn from the set of practices and content knowledge important to introductory materials science (shown in Table 1), and they are consistent with the learning objectives of the *Design a Model House* curriculum. Both the interview and workbook data sources were designed to measure these practices.

Table 3

*Construct Analysis with Examples of Introductory Materials Science Practices*

| | Science Practices Related to Material and Object Properties | | |
|---|---|---|---|
| STUDENTS:<br>The student has… | SELECTING:<br>To select a material for a specific task, the response… | TESTING:<br>To discuss how to test for a material property, the response… | DESCRIBING:<br>To describe a specific material sample, the response… |
| Highly proficient abilities related to material and object properties | Names appropriate materials and provides scientifically accurate rationale.<br>*e.g., I choose foam for the bath toy because it has very low density and is buoyant in water.* | Proposes and justifies an effective method that does not require full prototyping.<br>*e.g., Test for the strength of that plastic by measuring its compression distance under a 1-kg weight.* | Accurately uses multiple intensive and extensive properties.<br>*e.g., That cube of wood is lightweight and about 2 inches long on each side. It is brown and buoyant in water.* |
| ↑ | Names somewhat appropriate materials and provides partial rationale.<br>*e.g., I choose clay for the bath toy because I've made a clay boat before.* | Proposes and justifies an effective method, but it requires full prototyping.<br>*e.g., Test for the strength of that plastic by building a bed out of it and seeing if the bed collapses.* | Primarily describes surface features and extensive properties.<br>*e.g., That cube of wood has rough spots on it. It could give splinters. It feels heavy.* |
| Very limited abilities related to material and object properties | Provides no information relevant to selecting materials for the task.<br>*e.g., I would choose wood for the bath toy because it's the best choice.* | Provides no information relevant to measuring the property.<br>*e.g., I know that plastic has a lot of strength.* | Provides no information relevant to the material.<br>*e.g., I like how that cube of wood feels.* |

The practice of *describing* requires the ability to specify the characteristics of objects and materials with accuracy, clarity, and detail (Minogue, 2008; Smith et al., 2006). Suppose that a child has been given a cube-shaped sample of wood and has been asked to describe it in as much detail as possible. Proficient responses would accurately include entity-relevant, extensive properties as well as substance-relevant, intensive properties that refer to the invariant qualities of its constituent material (Au, 1994; Schwartz, 1996).

The practice of *testing* requires the ability to determine specific properties of objects and materials (Dowling, 1998; NRC, 1996; Smith et al., 2006). Suppose that the child mentioned above has now been asked to determine whether the wood is strong enough to be used for a bed. A highly proficient response would propose an effective but parsimonious method for determining strength and would clearly explain how to carry out this method (Smith et al., 2006).

The practice of *selecting* objects and materials requires the ability to choose appropriate materials and objects for a specific task (AAAS, 1993; Massachusetts Department of Education, 2001; NRC, 1996). Now suppose that the child has been asked to select a good material for a floating bath toy. Proficient responses would name an appropriately buoyant material and justify the choice by accurately describing properties that are relevant to flotation and children's play.

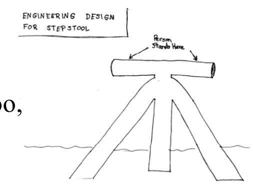
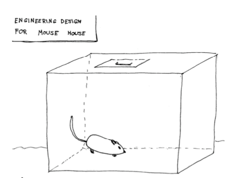
It is important to note that proficiency at one of these practices does not require or imply proficiency at the other two practices. For example, some children can select appropriate materials for a task, but they cannot describe properties of those materials or propose methods for testing those materials' properties (Krnel et al., 2003). We might also expect that some children can describe materials according to intensive and extensive properties without knowing which properties are important to a given design task. Still other children might be able to test for properties without being able to apply the test results to material selection. Different strategies for science instruction may influence these three practices in different ways. Accordingly, this study's instruments were designed to measure each practice separately. In both the interviews and workbooks, children were given a distinct opportunity to engage in each practice.

*Clinical interviews.* A semi-structured (Merriam, 1998) clinical interview (Brizuela, 2004; Ginsburg, 1997; Piaget, 1929) was used to answer this study's first research question: how do children approach the task of selecting objects and materials for design goals, and do their approaches change after design-based science instruction? In a clinical interview, each participant responds to the predetermined same set of prompts, but the interviewer probes more deeply when it is difficult to make sense of a participant's language or reasoning. The interview is "clinically" tailored to each individual participant. This flexibility makes the clinical interview a useful tool for in-depth exploration of a child's ideas about a science topic.

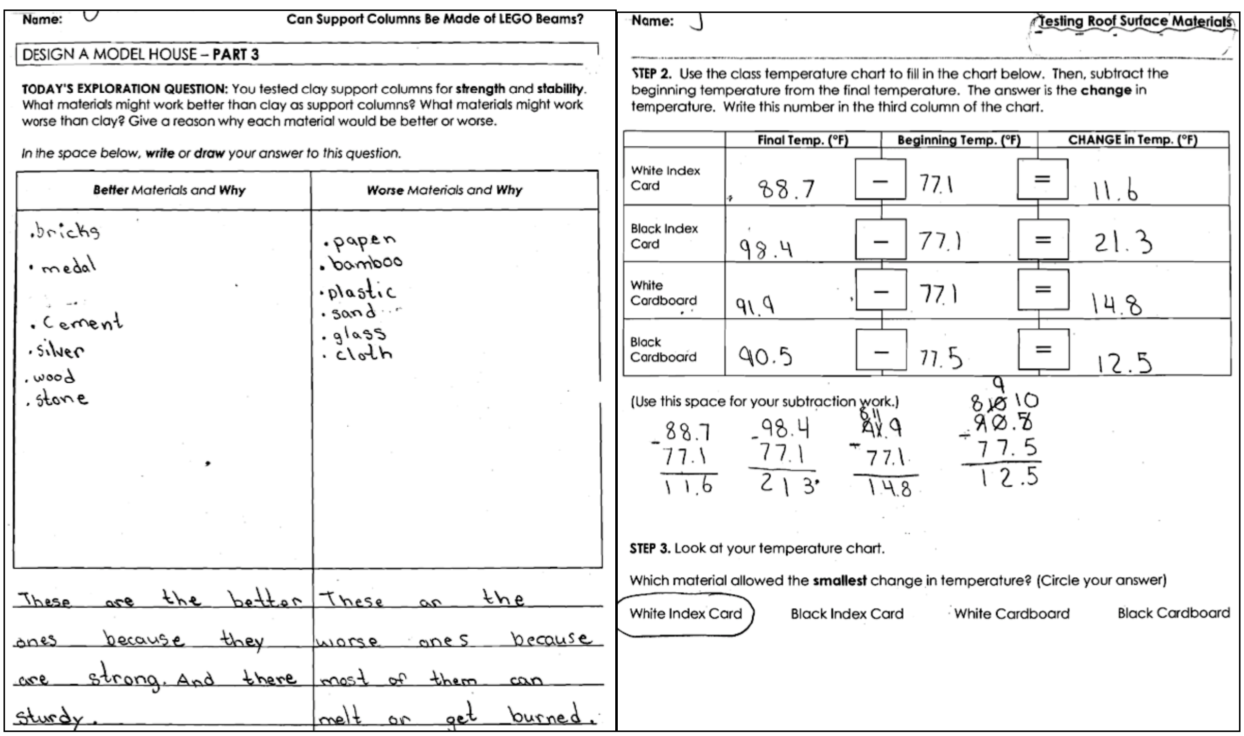
The interview used in this study consisted of two material selection design tasks. To begin each task, the interviewer stated a design goal, displayed a design schematic, and gave the student two material samples. The interviewer then asked the student to select, propose testing techniques for, and describe the material that would best accomplish the goal. For the first design task, students selected either bamboo pole or polyvinyl chloride (PVC) plastic pipe for the legs of a sturdy stepstool. In addition to providing an opportunity for students to select, test, and describe materials, the stepstool task was designed to elicit students' ideas about the property of strength. For the second task, students selected either aluminum sheeting or foam sheeting for the walls of an insulated pet habitat. This habitat task was designed to focus students' thinking on the property of insulation. Table 4 outlines the topics and practices assessed by the interview protocol. Both strength and insulation were properties addressed during instruction (both were required characteristics of the model houses), but students had more preinstruction everyday experience with strength. This allowed for their ideas about a familiar property to be compared with their ideas about a less familiar property.

The interviewer presented both tasks as parts of engineering design projects, and each interview lasted approximately twenty minutes. The preinstruction interviews took place the week before the students began working on the curriculum module in class. The postinstruction interviews took place over the two-week period after the students completed their model houses. The pre- and postinstruction protocols were identical.

Table 4
*Summary of Topics and Practices Assessed by Clinical Interviews*

| Topic | Practice | Scenario and props presented to students | Examples of additional spoken prompts |
|---|---|---|---|
| **Strength** (Compressive yield strength) | **Selecting** materials and objects for a particular task | Your job is to pick the material for this stepstool design. The stepstool shouldn't collapse when someone steps on it. Imagine you have to choose between these two materials:<br>• 12" length of PVC, 1.5" in diameter<br>• 12" length of bamboo, 1.5" in diameter<br> | Which material would you **pick** as the better material for the stepstool?<br><br>Why do you think this is better for the stepstool? |
| | **Testing** for a particular property | Now the other engineers want to know exactly how **strong** these two materials were. What does it mean for something to be strong? (Provide definition if student does not know.) | Tell me how you would **measure** how **strong** the materials are.<br><br>And are there any other **tests** you would do to figure out how strong the materials are? |
| | **Describing** objects and materials | Now the other engineers need to find the material that you picked in a big supply room. | **Describe** the strong material to me with as many details as you can think of. |
| **Insulation** (Inverse of thermal conductivity) | **Selecting** materials and objects for a particular task | Now imagine you have a new engineering job. Your job is to pick the material for the walls of this mouse habitat. The walls should keep the inside of the habitat warm. Imagine you have to choose between these two materials.<br>• 6"x9" Styrofoam tray<br>• 6"x9" aluminum tray<br> | Which material do you think you would **pick** as the better material for the mouse house walls?<br><br>Why do you think this is better for the walls? |
| | **Testing** for a particular property | Now imagine the other engineers want to know exactly how **insulating** these two materials are. What does it mean for something to be insulating?(Provide definition if student does not know.) | Tell me how you would **measure** how insulating the materials are.<br><br>And are there any other **tests** you would do to figure out how insulating the materials are? |
| | **Describing** objects and materials | Now the other engineers need to find the material that you picked in a big supply room. | **Describe** this material to me with as many details as you can think of. |

*Workbooks.* Student workbooks provided data for this study's second research question: how do the pre/post changes in children's interview performance compare to their engagement with and performance on materials science practices throughout the curriculum module? Used throughout the entire *Design a Model House* curriculum module, these workbooks were highly structured paper-and-pencil booklets that provided introductory questions, investigation instructions, and data recording prompts for each of the nine instructional activities of the module. The 21-page workbook included 14 pages with writing and drawing tasks related to the properties of strength and insulation. Six pages focused on the property of strength. Five pages focused on the property of insulation. Three pages prompted students to consider multiple properties simultaneously. Because strength and insulation were the two properties addressed in the clinical interview, researchers reviewed students' work only for these 14 pages. Figure 1 shows two page samples.


Name:

**Can Support Columns Be Made of LEGO Beams?**

DESIGN A MODEL HOUSE – **PART 3**

**TODAY'S EXPLORATION QUESTION:** You tested clay support columns for **strength** and **stability**. What materials might work better than clay as support columns? What materials might work worse than clay? Give a reason why each material would be better or worse.

*In the space below,* ***write*** *or* ***draw*** *your answer to this question.*

| **Better** Materials and **Why** | **Worse** Materials and **Why** |
| --- | --- |
| .bricks<br>• medal<br>• Cement<br>• silver<br>• wood<br>• stone | • papen<br>• bamboo<br>• plastic<br>• sand<br>• glass<br>• cloth |
| These are the better ones because they are strong. And there sturdy. | These ar the worse ones because most of them can melt or get burned. |

Name:

**Testing Roof Surface Materials**

**STEP 2.** Use the class temperature chart to fill in the chart below. Then, subtract the beginning temperature from the final temperature. The answer is the **change** in temperature. Write this number in the third column of the chart.

| | Final Temp. (°F) | | Beginning Temp. (°F) | | CHANGE in Temp. (°F) |
| --- | --- | --- | --- | --- | --- |
| White Index Card | 88.7 | – | 77.1 | = | 11.6 |
| Black Index Card | 98.4 | – | 77.1 | = | 21.3 |
| White Cardboard | 91.9 | – | 77.1 | = | 14.8 |
| Black Cardboard | 90.5 | – | 77.5 | = | 12.5 |

(Use this space for your subtraction work.)

88.7 − 77.1 = 11.6  98.4 − 77.1 = 21.3  91.9 − 77.1 = 14.8  90.8 − 77.5 = 12.5

**STEP 3.** Look at your temperature chart.

Which material allowed the **smallest** change in temperature? (Circle your answer)

White Index Card  Black Index Card  White Cardboard  Black Cardboard

*Figure 1.* Sample of workbook pages: The page on the left is related to the property of strength; the page on the right is related to the property of insulation.

*Data Reduction and Analysis*

The data reduction and analysis processes were designed to measure students' performance on three science practices: a) selecting and reasoning about objects and materials, b) testing the properties of objects and materials, and c) describing objects and materials. In scoring a workbook, raters examined students' writing and drawing to find indicators of proficient engagement in these practices. In scoring an interview, raters examined transcripts of students' talk rather than students' writing and drawing, but they were again seeking indicators of proficiency at selecting, testing, and describing objects and materials.

*Interview scoring.* All nine preinstruction and nine postinstruction interviews were transcribed in their entirety. Each transcript was then separated into six segments: the *select, test*, and *describe* sub-tasks within the strong stepstool design task, and the *select, test,* and *describe* sub-tasks within the insulated habitat design task. Each of the six segments was assigned a set of qualitative, descriptive codes (Glaser & Strauss, 1967; Miles & Huberman, 1994) as well as a numerical score, as described below and in Table 5. All interview responses were coded and scored by two independent raters. The interrater reliability for the numerical scores was 0.87 (Pearson correlation coefficient). Appendix C includes the complete scoring rubric for the interview responses, and Table 5 and the following paragraphs briefly discuss the coding and scoring of each sub-task.

For the *select* sub-tasks within both the stepstool task and the habitat task, students were prompted to state their rationale for choosing a certain material. Each rationale was coded for three factors: a) its relevance to the design goal (Fortus et al., 2005), b) its accuracy for the selected material sample (Minogue, 2008), and c) its type (whether the student cited an intensive property [Schwartz, 1996], extensive property [Schwartz, 1996], material kind [Dickinson, 1987], or irrelevant concern as the reason for choosing a material). Responses earned 2 points for each relevant, accurate rationale and 1 point for each relevant but inaccurate rationale. The maximum score for each select sub-task was 8.

Table 5
*Generalized Version of the Scoring Rubric for Interview Responses*

| **Sub-Task** | **Score Range** | **Scoring Rubric in Brief** | **Coding Scheme in Brief** |
|---|---|---|---|
| **Selecting** | 0 to 8 | 2 pt for each rationale that is both relevant to the design goal and accurate for the selected material sample<br>1 pt for each rationale that is relevant to the design goal but not accurate for the selected material sample<br>0 pt for rationales that inaccurately characterize the material or that are irrelevant to design goal | Each rationale coded for 3 factors:<br>• *Relevant* or *Irrelevant* (to design goal)<br>• *Accurate* or *Inaccurate* (for material sample)<br>• *Intensive*, *Extensive*, *Material Kind*, *Experience*, or *Irrelevant* (for rationale type) |
| **Testing** | 0 to 8 | 2 pt for each proposed material sample test that is relevant to design goal and that would provide evidence central to the desired property<br>1 pt for each proposed test that is relevant and central, but requires full prototyping of design product<br>0 pt for tests that would not produce central evidence or that are not relevant to design goal | Each proposed test coded for 3 factors:<br>• *Relevant* or *Irrelevant* (to specified property)<br>• *Central* or *Non-Central* (to determining property)<br>• *Prototype* or *No Prototype* (indicating whether building is required) |
| **Describing** | 0 to 2 | 1 pt for first accurate intensive property<br>1 pt for first accurate extensive property<br>0 pt for identifying surface features (e.g., dents, scratches) | Each descriptor coded as: *Intensive Property, Extensive Property, Surface Appearance, Material Kind, Function, Simile,* or *Inaccurate* |

For the *test* sub-tasks, students were prompted to propose techniques for measuring or testing a given property (strength for the stepstool, insulation for the habitat). Each proposed test or measurement technique was coded for three factors: a) its relevancy for the design goal (Fortus et al., 2005), b) its centrality to determining the property (Smith et al., 2006), and c) whether or not it required full prototyping of the design product, such as building an entire stepstool or pet habitat (Dym & Little, 2004). Responses earned 2 points for each relevant, productive test that did not require prototyping, and 1 point for proposing to build and test a full prototype. The maximum score for each measure sub-task was 8.

For the *describe* sub-tasks, students were prompted to describe a material sample with as many details as possible. Each descriptor was coded into one of the following categories: intensive property (Schwartz, 1996), extensive property (Schwartz, 1996), surface appearance (Smith, Carey, & Wiser, 1985), material kind (Dickinson, 1987), function (Dickinson, 1987), simile or metaphor (e.g., it looks like a log), or inaccurate. Responses earned 1 point for the first intensive property descriptor and 1 point for the first extensive property descriptor. The maximum score for each describe sub-task was 2.

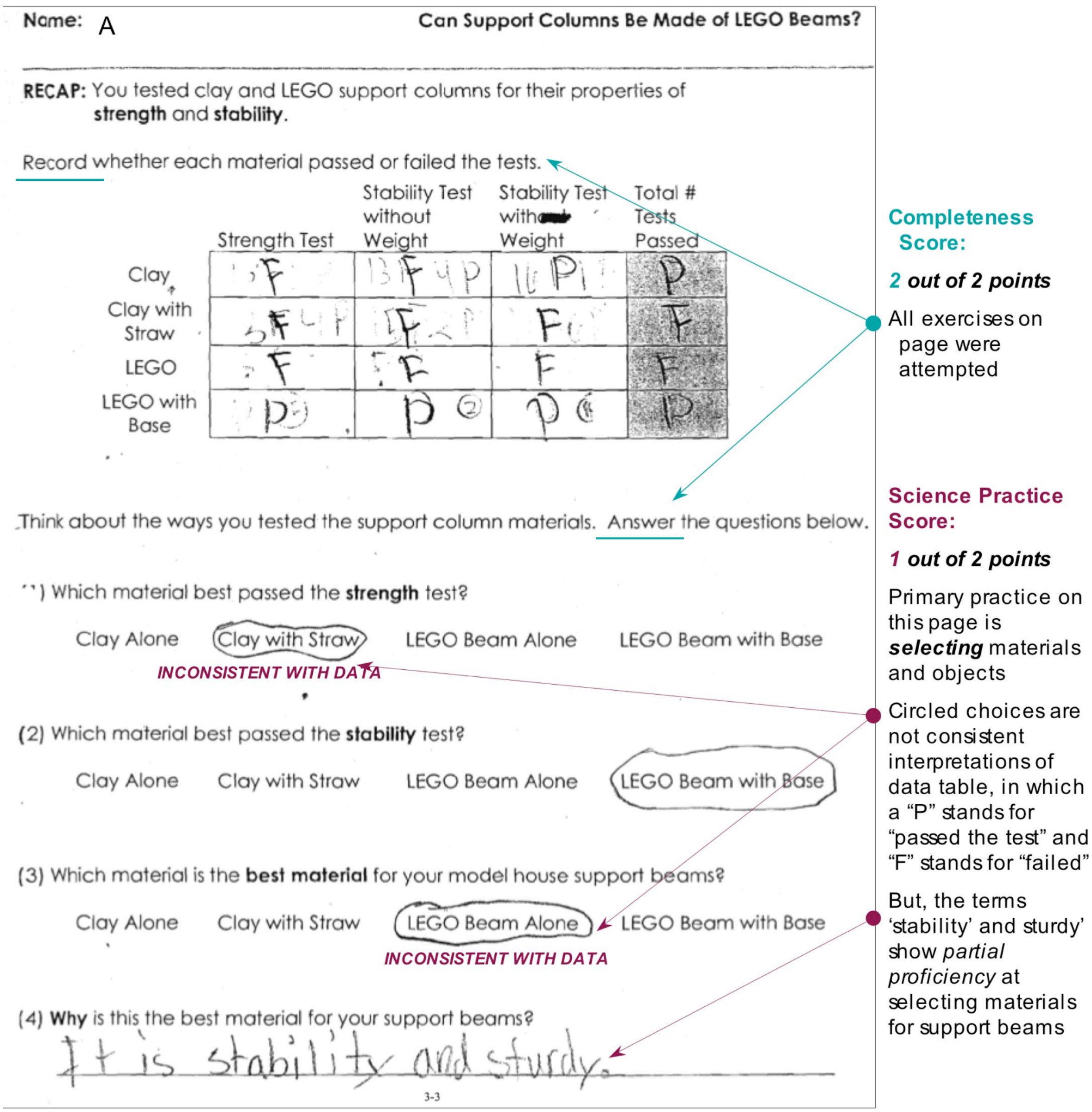
Assigning a numerical score of 0, 1, or 2 to each rationale, test, or description was useful for tracking progress at the specific materials science practices of selecting, testing, and describing. Consistent with the construct map shown in Table 3, where the upward direction shows increasing abilities at materials science, an upward shift in an individual student's scores implied an improvement in that student's materials science proficiency. By scoring the responses to each sub-task separately, I could recognize improvement at one practice even if performance on the other two practices did not improve. Numerical scoring also allowed me to compare performance on the stepstool design task with performance on the habitat design task, to determine the extent to which a particular property (strength or insulation) confounded students' materials science practices. The numerical scoring also enabled me to compare students' interview performances with their workbook scores.

To statistically compare scores from preinstruction interviews with scores from postinstruction interviews, nonparametric Wilcoxon signed-rank tests were conducted. Nonparametric statistical inference tests were used because the small sample size ($n = 9$) and non-normal distribution of scores violated the assumptions of the parametric paired sample t-test. The signed-rank test is a nonparametric test of significant difference between paired samples of an interval measure.

*Workbook scoring.* Each of the 14 workbook pages on strength and insulation was assigned two different scores. The first score for each page was the *activity completeness score*, which measured how fully the student attempted the tasks on the page (Baxter, Bass, Glaser, 2000). This score is a measure of behavioral engagement, which includes the student's response to teacher instructions (Fredricks, Blumenfeld, & Paris, 2004). A response earned 0 points for completeness if less than one third of the page had been completed, 1 point if between one third and two thirds of the page had been completed, and 2 points if two thirds or more of the page had been completed. For example, a page with three open-ended questions would earn 2 points if the student had answered two or three questions, 1 point if the student had answered one question, and 0 points if the student had answered no questions.

The second score for each workbook page was the *science practice score*, which measured proficiency at the key science practice featured in the page's tasks. Each page focused on one of three practices: a) reasoning about the relationship between materials' and objects' characteristics and their suitability for a design goal (*selecting materials*), b) representing experimental procedures and results (*testing materials*), and c) interpreting experimental data to make a claim about material characteristics (*describing materials*). A workbook response earned 0 points for no proficiency at the science practice, 1 point for partial proficiency, and 2 points for full proficiency at the science practice. For each workbook page, a specific scoring rubric stated the page's key science practice and described what could be considered partial or full proficiency at that practice (see Appendix D).

Two raters scored each workbook response for both completeness and science practice. Their interrater reliability for the scores was 0.86 (Pearson correlation; exact match proportion was 0.85). Taken together, the completeness score and science practice score represent a student's level of performance on a workbook page. Figure 2 illustrates the scoring of a typical workbook page.



Name: A

**Can Support Columns Be Made of LEGO Beams?**

**RECAP:** You tested clay and LEGO support columns for their properties of **strength** and **stability**.

Record whether each material passed or failed the tests.

| | Strength Test | Stability Test without Weight | Stability Test with[redacted] Weight | Total # Tests Passed |
|---|---|---|---|---|
| Clay | F | F | P | P |
| Clay with Straw | F | F | F | F |
| LEGO | F | F | F | F |
| LEGO with Base | P | P | P | P |

Think about the ways you tested the support column materials. Answer the questions below.

(1) Which material best passed the **strength** test?

Clay Alone  Clay with Straw  LEGO Beam Alone  LEGO Beam with Base

INCONSISTENT WITH DATA

(2) Which material best passed the **stability** test?

Clay Alone  Clay with Straw  LEGO Beam Alone  LEGO Beam with Base

(3) Which material is the **best material** for your model house support beams?

Clay Alone  Clay with Straw  LEGO Beam Alone  LEGO Beam with Base

INCONSISTENT WITH DATA

(4) **Why** is this the best material for your support beams?

It is stability and sturdy.

3-3



*Figure 2.* Sample scoring of workbook page 3-3, a page featuring the property of strength and the science practice of selecting materials and objects. This page, out of Ava's workbook, was scored a 2 for completeness and a 1 for science practice proficiency.

# Results

Interviews and workbooks were analyzed for evidence related to two main questions: a) How did the children approach the material selection design tasks in the interviews, before and after participating in the curriculum module? b) How did the changes in interview performance compare to the children's performance on materials science practices throughout the module? Extended examples from the work of four prototypical students reveal how students' performance on the workbook activities related to their interview responses. However, before describing those examples, I will summarize the main findings that stem from analyzing the interview responses as a stand-alone dataset. These findings include positive patterns of difference from the preinstruction to the postinstruction interviews. After reviewing the interview analyses, I will present correlations among the interviews and workbooks, and then I will use both data sources to describe four examples of student trajectories.

## *Design Task Approaches in Clinical Interviews*

By prompting students to talk through the process of selecting materials for design tasks, the clinical interviews showed that these third-grade children were able to demonstrate and articulate a rich spectrum of practices and ideas related to the properties of objects and materials. However, children's responses during the interviews' two design tasks were quite distinct. Their approaches to the stepstool task were more scientifically normative but less diverse than their approaches to the habitat task. When responding to the habitat task, the children often strayed from the stated design goal, introduced their own goals, and described properties and tests that would meet these more personally meaningful goals.

*Material selection rationale.* Each design task began by prompting students to select the better material for a design goal and explain the rationale for their selection. Table 6 shows that for this *select* sub-task, the insulated habitat goal elicited fewer relevant and accurate responses than did the strong stepstool goal. That is, for the habitat task, the students offered fewer statements that were both relevant to the design goal and accurate for the material they had selected. However, the students offered a greater total number of statements for habitat material selections than they did for stepstool material selections. This contrast occurred in both the preinstruction and postinstruction interviews.

Brandon's material selection reasoning at preinstruction illustrate the difference between the two design tasks. Brandon stated only two reasons for choosing  PVC for the stepstool, but both reasons were directly related to the goal of a collapse-proof stepstool: PVC is "strong" and "sturdy." In contrast, when Brandon chose metal for the pet habitat, he gave three reasons, and none was directly related to the goal of an insulated habitat wall. He explained that he chose metal because it is "hard," "heavy," and "the mouse can't bite through it."

Table 6
*Material Selection Rationales Offered by All Students*

| | Rationales for Material Selection | | List of All Rationales | |
|---|---|---|---|---|
| | No. of Distinct Rationales | Percentage Relevant and Accurate | Relevant and Accurate | *Irrelevant or Inaccurate* |
| Stepstool Task | | | | |
| Pre-instruction | 8 | 3/8 (38%) | • Strong<br>• Sturdy<br>• Holds a lot of weight | • *Heavy*<br>• *Stable*<br>• *Metal*<br>• *Wooden*<br>• *Paintable* |
| Post-instruction | 12 | 5/12 (42%) | • Strong<br>• Sturdy<br>• Holds a lot of weight<br>• Hard<br>• Tear-resistant | • *Heavy*<br>• *Stable*<br>• *Plastic*<br>• *Splinter-free*<br>• *Layered*<br>• *Attachable*<br>• *Thick* |
| Habitat Task | | | | |
| Pre-instruction | 10 | 2/10 (20%) | • Retains heat<br>• Blocks coldness | • *Heavy*<br>• *Strong*<br>• *Unbreakable*<br>• *Hard*<br>• *Thick*<br>• *Cut-able*<br>• *Lightweight*<br>• *Feels warm* |
| Post-instruction | 15 | 2/15 (13%) | • Retains heat<br>• Proven as insulator | • *Heavy*<br>• *Strong*<br>• *Unbreakable*<br>• *Hard*<br>• *Thick*<br>• *Cut-able*<br>• *Feels warm*<br>• *Heats up quickly*<br>• *Shiny*<br>• *Metal*<br>• *Soft*<br>• *Melt-proof*<br>• *Sound-proof* |

**Note:** Rationale statements made by multiple students are counted and listed only once, but this table includes the full list of these distinct rationale statements. Synonymous statements have been collapsed into one entry; e.g., the statements "unbreakable," "not breakable," and "can't be broken easily" are collapsed into the entry "unbreakable."

Table 7

*Material Property Tests Proposed by the Overall Group of Students*

| | Tests of Material Properties | | List of All Proposed Tests | |
|---|---|---|---|---|
| | Number of Distinct Proposals | Percentage Relevant to Goal and Central to Property | Relevant and Central | *Irrelevant or Non-central* |
| Stepstool Task | | | | |
| Pre-instruction | 8 | (7/8) 88% | • Build prototype and step on it<br>• Build prototype and put rocks on it<br>• Drop test<br>• Bend until break<br>• Tap until break<br>• Impact with rock<br>• Apply body weight | • *Weigh on scale* |
| Post-instruction | 12 | (10/12) 83% | • Build prototype and step on it<br>• Build prototype and do compression test<br>• Drop test<br>• Squeeze test<br>• Impact with rock<br>• Apply body weight<br>• Apply test weight<br>• Impact with hammer<br>• Impact with golf ball<br>• Flick test | • *Weigh on scale*<br>• *Measure thickness* |
| Habitat Task | | | | |
| Pre-instruction | 11 | (2/11) 18% | • Build and measure air temp over time<br>• build prototype and test with mouse | • *Expose to heat and measure temp [looking for warmth]*<br>• *Build and check for breathing holes*<br>• *Measure height*<br>• *Measure thickness*<br>• *Measure length*<br>• *Weigh on scale*<br>• *Check for melting*<br>• *Impact with hammer*<br>• *Try to poke holes* |
| Post-instruction | 16 | (4/16) 25% | • Build and measure air temp over time<br>• Build and test with mouse<br>• Expose to light and measure temp on each side over time<br>• Form into cylinder and expose to light and measure temp | • *Feel for warmth*<br>• *Expose to heat and measure temp [looking for warmth]*<br>• *Put on overhead projector*<br>• *Build and check for breathing holes*<br>• *Weigh on scale*<br>• *Try to break*<br>• *Drop test*<br>• *Impact with hammer*<br>• *Apply test weight*<br>• *Expose to light and check for light transmission*<br>• *Push test*<br>• *Expose to light and feel for warmth*<br>• *Add door to house*<br>• *Add tape to house* |

**Note:** Tests proposed by multiple students are counted and listed only once, this table includes the full list of these distinct proposed tests. Synonymous tests have been collapsed into one entry; e.g., the proposals to "add a test weight," "put weight on top," and "see if it can hold a weight," are collapsed into the entry "apply test weight."

*Proposals for testing material properties.* Table 7 illustrates that for the *test* sub-task, the insulated habitat task elicited fewer proposals that were relevant to the design goal and central to the desired property than did the strong stepstool task. The students more frequently proposed scientifically normative methods for determining strength than they did for determining insulation. However, the students offered a greater number of distinct ideas for testing insulation than they did for testing strength. These differences appeared during both the preinstruction and postinstruction interviews.

The tests proposed by Julia in her postinstruction interview exhibit these distinctions between the two design tasks. Within the stepstool task, Julia suggested three appropriate, useful tests for determining PVC's strength: "use a hammer to bang it and see if it doesn't break," "put a rock on it and see if it breaks," and "try dropping it on the floor." In contrast, within the habitat task, Julia's list of tests was more extensive, but fewer of the tests were relevant and central to the property she had been asked to determine, insulation. She said that to test for insulation, she would "touch [the metal] and see if it feels hot," "use a rock or a hammer," "put something heavy in the middle and see if it breaks," "try to bend it," "build a house and try it," and "make a little door to let air in and out."

Tables 6 and 7 are intended to highlight disparities between the strength and insulation tasks, rather than to display pre/post change. In these tables, any response given by multiple students is listed as only one distinct response. The numbers in the tables were calculated by counting the number of distinct responses and by computing the percentage that were "correct" (i.e., relevant and accurate for the material being selected, or relevant and central to the property being tested). Because Tables 6 and 7 count each correct response only once, they do not indicate whether individual or average student performance improved. For example, the correct response that the stepstool material should be “strong” was given by 3 students at preinstruction and by 6 students at postinstruction, but “strong” is listed just once in both the preinstruction and postinstruction rows of the table. Nonetheless, the slight pre/post increases in Tables 6 and 7 show that although they maintained a number of incorrect ideas, the students as a group did generate new correct responses at postinstruction.

*Descriptions of materials and objects.* Unlike the *select* and *test* portions of the interviews, the *describe* sub-task elicited similarly accurate ideas during both the stepstool task and the habitat task. Students were prompted to describe each material sample (PVC and bamboo in the stepstool task; aluminum and foam in the habitat task) with as many details as possible. As a group, the students generated between 25 and 35 distinct descriptors for each of the two stepstool materials and for each of the two habitat materials, at both preinstruction and postinstruction.

At preinstruction, 34 of the total 116 (29%) distinct material descriptors were intensive properties, like “weak,” “white,” and “hard,” and 36 (31%) were extensive properties, like “skinny,” “rectangular,” and “heavy.” Only two of the preinstruction descriptors were inaccurate: Katie described the foam sheeting by saying “warm air goes through it,” and “it won’t burn down.” The postinstruction counts and percentages were similar, with 29 of the 114 (25%) distinct descriptors as intensive properties, and 34 (30%) as extensive properties. Only one of the postinstruction descriptors was inaccurate. Again one of Katie’s responses, it was a description of metal sheeting as being “easy to melt.” (See Appendix E for a list of all descriptors generated by students.)

*Changes from Preinstruction to Postinstruction*

On average, the students’ postinstruction interview responses earned significantly higher scores than their preinstruction responses ($p < .05$, $n = 9$, 2-tailed Wilcoxon test).[5] There are three main ways to parse students’ performance during the clinical interviews: a) their total scores, b) their scores for each of the two design tasks, and c) their scores for each of the three key science practices of selecting, testing, and describing (Figures 3 and 4). Table 8 indicates that for all six of these measures, the postinstruction means were higher than the preinstruction means. For four of the measures – the total scores, stepstool task scores, habitat task scores, and “test” sub-task scores – the pre/post increases were statistically significant according to a Wilcoxon signed-rank test of differences ($p < .05$, $n = 9$, 2-tailed tests).

Table 8
*Average Pre-Instruction, Post-Instruction, and Gain Scores on Design Task Interviews*

| Measure | Pre-instruction Mean (SD) | Post-instruction Mean (SD) | Average Paired Gain, Pre/Post (SD) | Effect Size, Cohen’s d |
|---|---|---|---|---|
| Total Interview Points (out of possible 36) | 8.8 (3.4) | 13.8 (3.9) | 5.0* (3.7) | 1.34 |
| Part I: Stepstool Task Points (out of 18) | 6.2 (2.4) | 10.0 (3.8) | 3.8* (3.6) | 1.27 |
| Part II: Habitat Task Points (out of 18) | 2.6 (1.5) | 3.8 (1.8) | 1.2* (1.1) | 1.27 |
| Points on “Select” Sub-Tasks (out of 16) | 2.7 (2.8) | 4.3 (2.3) | 1.7 (3.0) | 0.56 |
| Points on “Test” Sub-Tasks (out of 16) | 2.6 (1.4) | 5.6 (2.4) | 3.0* (1.7) | 1.73 |
| Points on “Describe” Sub-Tasks (out of 4) | 3.6 (0.7) | 3.9 (0.3) | .33 (.87) | 0.38 |

**Key:** * With 2-tailed Wilcoxon signed ranks test, average gain is significant at the $p < .05$ level, $n = 9$.

**Note:** Effect size is calculated according to the Cohen’s *d* statistic, where $d$ = (average of paired differences) ÷ (standard deviation of paired differences)

[5] A Wilcoxon signed-rank test was conducted because the small sample size ($n = 9$) and non-normal distribution of scores violated the assumptions of the parametric paired sample t-test. The signed-rank test is a nonparametric test of significant difference between paired samples of an interval measure.

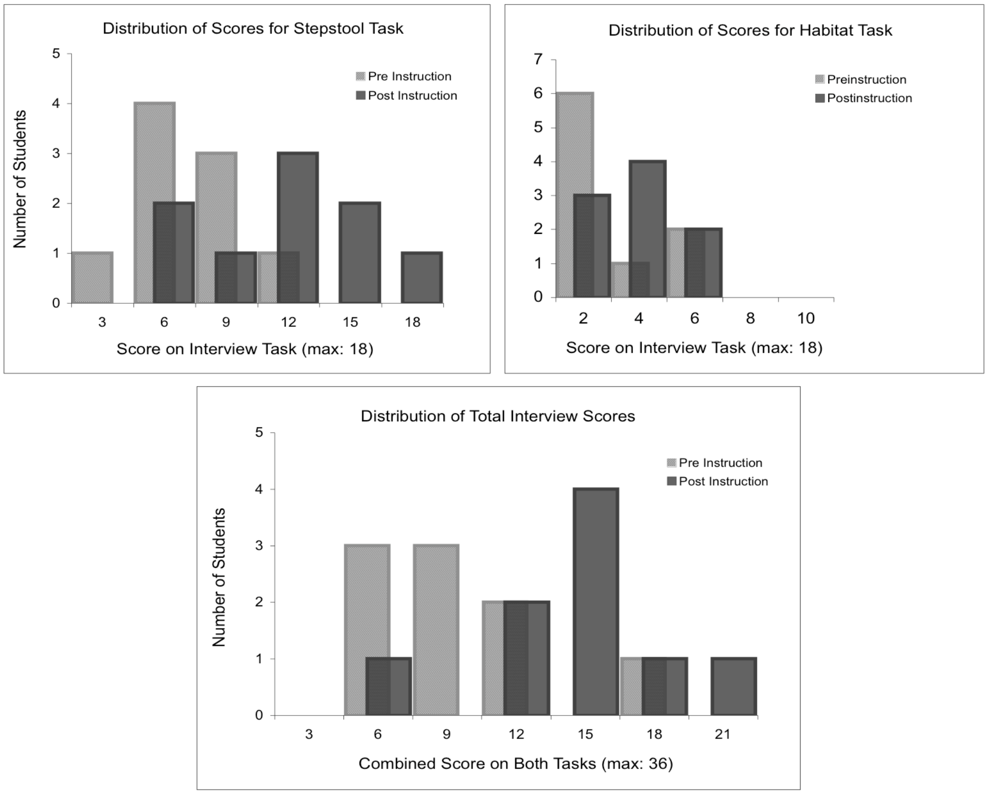



*Figure 3.* Preinstruction and postinstruction histograms of the students' scores on the stepstool task, habitat task, and total interview (sum of the stepstool and habitat scores).

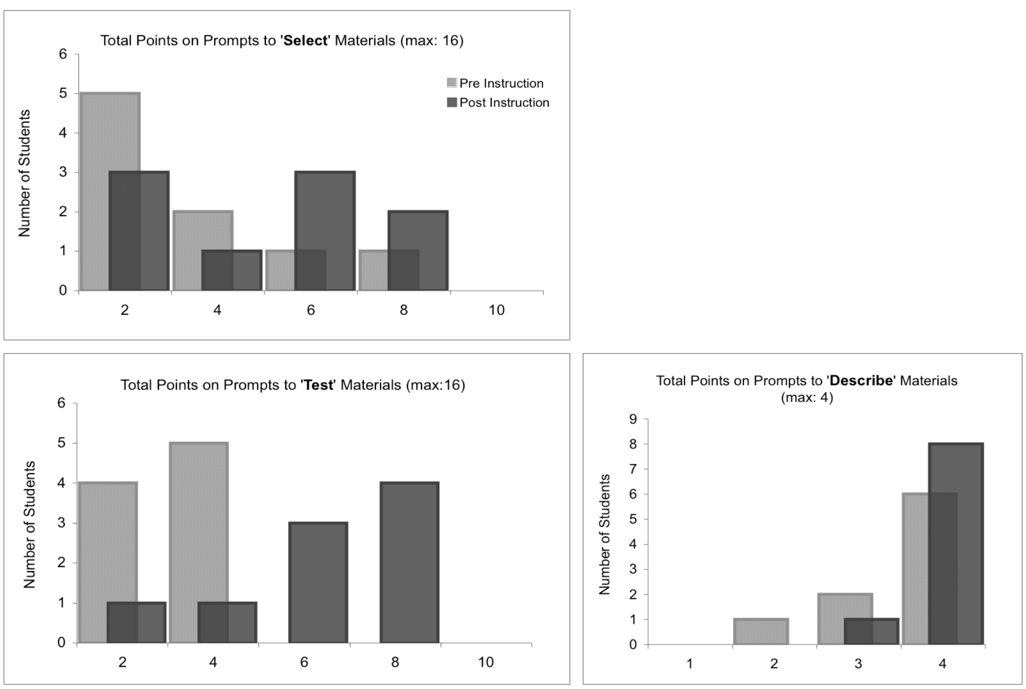



*Figure 4.* Preinstruction and postinstruction histograms of the students' scores in each of the interview's science-practice categories: selecting, testing, and describing materials.

In addition to improvements in mean scores, there were also positive pre/post changes in the numbers of students who provided particular kinds of selection rationales and material property tests. Table 9 displays additional evidence of learning about material selection. First, after instruction, students simply had more to say about why they chose a certain material: the majority of the students offered more rationales for their material selections during the post-interview than during the pre-interview. Second, at postinstruction, more students provided rationales that were relevant to the design goal. A third positive change occurred in the number of students who provided at least one rationale that was accurate to the material sample and also relevant to the design goal. Finally, there was an increase in the number of students who avoided simply restating the design goal as a rationale (i.e., before instruction, more students made statements like, "I would choose the bamboo because then the stepstool won't collapse.").

Table 9
*Frequency Table for Key Material Selection Practices*

| | Students who provided greater number of rationales at postinstruction | Students who selected most appropriate material for the design goal[a] | | Students who provided at least one relevant rationale (accurate or inaccurate) | | Students who provided at least one accurate and relevant rationale | | Students who avoided the circular reasoning of citing design goal as rationale | |
|---|---|---|---|---|---|---|---|---|---|
| | | Pre | Post | Pre | Post | Pre | Post | Pre | Post |
| Stepstool Task | 5 | 7 | 7[b] | 6 | 8 | 5 | 8 | 4 | 8 |
| Habitat Task | 7 | 5 | 4 | 2 | 5 | 1 | 2 | 7 | 7 |

a PVC was more appropriate for the stepstool task; foam was more appropriate for the habitat task.
b One student switched from PVC to bamboo; one switched from bamboo to PVC
*Note.* Values represent frequencies; n = 9.

As shown in Table 10, further analysis of the *test* sub-tasks provides additional evidence of learning about determining physical properties by conducting tests. First, at postinstruction, almost all students proposed a greater number of distinct tests for strength than they had at preinstruction, and two thirds of students proposed a greater number of distinct tests for insulation. Second, after instruction, more students proposed tests that could be conducted without full prototype construction. Further, more students proposed non-prototype tests that were central to determining the given property. Another noteworthy change occurred in the number of students who avoided using size and weight measurements to determine strength and insulation[6]. For both design tasks, there was a pre/post increase in this number.

Table 10
*Frequency Table for Key Material Property Testing Practices*

| | Students who proposed greater number of tests at postinstruction | Students who proposed tests that do not require prototype building | | Students who proposed centrally important tests that do not require prototype building | | Students who did not propose measuring size or weight to find strength or insulation | |
|---|---|---|---|---|---|---|---|
| | | Pre | Post | Pre | Post | Pre | Post |
| Stepstool Task | 8 | 8 | 9 | 6 | 7 | 6 | 8 |
| Habitat Task | 6 | 7 | 9 | 0 | 2 | 4 | 7 |

*Note.* Values represent frequencies; n = 9.

*Connecting Pre/Post Changes to Instructional Experiences*

This study's second research question asks, how does the change in children's interview performance compare to their performance on the corresponding instructional activities? The correlations between interview gains and workbook scores partially answer this question.

*Workbooks: overall performance.* Performance on the curriculum module workbooks varied greatly across the nine students and across tasks. The maximum possible workbook score

[6] Students who suggested weighing the material or finding out how heavy it was might have been referring to an undifferentiated weight/density concept (Smith, Carey, & Wiser, 1985). However, even if students were proposing to measure density, they were still incorrect. Density alone is not central to determining strength or to determining insulation (e.g., liquid water is dense but not strong nor a good insulator; on the other hand; carbon fiber is *not* dense but it *is* strong and a good insulator).

was 60 points, but the students' actual workbook scores ranged from 18 to 51 points (see Appendix F).

*Correlating interview and workbook performance.* Figure 5 shows that the students' interview scores were positively correlated with their workbook scores. As workbook completion and proficiency increased, the magnitude of pre/post interview gain also tended to increase. Greater gains on the individual stepstool and habitat interview tasks, respectively, tended to be associated with higher completion of and proficiency on the corresponding strength and insulation workbook tasks (Spearman $\rho = 0.478$, $p < .05$, $n = 18$). Interview gain was calculated by subtracting a student's preinstruction score for an interview design task from the postinstruction score for that task.

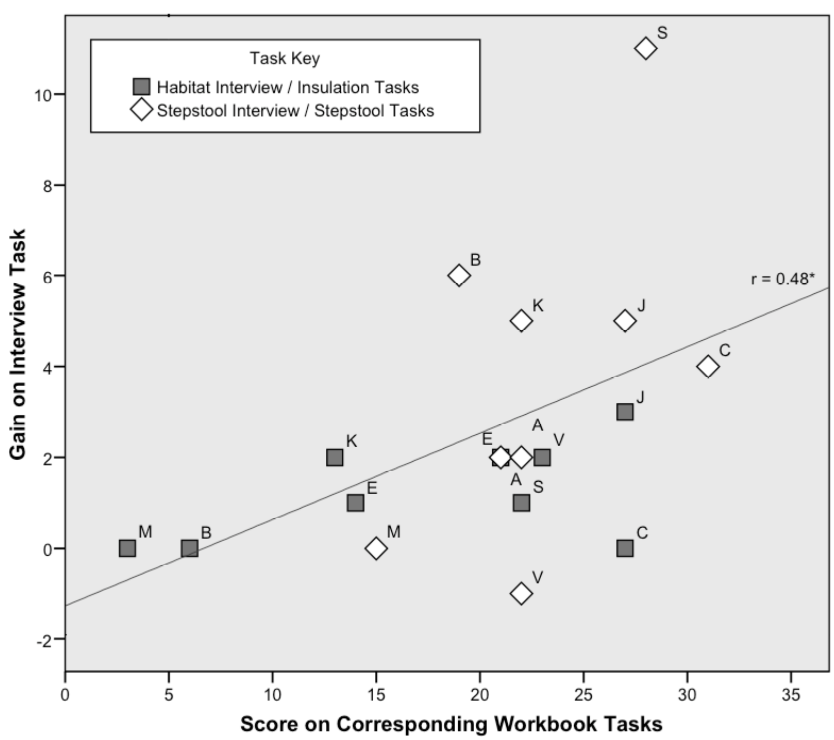



*Figure 5.* Students' gains on the interview design tasks plotted against their scores on the corresponding workbook exercises (numerical data are shown in Table 11). The letters adjacent to the markers identify individual students.

More evidence of the link between interview performance and instructional engagement comes from examining in greater detail the workbook and interview responses of individual students. In the next section, I use the work of four students to describe qualitatively the nature of the connections between the interviews and workbooks. It may be helpful, however, to pause and review the two main findings that have emerged so far.

First, the data have shown that students performed quite disparately on the stepstool task compared to the habitat task. The habitat task elicited ideas that were less scientifically normative but more diverse than the ideas elicited by the stepstool task. For the habitat task, students often strayed from the stated design goal, introduced their own goals, and described properties and tests that would meet these more personally meaningful goals.

Second, the data have shown that on average, students' postinstruction interview responses exhibited more proficient science practices (and earned significantly higher scores) than their preinstruction responses. Greater gains on the interview tasks tended to be associated with higher completion of and proficiency on the workbook tasks. However, all students showed some improvement on the interview tasks. At postinstruction, 8 of the 9 students proposed a greater number of distinct tests for strength than they had at preinstruction, and 6 of the 9 students proposed a greater number of distinct tests for insulation. Additionally, more students proposed tests that could be conducted without full prototype construction, and more students proposed non-prototype tests that were central to determining strength and insulation.

*Deepening the Interview-Instruction Connection: Examples from Individual Students*

An examination of the data from four individual students provides more evidence to link interview performance with instructional engagement. Four instructive examples are those of a)

Macon, the student who improved the least on both interview tasks, b) Stella, the student who improved the most on the stepstool interview task, c) Julia, the student who improved the most on the habitat interview task, and d) Chinelle, the student who received the highest workbook score. To illustrate how these exemplar students were selected, Table 11 highlights the highest and lowest scores within each set of tasks.

Table 11
*Interview Gains and Workbook Scores by Individual Student*

| | | | Interview Gain by Task | | Workbook Score by Task | |
|---|---|---|---|---|---|---|
| Student | Sex | Typical Science Test Achievement | Stepstool | Habitat | Strength | Insulation |
| Ava | F | Middle | 2 | 2 | 22 | 21 |
| Brian | M | Middle | 6 | 0 | 19 | 6 |
| **Chinelle** | F | Low | 4 | 0 | **31** | **27** |
| Elisa | F | Low | 2 | 1 | 21 | 14 |
| **Julia** | F | Low | 5 | **3** | 27 | 27 |
| Katie | F | Middle | 5 | 2 | 22 | 13 |
| **Macon** | M | Middle | **0** | **0** | **15** | **3** |
| **Stella** | F | Low | **11** | 1 | 28 | 22 |
| Viola | F | High | -1 | 2 | 22 | 23 |

*Macon: No improvement on interview tasks.* Macon's scores did not increase on either the stepstool task or the habitat task, and his total interview score at postinstruction was the lowest of all nine students' scores. Both before and after instruction, Macon's responses comprised a narrow set of ideas focused on the property of weight. For the strong stepstool task, he selected PVC because it felt heavier. Then he proposed weighing the PVC on a scale to test for its strength. The excerpt below is from the pre-interview, but these types of responses persisted through the post-interview.

M: [Chooses PVC pipe for as better material for stepstool leg.]
I: Okay. Why do you think that would be the better material for the stool?

M: Because it's metal [incorrectly identifying PVC plastic as metal].
I: Okay, because it's metal. Why would that make it better?
M: Because metal's stronger than wood [3-second pause].
I: Can you tell me more about that? [5-second pause; M shrugs shoulders] What, why is metal stronger than wood?
M: Because it's heavier. And heavier stuff is stronger.
I: ….Okay.[7] What is something you could do to measure exactly how strong this material is? So that you could tell -
M: You could weigh it?....Um, you can put it on, like, one of those weigh things that people stand on.
I: And then what would that tell you about how strong it is?
M: It would, it would, because heavier stuff makes it stronger so, if it's heavy, if it's heavier than this [picks up bamboo], then it [PVC] would be stronger.
I: …Okay, any other tests that you could do to tell how strong this material is? [5-second pause] Is there any other thing you could do to test the strength of this one, to tell the other engineers about the strength?
M: [3-second pause] No.

For the insulated habitat task, Macon also focused on weight as a priority in establishing the property of a material. He chose aluminum because it felt heavier, and he suggested weighing the aluminum to determine its insulation ability. Again, he offered these responses at both preinstruction and postinstruction. To Macon, the property of weight could be used to determine both strength and insulation. The following excerpt is from his pre-interview but is representative of his post-interview responses as well.

M: This one [points to metal tray].
I: Why do you think this would be the better material for -
M: Cuz it's heavy and stronger.
I: And why would heavier and stronger make it better?
M: Because, if it's thicker, then it might be heavy and strong, and if it's heavy and strong it makes stuff warm.
I: Okay, why do you think that something heavy and strong makes stuff warm?
M: Because it's thick.
I: And what does the thickness do with the warmth?
M: It traps it all inside of something.
I: Oh, okay. And now imagine the other engineers, so you picked this metal, the other engineers want to know exactly how good it is at insulating. …It

[7] Ellipses (…) indicate that a repetitive portion of the utterance has been omitted from this excerpt to save space. Ellipses do not indicate a pause in speech.

means how good it is at keeping the temperature the same…inside the house. How could you measure how insulating this material is?
M: You could weigh it on a scale.
I: And if it weighed more, what would that mean about how insulating it is?
M: It would mean that it would be warmer.

Macon's workbook performance aligned with his lack of improvement on the interview tasks. His workbook score, 18 out of the possible 60 points, was the lowest of all nine students' scores. He responded to very few of the exploration or follow-up prompts, and as a result, his workbook contains very few complete phrases or drawings. His workbook also contains very little data from the class's tests of material properties, with one major exception: Macon recorded all the results from the sound absorption test and responded to the follow-up prompt about soundproof materials. Perhaps as a result, the one new idea posed by Macon in his postinstruction interview involved sound absorption. Despite several reminders that the design goal for the second task was a well-insulated habitat, Macon based his initial material selection for that task on the samples' soundproofing qualities. He declared that the mouse would not want to be bothered by outside noises, and foam was the best wall material because its sound absorption ability had been proved in class.

In Macon's example, a lack of improvement on the interview followed minimal engagement with the instructional activities. During science instruction, he was present in the classroom and had the opportunity to participate with his teacher and classmates in the shared events of selecting and testing materials. However, because Macon produced very few written or drawn representations of the steps they took to select and test materials, we cannot determine the extent of his participation in those events. It is possible that by not carrying out the tasks of writing and drawing about his class's science investigations, Macon's chances for learning the practices of materials science were limited.

*Stella: Greatest improvement on stepstool task.* Of all nine interviewed students, Stella showed the greatest pre/post improvement on the stepstool task. Both before and after the module, she expressed many ideas for both design tasks, but only on the stepstool task did her responses substantially improve in their level of sophistication. At preinstruction, Stella chose bamboo for the stepstool simply because of its material kind: "because it's wooden." To test for the bamboo's strength, she proposed standing on it and trying to break it in half. To describe the bamboo, she offered only four descriptive terms, and none of them were extensive properties.

> I: I want to know which you think would be the better material to use to make the stepstool.
> S: This [bamboo] one because it wouldn't break… because this [bamboo] one's wooden….This [PVC] one would break because… this [PVC] one's, like, hard, but it's not wooden.
> I: ….Yeah, okay, so what could you do to tell the engineers exactly how strong this [bamboo] is? How could you measure its strength?
> S: You can stand on it.
> I: And what would that tell you about how strong it is?
> S: Uh, um, tell you if its sturdy or not sturdy, and it's like, uh, if it's like, uh, not sturdy to stand on it, and it like, start standing on it and you'll fall.
> I: Okay, any other tests you could do to find out how strong it is?
> S: Like if you can break it in half and see if you can stand on it and it won't break apart.

At postinstruction, Stella switched to choosing PVC for the stepstool, and she based her choice on its intensive properties: it was "strong," "sturdy," and "hard." To test for strength, she suggested a number of useful tests, including, once again, standing on the PVC, hitting it with a golf ball, compressing it, and building a prototype stepstool. She described the bamboo in 10 different ways, which included both intensive and extensive properties.

> I: So which do you think would be better?
> S: This one [PVC].
> I: Okay, why would that be better?
> S: Because, um this [bamboo] one like is wood, and this [PVC] one um, is like stronger, cuz when you stand on it, you're um, you won't fall….
> I: Why do you think this one [PVC] is stronger?

S: Because, um, it has, um, good materials [3-sec pause] and, [3-sec pause] it's sturdy….Like, um, it's hard, and um.
I: ….So imagine the other engineers say, okay, you picked this material, but we want to know exactly how strong it is. What could you do to measure exactly how strong this material is?
S: You can, um [3-sec pause], put it like, um, on something, and, um make sure it doesn't break with anything…. Like you can um [3-sec pause], like um, you can uh, you can try to stand on it.
I: …. Are there any other tests that you could do with this material, to see how strong it is?
S: You can um, make something out of it, and then you can like um, like do it with, that thing we did for a golf ball, you can….We hit it with a ball.
I: What did you do in that, is there anything you did in that project where you tested for strength, that you could use to test for the strength of this?
S: Um, you can like um, we pressed it down. [Holds PVC upright, pushes on top] You can um, you can uh [6-sec pause] you can [12-sec pause] you can make it out a stepstool, and then you can try it.

Like Macon's workbook, Stella's workbook corresponds well with her interview performance. Out of the 60 possible points, Stella earned 43 points, the third highest total score. On the tasks related to strength, she earned the second highest score. With the exception of the tasks for the second lesson, Stella's responses to prompts about strength were complete, and they indicated proficiency at materials science practices. Her writing and drawing indicate that she was attempting to connect materials' characteristics to their appropriateness for certain tasks. In response to a prompt about useful materials for support columns, she listed, "brick, wood, metal, gold, stone, silver, and cement," because these materials are "sturdy, strong, stable, able to support weight."

Stella's much improved stepstool task performance and her high level of workbook engagement indicate that she learned a great deal about the property of strength, and that the opportunity to write and draw about strength in several different contexts enabled her to construct and explore new ideas. However, her much more moderate improvement on the habitat interview task suggests that students' learning about materials science practices may not immediately generalize to all material properties. Stella was able to apply her classroom experiences to a novel interview task involving the property of strength, but she showed less proficiency at selecting and testing materials for the novel task focused on insulation.

*Julia: Greatest improvement on habitat task.* Julia's gain on the habitat task was the highest among all nine students. She also improved substantially on the stepstool task, and she was one of only three students who improved on all three sub-tasks: selecting, testing/measuring, and describing materials[8]. Julia, an English language learner, shifted from providing just a few vague ideas in her preinstruction interview to offering many well-articulated ideas at postinstruction. For the preinstruction habitat task, she chose metal for the habitat walls because it "can't break," and she proposed testing the metal with a hammer to see if it was strong. Neither her selection rationale nor her proposed test was relevant to the design goal of insulation. Further, for the describe sub-task, she generated only three descriptors for the metal.

> I: Why would you pick that [metal] one as the better material?
> J: Because this [metal] one, it can't break and this [foam] one is kind, could break, if you let it, do it like this.
> I: Mm, this [foam] white one could break.
> J: Yes.
> I: What if you had to think mostly about the job of keeping the mouse house warm, would you still pick this [metal] one, or would you pick this [foam] one?
> J: Still pick this [metal] one.
> I: Okay, tell me about that.
> J: Because [3-sec pause], um, if you like, made it with, you could like, maybe you could like make a door with it?

At postinstruction, Julia still cited metal's strength as part of her selection rationale and as a property to be tested, but she now introduced other rationales and tests that did relate to

---

[8] Julia, Stella, and Katie are the three students who have higher scores at postinstruction for all three sub-tasks, when the stepstool and habitat tasks are added together. When the stepstool and habitat tasks are considered separately, Julia has a higher score for the selecting and testing sub-tasks within both tasks. She has a higher describing sub-task score within the habitat task, and she maintained the maximum possible describing sub-task score within the stepstool task.

insulation. She reasoned in a relevant (though inaccurate) manner that metal "won't let the cold in," and she suggested touching the metal to see if it felt hot or cold as well as building a prototype house to test its insulation abilities. She described the metal in seven different ways.

> I: So why would you pick this one [metal] as the better material for keeping the temperature warm?
> J: Because, um, when you put it in a house, this one [metal] will make it hot, and this [foam] will not, because it will like, cover the hotness and somehow it's gonna go in, the cold.
> I: ….So how could you measure exactly how good this material [metal] is at keeping the temperature the same? ….
> J: Um, I would [6-sec pause], um [6-sec pause], I would like put it on, and touch it, and if it's cold, and maybe when you build it, it will make it hot because it lets in the –
> I: So when you touch it, it feels kind of cold
> J: Yeah.
> I: But when you use it to build something –
> J: Yeah, inside.
> I: It would keep it warm?
> J: Yeah. For the mouse.

As did Macon and Stella, Julia engaged with her workbook tasks in a manner that aligned well her pre/post interview changes. Julia's workbook earned the second highest total score and the second highest score for tasks related to insulation. She engaged in all of the exploration prompts, which allowed her to practice reasoning about material and object selections. For example, she reasoned that a winter jacket keeps you warm because "it is thick and [it] keep[s] your body heated by sorta stick on to you." Julia also completely and accurately recorded results for both tests related to insulation (for the model house's wall material and roof material). Her computations of temperature changes were correct, and her conclusions about the most insulating materials were consistent with the data.

*Chinelle: Highest workbook score.* In the previous three examples, each student's level of workbook engagement aligned to his or her interview performance. Chinelle's example is less straightforward. She earned the highest total workbook score (51 points). If her results followed the pattern of Macon, Stella, and Julia, the quality of her workbook would predict a great improvement on her interview responses. But instead, Chinelle's 4-point interview gain was below the average of 4.9 points, and her post-interview score, 10 points, was the second lowest among the students.

These mismatched indicators beg the question: why did Chinelle's engagement with the instructional activities *not* lead to greater pre/post interview improvement? One possible explanation is that Chinelle's interview gain was small because she performed so well on the pre-interview that there was little room for improvement. However, her pre-interview score of 6 points was among the lowest. Another explanation for Chinelle's lack of pre/post improvement is that she somehow failed to engage in a critical instructional activity. While Chinelle's workbook score was the highest among all the students, it was still 9 points shy of the total possible 60 points. Perhaps one of the activities that accounted for those missing 9 points was essential for establishing newly learned science practices related to materials and objects.

Indeed, 3 of Chinelle's missing points were on the very last workbook page, which prompted the students to review how they had treated a given property during their model house design.[9] The prompts were, "How did you test for the property of ____?" and "Why is the property of ___ important to your house?" (see Figure 6). Chinelle was assigned the property of insulation, but her page was only partially complete and gave no indication of proficiency at science practices. She wrote only, "The white is a good insulator." This work earned only 1 out of 4 possible points. In contrast, Stella and Julia earned scores of 3 and 4 points, respectively, on their work for this page. Stella wrote that to test for strength, "you have to push down and see if it pushes," and that strength is important for a house because you have "to keep it up." Julia

---

[9] Each individual student was assigned a different property by the teacher. The lesson plan called for all students to share their responses so that all properties could be reviewed by the class as a whole.

recalled that for her reflectivity test, she had "got foam, [and] cardboard of the color black or white," and that reflectivity is important because one needs to "let the heat out not in the house."

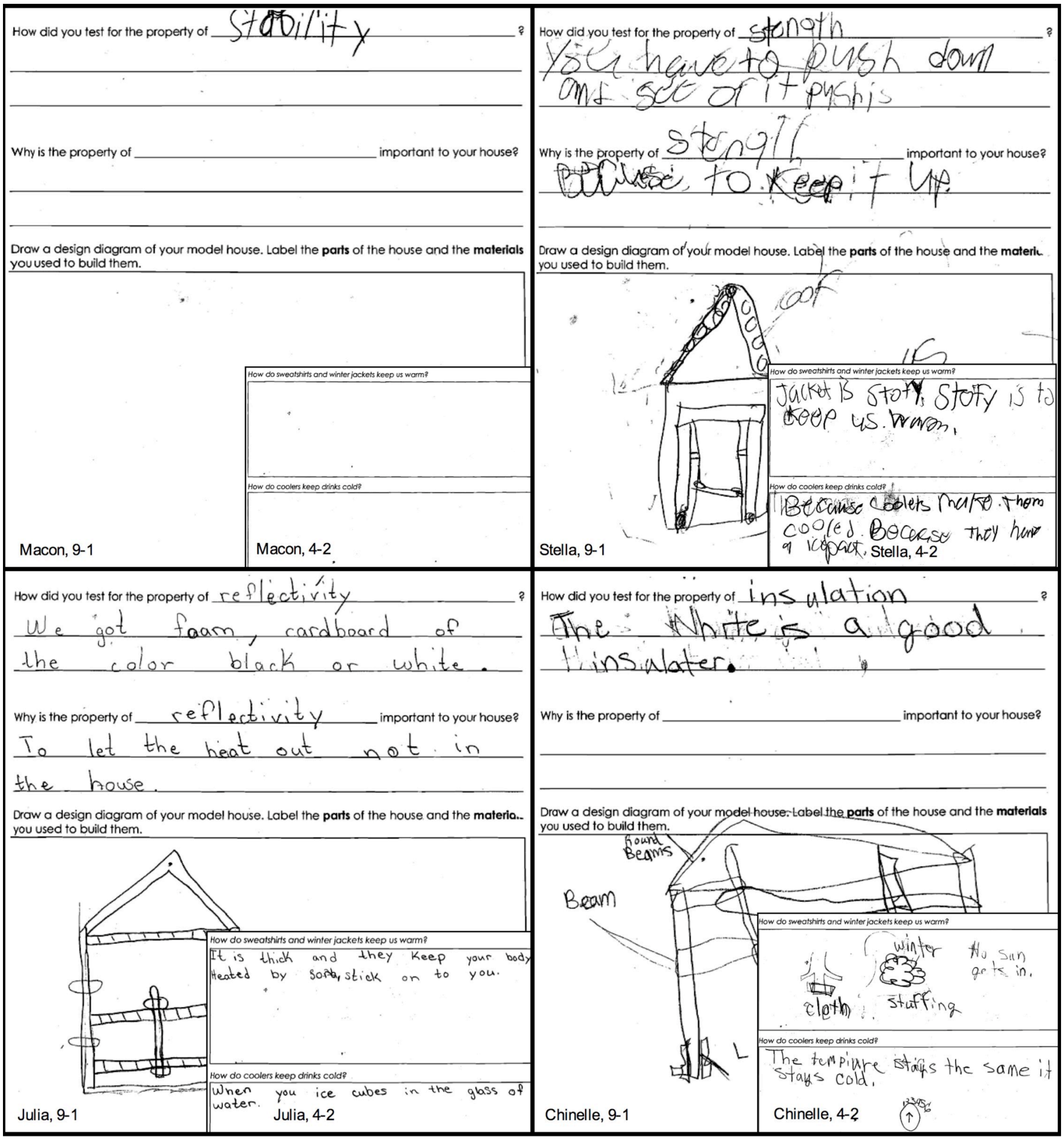



*Figure 6.* Workbook pages Macon, Stella, Julia, and Chinelle. Page 9-1 is the final page. The inset page 4-2 is an exploration question from the middle of the unit; it focuses on insulation.

Julia used her final workbook page to write about temperature regulation for the model house, and she performed well on the habitat task in her post-interview. Likewise, Stella concluded her workbook by writing accurately about the property of strength, and she expressed sophisticated ideas about strength in her post-interview. Chinelle, on the other hand, did not take advantage of the last workbook task, and her subsequent post-interview performance showed relatively little improvement. Macon, similarly, earned 0 points on his last workbook page, and he showed negligible improvement on the interview assessment. These patterns suggest that the closing instructional activity, which tasked students with reviewing the actions and reasoning that led to their model house design decisions, was a critical factor in reinforcing their newly learned materials science practices.

The more general connection between workbook engagement and interview performance, however, was that students who took advantage of the workbook's opportunities to record observations and reason about material properties showed greater gains on the stepstool and habitat design tasks. Students who completed more of the workbook and showed partial or full proficiency at its required science practices were more likely to improve their abilities to select, test, and describe the materials for the stepstool and habitat.

# Discussion

This study has focused on the science-as-practice aspect of science learning, and in particular, on third-grade children's learning of the practices of introductory materials science. In the curriculum implemented in this study, *Design a Model House,* students were given opportunities to participate in the practices of selecting, testing, and describing materials. Driver et al. (1994) suggest that science learning is facilitated when "a teacher, familiar with the scientific way-of-seeing, makes the cultural tools of science available to learners and supports their (re)construction of the ideas through discourse about shared events" (p. 11). The "shared events" in this study's curriculum involved explorations and tests of material samples, and the use of those materials to construct the components of a miniature model house.

The design tasks incorporated into this study's interview and workbook assessment tools served to situate the children's thinking about materials, objects, and their properties. The stepstool, habitat, and model house design scenarios provided authentic contexts for engaging in materials science practices and for reasoning about materials and objects (Brown, Collins, & Duguid, 1989; Roth, 1996). Several intriguing patterns emerged from the interview responses and workbook performances elicited by these design scenarios.

### *The Children's Approaches to Materials Selection Design Tasks*

One pattern was the great difference in performance between the stepstool and habitat tasks. The design goal of an insulated pet habitat was not well understood by all the children; it was not a goal that made sense to all of them. Even when the interviewer avoided the term "insulation" and posed the design goal as choosing a wall material that would "keep a warm temperature inside the habitat," the children chose to select materials for and propose tests for other attributes of pet habitats. In contrast, for the stepstool task, they remained focused on the interviewer-posed attribute of strength. The goal of a strong stepstool made sense to the children, but the notion that insulation (i.e., maintenance of a particular temperature) is an important design goal did not resonate with their vision of adequate pet habitats. It also is possible that because both the curriculum and the first interview task emphasized the design goal of strength, the children concluded that it was important to incorporate strength testing into any material selection task. In fact, more than half of the students suggested testing the habitat materials for strength, even though the stated design goal was a well-insulated habitat, not a strong habitat.

The second pattern emerging from this study is that the children privileged perceptually accessible properties (Smith, 2007), including thickness, surface texture, felt weight, and felt temperature, as they constructed their rationales for material selection. They also privileged perceptually accessible properties as they proposed methods for testing strength and insulation, two properties that are not immediately observable. This preference for perceptual accessibility is consistent with previous findings (Gustafson et al., 1999; Lewis & Linn, 1994), and it can help to explain why, despite the substantial gains from pre to postinstruction, the highest postinstruction interview scores were still less than 60% of the maximum possible score. The properties that the children were asked to consider were complex physical properties, determined by underlying molecular structure, bonding, and manufacturing methods. Primary grade children lack experience with these invisible determinants of complex properties (Nakhleh & Samarapungavan, 1999). As a result, when the children in this study tried to select and measure materials that would be strong and insulating, they based their ideas on directly observable properties rather than on underlying intensive properties. Their reliance on immediate perception placed an upper boundary on the material selection and property measurement abilities that they would be able to show – which was reflected in their scores at 60% of the maximum possible in the postinstruction interviews.

A third pattern within this study's findings is that despite their preference for perceptually accessible properties, all the children included at least one intensive property in their postinstruction rationales for material selection, and all but one child proposed at least one useful material test for an intensive property. These results indicate that 8- and 9-year-old children are able at least to *recognize* and *test* intensive properties, even if they cannot generate mechanistic explanations for those properties' existence. Smith, Carey, and Wiser (1985) conducted a study in which most 6- to 9-year-olds simply invoked material kind to explain why a clay ball is heavier than a wax ball. That is, they said that one ball was heavier because it was made of different stuff than the other ball. Smith et al.'s finding could lead one to predict that the students in the present study would simplistically invoke material kind and state that the bamboo was weaker just because it was made of wood, and the aluminum tray was less insulating just because it was made of metal. However, at postinstruction, very few of the students appealed to material kind to explain why one sample was weaker or less insulating than the other. Instead, they made serious attempts to construct a causal relationship between immediately accessible material and object properties and the less accessible properties of strength and insulation. This might be because by the end of the module, the students were starting to connect their observations of perceptual properties with their developing knowledge of testable or measurable properties. Wiser and Smith (2008) suggest that at the elementary level, students should start to account theoretically for properties of matter by coordinating descriptions at two levels: perceptual appearances and measurable properties or variables. In this study, the students' increased use of intensive properties during the postinstruction *select* and *test* sub-tasks provides some evidence that they were starting to carry out this kind of coordination of properties.

*How the Children's Approaches Changed Over Time*

The most general gains occurred on the task of proposing methods to *test* for material properties. In the postinstruction interview, students exhibited greater understanding of the use of engineering tests to determine a material sample's appropriateness for a particular design goal. It is important to note, however, that understanding how to test for a material's strength or insulation ability does not necessarily translate into understanding the properties determined by the tests, nor into understanding the physical principles underlying the tests.

Another major characteristic of the pre/post changes seen in this study was the appearance of synthetic models in the children's ideas. When instruction from a teacher (or interviewer) causes a student to find incoherencies in his or her commonsense theory for making meaning of something, then the student might be inclined to create what Vosniadou (2002) calls a "synthetic model." This happens when "information received through instruction seems to become assimilated to the initial explanatory framework creating synthetic or internally inconsistent models" (Vosniadou, 2002, p. 62). One kind of evidence that a child holds a synthetic model is a confusing later explanation that contradicts an earlier more understandable (but possibly incorrect) explanation.

From among this study's nine students, there were at least two examples of synthetic models in students' postinstruction interview responses. The first synthetic model that emerged was Katie's idea that a material's ability to block light transmission determines its ability to block heat transfer. In both her workbook and postinstruction interview, Katie suggested that if she held a material up to a light and saw light passing through, she would know that the material was a poor insulator. This model of equating transparency with thermal conductivity could have been based on the curricular activity in which a lamp provides the heat for an insulation test. Katie possibly viewed the lamp's primary function as a light source rather than a heat source.

The second apparent synthetic model was the idea, put forth by both Stella and Julia, that air flow into a space is equivalent to heat input into the space. At postinstruction, both of these students proposed methods for insuring that adequate air could enter the pet habitat. Stella wanted to make holes in the habitat walls, and Julia wanted to create an open door. They reasoned that if the air could flow into the habitat, then the habitat would be warm. I am unable to point to a instructional experience that might be related to this model for warmth maintenance, but it could be based on home experiences of opening doors and windows to regulate the indoor temperature. Their idea – that air flow is required to heat a space – is consistent with nonnormative ideas uncovered by previous research, including the ideas that heat itself is a gaseous substance (Erickson, 1979; Paik et al., 2007) and that materials with holes allow for better passage of heat (Clark, 2006).

*Connections Between Interview  Changes and Instructional Experiences*

From the four extended examples of students' interviews and workbooks, some tentative conjectures can be made about how writing and drawing tasks, such as those prompted in students' workbooks, might be associated with students' learning about the practices of materials science. First, the students who take the least advantage of opportunities to record their observations and reasoning about materials and objects (like Macon in this study) are likely to be those who exhibit the fewest changes in their approaches to materials selection and materials testing. However, as in the case of Macon's learning about soundproofing, small bursts of effort may lead to isolated sets of new or newly adjusted ideas.

Likewise, the students who take greatest advantage of opportunities to record their observations and reasoning (like Julia and Stella in this study) are likely to be the students who show the greatest change in their approaches to the selection and testing of materials. Along this same vein, if a student's writing and drawing is especially focused on the property of strength, and it culminates with a summary of ideas about strength (like Stella's work in this study), that student's performance on the stepstool design task may improve to a greater degree than on the habitat design task. This is consistent with Olson's (1996) assertion that writing about abstract notions, rather than only speaking about them, has helped humans to advance new ideas and concepts. It also meshes with Olson's (1996) findings that when scientists began drawing pictures intentionally for the purpose of communicating with other scientists, entire fields of study, such as anatomy and botany, were advanced.

On the other hand, it is important to consider the possibility that a third-grade student's performance on the habitat design task would not improve even with tireless effort on the writing and drawing exercises related to insulation. Though intended to be primarily an assessment of science practices, the habitat task does require students to draw upon their ideas about heat and temperature. To earn maximum points on this task, students would have to differentiate between those two concepts, but abundant research has shown that in elementary school, most children conceive of heat as perceived hotness, and thus conflate heat and temperature into one undifferentiated heat/temperature concept (Shayer & Wylam, 1981; Wiser & Amin, 2001). Consequently, even after hours spent writing and drawing about thermally comfortable model houses, there may be a conceptual "cap" on children's ability to select, test, and describe materials in the context of designing an insulated structure.

Of course, even in the absence of conceptual challenges like those presented by thermal insulation, a complete workbook in and of itself does not necessarily lead to great changes in a student's practice of materials science. There are also specific kinds of tasks related to "writing and drawing" that may do more to foster the building of proficiency at materials selection, testing, and description. The closing workbook activity used in this study prompts the student to review the scientific reasoning behind his or her design decisions. In this activity, first the student must list the materials he or she chose for the model house and explain why those choices were made. Second, the student must describe how he or she tested for a particular property (assigned by the teacher) and explain why that property is important to the model house.

This kind of activity appears to be a critical exercise for learning materials science practices. If a student engages in almost all of the workbook tasks but does not participate in this closing reflection (like Chinelle in this study), he or she may miss a key chance to consolidate newly constructed ideas and establish newly learned practices. In this study, Julia and Stella earned the highest scores for the final workbook task, and they also showed the highest total interview gains. This pattern suggests that teachers should stress the importance of the final task in the *Design a Model House* workbook, and this task should be expanded to prompt each student to review more than one material property.

Overall, the *Design a Model House* workbooks may have helped to distribute cognition (Salomon et al., 1991) during the course of classroom instruction (Puntambekar & Kolodner, 2005). They served as a tool that enabled the students to track the tests they conducted on potential materials as well as the selection decisions they made based on the test results. In this way, the workbooks may have eased the cognitive effort required to connect test results back to experimental procedures, and to link test results forward to decisions about courses of action (Shepardson & Britsch, 2001).

Pea (1993) asserts that learning is fundamentally a conversational process in which negotiation of meaning and appropriation of others' discourse are necessary acts. Therefore, in order for students to learn science, the science teacher must provide access to scientific discourse, and students must *converse* in science in order to learn science. If writing is a model for conversation (Olson, 1996), and conversation is essential for science learning (Pea, 1993), then the writing and drawing exercises in science workbooks can be seen as models of the ongoing dialogue between student and teacher, and between student and fellow students. The production of written and drawn representations may serve to reorganize the ideas to which students are exposed during teacher-student and student-student conversations.

The distributed cognition notion of *cognitive residue* (Bell & Winn, 2000; Salomon, Perkins, & Globerson, 1991) might sharpen our thinking about the cognitive impact of writing and drawing exercises in science. Humans generally assume that when they appropriately use tools (such as the workbooks and pencils in this study) as part of a cognitive act, they are more productive and their intellectual capabilities are enhanced. However, what is not quite as obvious is that those tools can leave a sort of residue that supports intellectual activity later, even when the tools are no longer present:

> Exposure to artifacts, whether they have been designed to help us internalize and develop cognitive skills, like the Writing Partner, or not, like television, leaves a residue that can serve individuals well when they must perform tasks in the absence of the tool. (Bell & Winn, 2000, p. 131)

Theorists often focus on computers as technological tools, but this role could be played by many other human-created artifacts, including the representations children both interpret and produce in their science workbooks. For example, in the *Design a Model House* module, the students became accustomed to filling out a "Pass/Fail" chart to indicate whether the material samples under consideration for the model house passed particular material property tests. When the students participated in the postinstruction interview, they did not have a blank "Pass/Fail" chart in front of them. However, their experience with recording test results may have left a cognitive residue that enabled them to suggest tests with a clearly dichotomous pass/fail outcome for the material property in question. In this way, the production of written and drawn representations may play an amplification role in children's learning of science practices.

Previous research studies as well as this study's findings suggest a need for assessment instruments that are more sensitive to the kinds of science learning fostered by design-based science instruction. This study's clinical interviews generated sophisticated student performances and lent support to the notion that design-based science instruction should use design-based assessment tasks to measure science learning. Aligning assessment to instruction in this way allows us to capture the unique effects of design-based instruction on students' science practices.

In this study, the students engaged in classroom activities to achieve the design goal of constructing a stable, quiet, thermally comfortable model house. Throughout instruction, they had opportunities to write and draw about those activities in their science workbooks. After these experiences, eight of the nine students were better able to apply the practices of introductory materials science to two novel design goals, a stepstool and a pet habitat.

*Limitations of This Work*

The conclusions that can be drawn from this study's findings are limited in several ways. The first main set of limitations is due to the particular tasks and spoken prompts chosen for the clinical interviews.

First, since the interviews used particular material samples and posed particular design scenarios, they did not comprise an exhaustive assessment of introductory materials science. It is not possible to say how much students' selecting, testing, and describing practices would differ within design contexts other than strong stepstools and insulated habitats. These particular design scenarios involved two complex intensive properties which are difficult for children to access directly and which are often misassociated with observable extensive properties. One distracting extensive property is thickness, which many children assume incorrectly is of central importance to a sample's insulation. Another distracting extensive property is weight, which children often mistake as a determinant of a sample's strength. The existence of these distracter properties may have deflated the students' interview scores in this study. The interviews were also limited by their lack of control for students' familiarity with the two design scenarios. No task or score adjustments were made for students who had prior experiences with pet habitats or strong structures. One final weakness of the interviews was their lack of explicit distinction between materials and objects. The interviewer did not clarify whether the students should be considering the objects (plastic pipe, bamboo pole, metal tray, or foam tray) as a whole, or the materials from which the objects were made. Although this approach allowed for exploration of students' natural inclinations toward material properties or object properties, it prevented a more exhaustive assessment of their understanding within either category.

Four other important limitations are related to the overall design of this study. First, I did not collect information on the students' initial commonsense theories of matter. With this information, I would have been able to examine whether a student's commonsense theory of matter, including the presence or lack of material/object differentiation, is predictive of ability to select, test, and describe materials and objects and their properties. Second, the sample size for this study was small. With nine students, I could not use parametric statistical inference tests. Instead, I had to conduct more conservative and less powerful nonparametric tests, such as the Wilcoxon signed rank test. Third, I did not have a comparison group of students who participated only in the postinstruction interview. A comparison group would have allowed me to estimate how much of pre/post interview gains were due to repeated exposure to the interview tasks. The advantage of the identical pre/post tasks is that they enabled direct comparison of performance on the two interviews. However, when a student offered an idea at postinstruction that was similar to a preinstruction idea, I did not know whether the student was intentionally recalling what he or she had said earlier, or whether the student's idea had indeed undergone little change. The fourth limitation of the study design is that the connections made between interview gains and instructional experiences relies on an informed but unproven assumption. This assumption posits that completeness and proficiency of workbook responses is a good proxy for engagement in curricular activities. Without extensive field notes or video recorded data, I cannot be certain that workbook performance corresponds well with participation in classroom activity.

*Questions for Further Investigation*

As a result of the limitations discussed above and due to the limited scope of this study, several questions remain for further exploration. First, it would be instructive to change the deign scenario for the insulation interview task. What ideas about insulation or thermal conductivity would students express if they were asked to choose materials for a hot chocolate container rather than for the walls of a pet habitat? Second, the evidence provided by the one classroom considered in this study should be compared with evidence from other classrooms implementing this curriculum. To what extent are students' interview responses and workbook performances associated with their teacher's emphasis on particular concepts, processes, and goals?

Additional work to be done includes making curricular changes, creating additional design tasks for the interview, and more completely mapping out the conceptual field that includes material selection and testing. While this study was not intended to be an evaluation of the *Design a Model House* curriculum module, its findings do indicate that the module should provide students with more opportunities to explore the design goal of insulation. This is consistent with the literature (e.g., Wiser & Amin, 2001) on the substantial effort required to facilitate learning about heat and temperature. Along with improvements to the curriculum module, enhancements could also be made to the clinical interviews to provide the students with more exploration opportunities. Additional interview tasks should probe students' material selection approaches for design goals that involve other intensive and extensive properties, such as density, weight, tensile strength, surface hardness, and elasticity. A final area for future work is to map out where an improved understanding of material selection and material property testing can take students. For what learning experiences are these students better prepared now that they are more proficient at selecting materials and proposing tests for material properties?

# References


American Association for the Advancement of Science/AAAS. (1993). *Benchmarks for science literacy.* New York: Oxford University Press, Inc. Accessed online at http://www.project2061.org/publications/bsl/online/ch15/findings.htm#Ch3. Last accessed 2 January 2006.

Acher, A., & Arcà, M. (2006). Children's representations in modeling scientific knowledge construction. In C. Andersen, M. N. Scheuer, M. P. Perez Echeveerria, & E. Teubal (Eds.), *Representational systems and practices as learning tools in different fields of knowledge.* Rotterdam, The Netherlands: Sense Publishers.

Acher, A., Arcà, M., & Sanmartí, N. (2007). Modeling as a teaching learning process for understanding materials: A case study in primary education. *Science Education.*, *91*, 398-418.

Allen, S. M., Thomas, E. L.(1999). *The Structure of Materials*. New York, NY: J. Wiley & Sons.

Ashby, M.F. & Jones, D.R.H. (2005). *Engineering materials 1: An introduction to properties, application, and design.* Elsevier.

Ashby, M., Shercliff, H., Cebon, D. (2007). *Materials: Engineering, science, processing, and design.* Burlington, MA: Butterworth-Heinneman.

Au, K.-F. (1994). Developing an intuitive understanding of substance kinds. *Cognitive Psychology, 27,* 71–111.

Baxter, G., Bass, K., Glaser, R. (2000). An analysis of notebook writing in elementary science classrooms. *CSE Education Technical Report 533.* University of California, Los Angeles.

Bell, P., & Winn, W. (2000). Distributed cognitions, by nature and by design. In D. Jonassen & S. M. Land (Eds.), *Theoretical foundations of learning environments.* Mahway, NJ: Lawrence Erlbaum Associates.

Bloom, B. (1976). *Human characteristics and school learning.* New York: John Wiley.

Brizuela, B. M. (2004). *Mathematical development in young children: exploring notations.* NY: Teachers College Press.

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

Callister, Jr., W.D. (2007). *Materials science and engineering: An introduction. (7th ed.).* New York, NY: J. Wiley & Sons.

Carey, S. (1985). *Conceptual change in childhood.* . Cambridge, MA: MIT Press.

Carey, S. (1992). The origin and evolution of everyday concepts. In R. Giere (Ed.), *Cognitive*

*Models of Science (Minnesota Studies in the Philosophy of Science, Vol. XV)* (pp. 89-128). Minneapolis: University of Minnesota Press.

Carey, S., & Spelke, E. S. (1996). Science and core knowledge. *Philosophy of Science, 63*(4), 515-533.

Clark, D. B. (2006). Longitudinal conceptual change in students' understanding of thermal equilibrium: An examination of the process of conceptual restructuring. *Cognition and Instruction, 24*(4), 467-563.

Cole, M., & Griffin, P. (1980). Cultural amplifiers reconsidered. In D. R. Olson (Ed.), *The social foundations of language and thought*. New York: W.W. Norton.

Constable, G., Somerville, B. (2003). *A century of innovation: Twenty engineering achievements that changed our lives.* Joseph Henry Press.

Dickinson, D. K. (1987). The development of a concept of material kind. *Science Education, 71*(4), 615 – 628

diSessa, A. A. (1983). Phenomenology and the evolution of intuition. In D. Gentner & A. Stevens (Eds.), *Mental Models* (pp. 15-33). Hillsdale, NJ: Lawrence Erlbaum.

Dowling, N.E. (1998). *Mechanical behavior of materials: Engineering methods for deformation, fracture, and fatigue. (2nd ed.)*. Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher, 23*(5), 5-12.

Dym, C.L., & Little, P. (2004). *Engineering design: A project-based introduction.* John Wiley & Sons, Inc.

Education Development Center/EDC. (2003). *Insights: Changes of state teacher's guide, (2nd ed.)*. Kendall/Hunt Publishing Company.

Education Development Center/EDC. (2003). *Insights: Structures with trusses teacher's guide, (2nd ed.)*. Kendall/Hunt Publishing Company.

Education Development Center/EDC. (2003). *Insights: Mysterious powder teacher's guide, (2nd ed.)*. Kendall/Hunt Publishing Company.

Erickson, G. L. (1979). Children's conceptions of heat and temperature. *Science Education, 63*(3), 221-230.

Fortus, D., Krajcik, J. S., Dershimer, R. C., Marx, R. W., & Mamlok-Naaman, R. (2005). Design-based science and real-world problem-solving. *International Journal of Science Education, 27*(7), 855-879.

Fredricks, J.A., Blumenfeld, P.C., Paris, A.H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Ginsburg, H. P. (1997). *Entering the child's mind: The clinical interview in psychological research and practice* (pp. 115-158). NY: Cambridge University Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* New York: Aldine de Gruyter.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist, 53*(1), 5-26.

Greeno, J. G. (2006). Learning in activity. In K. L. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 79-96). New York: Cambridge University Press.

Gustafson, B. J., Rowell, P. M., & Guilbert, S. M. (2000). Elementary children's awareness of strategies for testing structural strength: a three year study. *Journal of Technology Education, 11*(2), 5-22.

Gustafson, B. J., Rowell, P. M., & Rose, D. P. (1999). Elementary children's conceptions of structural stability: a three year study. *Journal of Technology Education, 11*(1), 27-44.

Haber-Schaim, U., Cutting, R., Kirksey, G., & Pratt, H. (1999). *Introductory Physical Science* (7th ed.). Lakewood, CO: Science Curriculum Inc.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, MA: MIT Press.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child, classification and seriation.* London: Routledge & Kegan Paul.

Johnson, P. (2000). Children's understanding of substances, Part 1: Recognising chemical change. *International Journal of Science Education, 22,* 719–737.

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences, 4*(1), 39-103.

Kolodner, J. L. (2006). Case-Based Reasoning. In K. L. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 225-242). Cambridge: Cambridge University Press.

Krnel, D., Glazar, S. S., Watson, R. (2003). The development of the concept of "matter": A cross-age study of how children classify materials*. Science Education, 87,* 621-639.

Krnel, D., Watson, R., & Glazar, S. S. (1998). Survey of research related to the development of the concept of "matter." *International Journal of Science Education, 20*(3), 257 – 289.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* New York: Cambridge University Press.

Lawrence Hall of Science/LHS. (2005). *Chemical interactions teacher's guide (Full Option Science System).* Delta Education.

Lawrence Hall of Science/LHS. (2005). *Earth materials teacher's guide (Full Option Science*

*System).* Delta Education.

Lawrence Hall of Science/LHS. (2005). *Measurement teacher's guide (Full Option Science System).* Delta Education.

Lawrence Hall of Science/LHS. (2005). *Mixtures and solutions teacher's guide (Full Option Science System).* Delta Education.

Lawrence Hall of Science/LHS. (2005). *Solids and liquids teacher's guide (Full Option Science System).* Delta Education.

Lawrence Hall of Science/LHS. (2005). *Wood and paper teacher's guide (Full Option Science System).* Delta Education.

Lee, O., Eichinger, D.C., Anderson, C. W., Berkheimer, G. D. (1993). Changing middle school students' conceptions of matter and molecules*. Journal of Research in Science Teaching, 30*(3), 249-270.

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex Publishing.

Lewis, E. L., & Linn, M. C. (1994). Heat energy and temperature concepts of adolescents, adults, and experts: Implications for curricular improvements. *31, 6*(657-678).

Massachusetts Department of Education (2001). *Massachusetts science and technology/engineering curriculum framework*.

McCloskey, M. (1983). Intuitive physics. *Scientific American, 248*, 122-113.

Merriam, S. B. (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Josey-Bass Publishers.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis (2nd ed.).* Thousand Oaks, CA: SAGE.

Minogue, J. (2008). A lemon of a lesson. *Science and Children, 42*(1), 25-27.

Minstrell, J. (1989). Teaching science for understanding. In L. Resnick & L. Klopfer (Eds.), *ASCD 1989 Yearbook: Toward the thinking curriculum: Current cognitive research.* Washington, D.C.: Association for Supervision and Curriculum Development.

Nakhleh, M. B., & Samarapungavan, A. (1999). Elementary school children's beliefs about matter. *Journal of Research in Science Teaching, 36,* 777–805.

National Research Council/NRC. (1996). *National science education standards.* Washington, D.C.: National Academy Press.

National Science Resources Center/NSRC. (2005). *Balancing and weighing teacher's guide. Science and Technology for Children.* Kendall/Hunt Publishing.

National Science Resources Center/NSRC. (2005). *Changes teacher's guide. Science and Technology for Children.* Kendall/Hunt Publishing.

National Science Resources Center/NSRC. (2005). *Chemical tests teacher's guide. Science and Technology for Children.* Kendall/Hunt Publishing.

National Science Resources Center/NSRC. (2005). *Comparing and measuring teacher's guide. Science and Technology for Children.* Kendall/Hunt Publishing.

National Science Resources Center/NSRC. (2005). *Floating and sinking teacher's guide. Science and Technology for Children.* Kendall/Hunt Publishing.

National Science Resources Center/NSRC. (2005). *Food chemistry teacher's guide. Science and Technology for Children.* Kendall/Hunt Publishing.

National Science Resources Center/NSRC. (2005). *Rocks and minerals teacher's guide. Science and Technology for Children.* Kendall/Hunt Publishing.

National Science Resources Center/NSRC. (2005). *Solids and liquids teacher's guide. Science and Technology for Children.* Kendall/Hunt Publishing.

National Science Resources Center/NSRC. (2006). *Properties of matter teacher's guide. Science and Technology for Children.* Kendall/Hunt Publishing.

Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal, 28*(1), 117-153.

Novick S., & Nussbaum, J. (1981). Pupils' understanding of the particulate theory of matter: A cross-age study. *Science Education*, *65*, 187-196.

Olson, D.R. (1996). *The world on paper: the conceptual and cognitive implications of writing and reading.* Cambridge, Cambridge University Press.

Paik, S.-H., Cho, B.-K., & Go, Y.-M. (2007). Korean 4- to 11-year-old student conceptions of heat and temperature. *Journal of Research in Science Teaching, 44*(2), 284-302.

Pea, R.D. (1993). Practices of distributed intelligence and designs for education. In *Distributed cognitions: Psychological and educational considerations* (pp. 47-87). Cambridge: Cambridge University Press.

Piaget, J. (1929). *The child's conception of the world*. NY: Harcourt, Brace Jovanovich.

Piaget, J. (1930). *The child's conception of physical causality*. London: Kegan Paul.

Posner, G. J., Strike, K. A., Hewson, P. W, & Gertzog, W. A. . (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211-227.

Puntambekar, S., Kolodner, J. (2005). Toward implementing distributed scaffolding: Helping students learn science from design. *Journal of Research in Science Teaching, 42*(2), 185-217.

Renstrom, L., Andersson, B., & Marton, F. (1990). Students' conceptions of matter. Journal of Educational Psychology, 82(3), 555-569.

Rogoff, B. (1991). *Apprenticeship in thinking: Cognitive development in social context.* USA: Oxford University Press.

Roth, W.-M. (1996). Art and artifact of children's designing: A situated cognition perspective. *Journal of the Learning Sciences, 5*(2), 129-166.

Salomon, G., Perkins, D.N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher, 20*(3), 2-9.

Schaffer, J.P., Saxena, A., Antolovich, S.D., Sanders, T.H., Warner, S.B. (1999). *The Science and Design of Engineering Materials (2nd ed.).* Boston: WCB/McGraw-Hill.

Schwartz, J. (1996). Semantic aspects of quantity. Unpublished manuscript.

Shayer, M., & Wylam, H. (1981). The development of the concepts of heat and temperature in 10-13 year-olds. *Journal of Research in Science Teaching, 18*(5), 419-434.

Shepardson, D. P., & Britsch, S. J. (2001). The role of children's journals in elementary school science activities. *Journal of Research in Science Teaching, 38*(1), 43-69.

Skinner, B. F. (1958). Teaching machines: From the experimental study of learning come devices which arrange optimal conditions for self-instruction. *Science, 128*(3330), 969-977.

Smith, C.L. (2007). Bootstrapping processes in the development of students' commonsense matter theories: Using analogical mappings, thought experiments, and learning to measure to promote conceptual restructuring. *Cognition and Instruction, 25*(4), 337-398.

Smith, C., Carey, S., & Wiser, M. (1985). On differentiation: A case study of the development of the concept of size, weight, and density. *Cognition, 21,* 177 – 237.

Smith, C., Maclin, D., Grosslight, L., Davis, H. (1997). *Teaching for understanding: a study of students' preinstruction theories of matter and a comparison of the effectiveness of two approaches to teaching about matter and density.* Cognition and Instruction, *15*(3), 317-393.

Smith, C., Snir, J., & Grosslight, L. (1992). Using conceptual models to facilitate conceptual change: The case of weight/density differentiation. *Cognition and Instruction, 9,* 221–283.

Smith, C., Wiser, M., Anderson, C., & Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and atomic-molecular theory. *Measurement, 14*(1–2), 1–98.

Smith, J. P., diSessa, A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences, 3*(2), 115-163.

Snir, J., Smith, C., & Raz, G. (2003). Linking phenomena with competing underlying models: A software tool for introducing students to the particulate model of matter. *Science Education, 87,* 794–830.

Spelke, E. S. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In R. Gelman & S. Carey (Eds.), *The epigenesis of mind. Essays on biology and cognition* (pp. 133-169). Hillsdale, NJ: Lawrence Erlbaum.

Stavy, R. (1991). Children's ideas about matter. *School Science and Mathematics*, *91*(6), 240 – 244.

Vosniadou, S. (2002). On the nature of naïve physics. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 61-76). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Vygotsky, L. S. (1962). The development of scientific concepts in childhood. In *Thought and language* (pp. 82-118). Cambridge, MA: MIT Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher mental processes.* Cambridge, MA: Harvard University Press.

Wilson, M. (2005). *Constructing measures: An item response modeling approach.* Lawrence Erlbaum.

Wiser, M., Amin, T. (2001). "Is heat hot?" Inducing conceptual change by integrating everyday and scientific perspectives on thermal phenomena. *Learning and Instruction*, *11*, 331-355.

Wiser, M., & Smith, C. L. (2008). Learning and teaching about matter in grades K-8: When should the atomic-molecular theory be introduced? In S. Vosniadou (Ed.), *The International Handbook of Conceptual Change.*: Routledge.

## Appendix A. National and State Standards of Learning Related to the Properties of Objects and Materials

| **National Benchmarks for Science Literacy** (AAAS, 1993) | **National Science Education Standards** (NRC, 1996) | **Massachusetts Curriculum Frameworks** (Massachusetts Department of Education, 2001) |
|---|---|---|
| **4.D. 2nd Grade**<br>Objects can be described in terms of the materials they are made of (clay, cloth, paper, etc.) and their physical properties (color, size, shape, weight, texture, flexibility, etc.).<br><br>**4.D.5th Grade**<br>When a new material is made by combining two or more materials, it has properties that are different from the original materials. For that reason, a lot of different materials can be made from a small number of basic kinds of materials.<br><br>**4.E.5th Grade**<br>When warmer things are put with cooler ones, the warm ones lose heat and the cool ones gain it until they are all at the same temperature. A warmer object can warm a cooler one by contact or at a distance.<br>Some materials conduct heat much better than others. Poor conductors can reduce heat loss. | **Content Standard B: Properties of Objects and Materials (K-4)**<br>Objects have many observable properties, including size, weight, shape, color, temperature, and the ability to react with other substances. Those properties can be measured using tools, such as rulers, balances, and thermometers.<br>Objects are made of one or more materials, such as paper, wood, and metal. Objects can be described by the properties of the materials from which they are made, and those properties can be used to separate or sort a group of objects or materials. | **Grades K-2, Strand 2: Physical Science**<br>Sort objects by observable properties such as size, shape, color, weight, and texture.<br><br>**Grades 3-5, Strand 3: Physical Science**<br>Differentiate between properties of objects (e.g., size, shape, weight) and properties of materials (e.g., color, texture, hardness).<br>Recognize that light travels in a straight line until it strikes an object or travels from one medium to another, and that light can be reflected, refracted, and absorbed. |

## Appendix B. Overview of the *Design a Model House* Curriculum Module

### ***Design a Model House*** **Learning Objectives**

*By the end of this module, students will be able to:*

1) Describe materials by the following properties: strength, insulating ability, sound absorption ability, color, and reflectivity.
2) Describe objects by the following properties: weight, length, height, shape, strength, and stability.
3) Identify the materials from which objects are made.
4) Separate or sort a group of objects or materials based on their properties.
5) Measure the properties of materials and objects with both manual and electronic tools, such as rulers, weights, digital thermometers, and sound meters.
6) Identify the properties (e.g., strength, insulating ability, shape) that are most important for a specific design task.
7) Select materials and objects that exhibit the desired properties for a specific design task.
8) a) Define engineering design as the process of creating solutions to human problems through creativity and the application of math and science knowledge.
   b) List and explain the following steps of the engineering design process:
      i. Identifying a problem
      ii. Researching possible solutions
      iii. Picking the best solution
      iv. Building a prototype
      v. Testing the prototype
      vi. Repeating any steps needed to improve the design

### ***Design a Model House*** **Module Overview**

| Lesson Title | Lesson Overview | Lesson Learning Objectives<br>*Students will be able to:* |
|---|---|---|
| 1<br>Introduction to the Properties of Materials | Students will be introduced to the overall design challenge for the unit. Students will practice recognizing properties of materials and objects by sorting a set of objects into groups of similar items. | *Separate or sort a group of objects or materials based on their properties.*<br><br>*Define engineering design as the process of creating solutions to human problems through creativity and the application of math and science knowledge.* |
| 2<br>Can Support Columns Be Made of Clay? | Students will construct model-house support columns out of modeling clay and test their strength and stability. | *Describe objects by the properties of stability and strength.*<br><br>*Measure the properties of materials and objects with manual tools such as test weights.*<br><br>*Identify the properties that are most important for a specific design task and select materials and objects that exhibit those properties.* |

| | | |
|---|---|---|
| 3<br>Can Support Columns Be Made of LEGO Beams? | Students will construct model-house support columns out of LEGO beams and test their strength and stability. | *Describe objects by the properties of stability and strength.*<br><br>*Measure the properties of materials and objects with manual tools such as test weights.*<br><br>*Identify the properties that are most important for a specific design task and select materials and objects that exhibit those properties.* |
| 4<br>Testing the Insulation Ability of Wall Materials | Students will test the thermal insulation ability (using digital LEGO thermometers) of cardboard and Styrofoam as possible model-house wall materials. | *Describe materials by the property of thermal insulation ability.*<br><br>*Measure the properties of materials and objects with electronic tools such as digital thermometers.*<br><br>*Identify the properties that are most important for a specific design task and select materials and objects that exhibit those properties.* |
| 5<br>Testing the Sound Absorption Ability of Wall Materials | Using LEGO sound sensors, students will test the sound absorption (soundproofing) ability of cardboard and Styrofoam as possible model-house wall materials. | *Describe materials by the property of sound absorption ability.*<br><br>*Measure the properties of materials and objects with electronic tools such as sound level sensors.*<br><br>*Identify the properties that are most important for a specific design task and select materials and objects that exhibit those properties.* |
| 6<br>Testing the Stability of Roof Shapes | Students will design and build roof frames out of LEGO pieces and determine the strongest roof frame shape. | *Describe objects by the properties of shape and stability.*<br><br>*Identify the properties that are most important for a specific design task and select materials and objects that exhibit those properties.* |

| | | |
|---|---|---|
| 7<br>Testing Roof Surface Materials | Students will test the strength and thermal insulation ability (using digital LEGO thermometers) of white and black index cards and cardboard as possible roof surface materials. | *Describe materials by the properties of strength, thermal insulation ability, reflectivity, and color.*<br><br>*Recognize that a materials' color often determines its reflectivity and that a material's reflectivity often determines its thermal insulation ability.*<br><br>*Measure the properties of materials and objects with both manual and electronic tools, such as test weights and thermometers.*<br><br>*Identify the properties that are most important for a specific design task and select materials and objects that exhibit those properties.* |
| 8<br>Building the Model Houses | Using the materials that performed best in the tests, students will build model houses by measuring, cutting, constructing, and assembling the columns, walls, roof frames, and roof surfaces. | *Identify the properties that are most important for a specific design task and select materials and objects that exhibit those properties.*<br><br>*Measure the properties of materials and objects with manual tools such as rulers.*<br><br>*Describe objects by the properties of weight, length, height, and shape.* |
| 9<br>Model House Wrap-Up | Students will create design posters that explain how specific properties are important to their houses. They will participate in a whole-class discussion to review how t he houses meet the overarching engineering design requirements. | *Identify the properties of the materials from which objects are made.*<br><br>*Recognize that selecting the best material is a process in which engineers often engage.*<br><br>*Define engineering design as the process of creating solutions t o human problems through creativity and the application of math and science knowledge.*<br><br>*List and explain the steps of the engineering design process.* |

# Appendix C. Scoring Rubric for Interview Responses

Table C1
*Complete Version of the Scoring Rubric for Interview Responses*

**Code Sheet and Rubric for the *Select* Sub-Task**

STEPSTOOL DESIGN GOAL: Stepstool shouldn't collapse or break when someone stands on it.
HABITAT DESIGN GOAL: There's a heater inside the habitat, and the habitat walls should keep the inside of the habitat warm.

CODING/SCORING INSTRUCTIONS:

- Read the design goals above and the rubric below (please carefully review all examples).
- Assign one of these codes to each numbered rationale on the data sheet.
    1. Decide if rationale is a) **rep**eat of design goal, b) **ext**ensive property, c) **int**ensive property, d) appeal to life **exper**ience, e) appeal to constituent **mat**erial, or f) **irr**rel/none of the above. Code with bolded letters.
    2. If it's (b), (c), or (d), decide if it's (1) accurate for sample and (2) relevant to design goal. Add **acc/inacc** and **rel/irr** codes.
    3. If it's (e), decide if it's accurate for sample and most ideal choice. Add **acc/inacc** code.
- Attach points to each rationale according to scoring column.
- Add up total points. –
    - 2 for each acc-rel
    - 1 for each inacc rel
    - 0 for irr, rep, acc only, inacc-irr, or acc-irr

| Code | Explanation of Code | Example Rationale for Stepstool Task | Example Rationale for Habitat Task | Pts |
|---|---|---|---|---|
| irr | Gives reasoning that is not relevant to design goal, and does not fit into any categories below. | Bamboo is from trees. | You can cut a door into the foam. | 0 |
| rep | Repeats the stated criteria or states that selected material already meets design goal: stepstool shouldn't collapse; habitat wall should keep the habitat warm on the inside | It won't collapse;<br>It won't break when you stand on it. | It feels warm;<br>It makes it warm inside. | 0 |
| exper acc rel | Gives reasoning from personal life experience, in the form of "I have seen…" or "When I used it for…", etc., and cited experience is accurate for the material and relevant to the design goal | | When I used it for my LEGO house, and put my hand inside, my hand got sweaty. | 2 |
| exper inacc rel | Gives reasoning from personal life experience, and cited experience is relevant to design goal, but it is not accurate for the material. | I've seen bamboo support really heavy people. | | 1 |
| exper acc irr | Gives reasoning from personal life experience, and cited experience is accurate for the material, but it is not relevant to the design goal. | I know that furniture can be made out of PVC, and it looks good! | | 0 |

| | | | | |
|---|---|---|---|---|
| exper inacc irr | Gives reasoning from personal <u>life experience</u>, but cited experience is <u>not relevant</u> to design goal, and it is <u>not accurate</u> for the material. | | I've seen metal melt, so the mouse would melt. | 0 |
| int acc rel | Claims that chosen material possesses an <u>intensive</u> property, and the stated property is <u>accurate</u> for the material and <u>relevant</u> to the design goal | It is stronger; It's kind of strong; It's sort of strong;<br>It is sturdy;<br>It is solid;<br>It can hold a lot of weight;<br>It won't break when it has a lot of weight on it.<br>*(referring to either PVC or bamboo)* | It keeps heat in *(referring to foam)*;<br><br>It blocks coldness *(referring to foam)*<br><br>(*similar to saying: it is a good insulator)* | 2 |
| int inacc rel | Claims that chosen material possesses an <u>intensive</u> property, and the property is <u>relevant</u> to the design goal but <u>not accurate for the material</u> or <u>not accurately linked to property</u> | The PVC is weak, so I would use the wood instead.<br>*(PVC is not weak)* | It keeps heat in (*referring to metal);*<br>It will get hot when light shines on it *(referring to metal)*<br>It is melt-proof *(referring to foam)* | 1 |
| int acc irr | Claims that chosen material possesses an <u>intensive</u> property, and the property is <u>accurate</u> to the material but <u>not relevant to the design goal</u> | It is white *(referring to PVC)* | It is shiny (*referring to metal)*<br>It is strong, sturdy, won't break *(referring to metal)* | 0 |
| int inacc irr | Claims that chosen material possesses an <u>intensive</u> property, and the property is <u>not accurate</u> to the material <u>nor relevant to the design goal</u> | | | 0 |
| ext acc rel | Claims that chosen material possesses an <u>extensive</u> property, and the stated property is <u>accurate for the material</u> and <u>relevant to the design goal.</u> | | | 2 |
| ext inacc rel | Claims that chosen material possesses an <u>extensive</u> property, and the stated property is <u>relevant to the design goal</u> but <u>not accurate for the material</u> | It is stable;<br>It can stand up without falling *(referring to either foam or bamboo)*<br>*(stability only describes entire objects; wood as a material is neither stable nor unstable)* | It is lightweight *(referring to metal)* | 1 |

| | | | | |
|---|---|---|---|---|
| ext<br>acc irr | Claims that chosen material possesses an extensive property, and the stated property is accurate for the material, but it does not directly relate to the design goal<br><br>-OR-<br>Claims that chosen material possesses an extensive property, and the stated property is accurate for the material itself, but the stated property does not actually determine the desired intensive property for the design goal. | It is thicker<br>(*thickness alone does not determine strength)*<br><br>-OR-<br><br>It is heavier, so it is stronger<br>*(weight alone does not determine strength)* | It is taller<br>(*size alone does not determine insulation)*<br>-OR-<br><br>It is thicker<br>*(thickness alone does not determine insulation)* | 0 |
| ext<br>inacc irr | Claims that chosen material possesses an extensive property, but the stated property is not accurate for the material and not relevant to the design goal. | | | 0 |
| mat<br>acc | Explicitly names the material substance and claims that it is simply the most appropriate one for the job, and the chosen material is indeed the ideal one for the task | Because it is like hard plastic | It is foam; foam is good at keeping things warm | **0** |
| mat<br>inacc | Student names the material substance and claims that it is simply the most appropriate one for the job, but the chosen material is not the ideal one for the task, or the material is inaccurately identified. | Wood is the best thing for stools;<br>Because it is metal | Pet houses should be metal | 0 |

## Code Sheet and Rubric for the *Test* Sub-Task

CODING/SCORING INSTRUCTIONS:

- Assign one of these codes to each testing action described by the student.
- Attach points to each test according to scoring column.
- Add up total points.
- NOTE: In the stepstool task, the property in question is *strength* (students are asked to measure/test for *strength)*. In the mouse house task, the property in question is *insulation (*students are asked to measure/test for *insulation* ability).

| **Code** | **Explanation of Code** | **Example of Test/Measure for Stepstool Task** | **Example of Test/Measure for Habitat Task** | **Point Value** |
|---|---|---|---|---|
| Irrelevant<br>irr | Proposed action is not a test or measurement; it would not produce information about the sample. | You could paint the bamboo. | You could add tape to the house. | 0 |

| | | | | |
|---|---|---|---|---|
| Product Test + p+ | Proposed action would produce useful information that would accurately reveal the property in question (*strength* or *insulation*), but requires full construction of product. | Build stool and see if it stays together when a person stands on it.<br><br>Build stool and put a heavy weight on it. | Build two mouse houses and measure temperature inside and out;<br>Build mouse house and see if mouse shivers. | **1 for each** |
| Product Test – p- | Proposed action would produce useful information, but it would not accurately reveal the property in question (*strength* or *insulation*), and it requires full construction of product. | Build stool and see how much it weighs. | Build two houses and see which one the mouse looks better in. | **0** |
| Material Test + m+ | Proposed action is a physical test of the material only, and the test will produce useful information would accurately reveal (or nearly accurately reveal)[a] the property in question (*strength* or *insulation*). | Try to break it in half and see if it breaks; squeeze it and see if it squishes. | Shine a light on the material and measure the temperature on both sides after some time has passed. | **2 for each** |
| Material Test – m- | Proposed action is a physical test of the material only and would produce useful information, but it would not accurately reveal[b] the property in question (*strength* or *insulation*). | Measure length or weight of material to figure out strength. | Measure weight to figure out insulation.<br><br>Put in the sun, and see which one gets hotter faster. That's the one to pick.<br><br>Feel which one is warmer. | **0** |

**a What kind of material test WOULD accurately reveal insulation? (i.e., be coded "m+")**

- Expose one surface of the material to heat or cold (i.e.,, set up a temperature gradient across the material using a hot or cold stimulus, which may be a lamp, sunlight, heater, ice, etc.), and measure over time either: a) the air temperature near the other surface or b) the temperature of the other surface itself. The **lower** the change in temperature over time, the **lower** the conductivity, the **greater** the insulation, and the more appropriate the material for a house wall.

**b What kinds of material tests would NOT accurately reveal insulation? (i.e., be coded "m-")**

- Determine the materials' temperatures at just one moment in time. Assume incorrectly that higher *instantaneous* temperature implies greater insulation ability (e.g., feel the metal and see if it is warm; if it is, use it for the house). In fact, temperature taken at one just moment in time does not provide information about insulation over a length of time.
- Expose the material to heat or cold, and measure or feel temperature at several instances over a length of time, but interpret results incorrectly such that a **greater** *change* in temperature implies **greater** insulation ability. (e.g., put the materials in the sun and see which one gets hotter faster)

**<u>Code Sheet and Rubric for the *Describe* Sub-Task</u>**

CODING/SCORING INSTRUCTIONS:

- Assign one of the codes to each of the descriptors.
- Assign points to each descriptor according to scoring column. Note that only the first descriptor in each 1-point category scores a point (i.e., even if 5 intensive properties are listed, only 1 point is earned for the intensive property category).
- Add up points.
- NOTE: Maximum possible description score is 2 (1 for intensive property + 1 for extensive property).
- 

| **Code for Descriptors** | **Explanation and Example of Descriptor Code** | **Point Value** |
|---|---|---|
| INT | Description is an intensive property (does not change with quantity of material): color, stiffness, strength, surface texture, hardness, temperature | **1**<br>(for first intensive property only) |
| EXT | Description is an extensive property (changes with quantity of material): shape, size, weight | **1**<br>(for first extensive property only) |
| MAT | Description identifies material substance: plastic, bamboo, wood, foam, metal | 0 |
| FEAT | Description points out a surface feature specific to that object, such as a scratch, dent, printed word or number, etc. | 0 |
| FUNC | Description refers to function of the material: can use it to bang things, can cook on it | 0 |
| SIM | Description is a simile or metaphor: like a log, like a railing | 0 |
| DUP | Description is a duplicate (synonym) of previously stated descriptor | 0 |
| IRREL | Description would not help to identify item | 0 |
| INACC | Description is inaccurate | 0 |

# Appendix D. Scoring Rubric for Student Workbook Pages

Table D1
*Summary of the Scoring Rubric for Workbook Pages*

| Workbook Activities | Topics | Science Practices | Completeness Score | Science Practice Score |
|---|---|---|---|---|
| 14 workbook pages related to topics addressed in interview<br><br>Each page = 1 activity | *Each page labeled:*<br>• Strength<br>• Insulation<br>• Both | *Each page labeled:*<br>• Describing - Reasoning about properties<br>• Measuring - Representing experiments<br>• Selecting - Interpreting experimental results | *Each page given:*<br><br>**0 pt** for completing none to 1/3 page<br>**1 pt** for completing 1/3 to 2/3 of page<br>**2 pt** for completing 2/3 to all of page | *Each page given:*<br>**0 pt** for no proficiency at science practice<br>**1 pt** for partial proficiency<br>**2 pt** for full proficiency |

Table D2
*Complete Scoring Rubric for Workbook Pages*

| **Activity Page** | **Key Topic** | **Key Science Practice** | **Completeness Score** (Extent to which page's instructions were followed)<br>**Science Practice Score** (Proficiency at page's key science practice [selecting, measuring, or describing materials and objects]) |
|---|---|---|---|
| 2.1 – columns prompt | Strength | Selecting | 0 = Neither drawing nor writing attempted<br>1 = Incomplete drawing without written explanation OR incomplete written explanation<br>2 = Complete drawing OR complete written explanation |
| | | | 0 = no consideration of vertical columns<br>1 = indicates consideration of vertical columns, but makes no claim about specific characteristics<br>2 = makes claim about characteristic(s) of support columns, and characteristic(s) is related to stability or strength |
| 2.3 – column test results | Strength | Testing | 0 = Attempted from none up to 1/3 of page<br>1 = Attempted from 1/3 up to 2/3 of page<br>2 = Attempted from 2/3 up to all of page |
| | | | 0 = from drawings/labels, reader can not decipher what happened when clay was flicked and squeezed<br>1 = at least 2 drawings/labels make it clear what happened when |

| | | | |
|---|---|---|---|
| | | | clay was flicked and squeezed<br>2 = at least 4 drawings/labels make it clear what happened when clay was flicked and squeezed |
| 3.1 – better than clay prompt | Strength | Selecting | 0 = No attempt at listing or explaining materials<br>1 = At least one material listed but no complete explanations<br>2 = At least one material listed on each side of page and at least one explanation on each side of page |
| | | | 0 = rationales for material choices are not relevant to strength of stability<br>1 = some rationale for material choices are relevant to strength or stability<br>2 = nearly all rationale for material choices are relevant to strength or stability |
| 3.3 – column summary chart | Strength | Describing | 0 = Attempted from none up to 1/3 of page<br>1 = Attempted from 1/3 up to 2/3 of page<br>2 = Attempted from 2/3 up to all of page |
| | | | 0 = circled choices are not consistent with data table, nor is rationale relevant to strength or stability<br>1 = either circled choices are consistent with data table, or rationale is relevant to strength or stability, but not both<br>2 = circled choices are consistent interpretations of data table, and rationale is relevant to strength or stability |
| 4.1 – ideas for insulation tests | Insulation | Selecting | 0 = No attempt at describing tests<br>1 = Attempted one test description<br>2 = Attempted two test descriptions |
| | | | 0 = suggested test(s) is not described clearly enough to be conducted by the reader<br>1 = suggested test(s) described clearly enough to be conducted by reader, but not relevant to insulation<br>2 = suggested test(s) could be carried out and is relevant to insulation |
| 4.2 – jackets and coolers prompts | Insulation | Selecting | 0 = No attempt at either of first two questions<br>1 = Attempted at least one of first two questions<br>2 = Attempted both of first two questions |
| | | | 0 = neither answer makes claim about the characteristics of materials that lead to temperature regulation<br>1 = one answer makes claim about the characteristics of materials that lead to temperature regulation<br>2 = both answers make claim about the characteristics of materials that lead to temperature regulation |
| 4.3 – wall insulation test results | Insulation | Testing | 0 = Attempted from none up to 1/3 of data table<br>1 = Attempted from 1/3 to 2/3 of data table<br>2 = Attempted from 2/3 up to all of data table |
| | | | 0 = computation work is not shown or is incorrect, or data are listed in wrong places<br>1 = partially correct, but some mistakes in computation or missing data<br>2 = computation work is shown and is correct |
| 5.3 – wall summary chart; choosing wall material | Insulation | Describing | 0 = Attempted from none up to 1/3 of page<br>1 = Attempted from 1/3 up to 2/3 of page<br>2 = Attempted from 2/3 up to all of page |
| | | | 0 = circled choices are not consistent with data table, nor is rationale relevant to insulation or sound absorption<br>1 = either circled choices are consistent interpretations of data table, or rationale is relevant to insulation or sound |

| | | | |
|---|---|---|---|
| | | | absorption, but not both<br>2 = circled choices are consistent interpretations of data table, and rationale is relevant to insulation or sound absorption |
| 6.1 – roof shape prompt | Strength | Selecting | 0 = Neither drawing nor writing attempted<br>1 = Incomplete drawing without written explanation OR incomplete written explanation<br>2 = Complete drawing OR complete written explanation<br>1 = indicates consideration of roof, but makes no claim about specific characteristics<br>2 = makes claim about characteristic(s) of roof, and characteristic(s) is related to stability or surface texture |
| 6.2 – roof stability test results | Strength | Testing | 0 = Attempted from none up to 1/3 of page<br>1 = Attempted from 1/3 up to 2/3 of page<br>2 = Attempted from 2/3 up to all of page<br>0 = from drawings/labels, reader can not decipher what shape was constructed nor what happened when test was conducted<br>1= either shape can be deciphered, or test results can be determined, but not both<br>2 = drawings/labels make it clear what shape was constructed and what happened when test was conducted |
| 7.1A – color-temperature prompt | Insulation | Selecting | 0 = No attempt to respond to prompt<br>1/2 = Makes claim, but no explanation<br>1 = Attempts explanation for claim<br>0 = explanation does not relate material characteristic(s) to temperature or insulation<br>1 = explanation relates material characteristic(s) to temperature or insulation |
| 7.1B – which-is-stronger prompt | Strength | Selecting | 0 = No attempt to respond to prompt<br>1/2 = Makes claim, but no explanation<br>1 = Attempts explanation for claim<br>0 = explanation does not relate material characteristic(s) to strength<br>1 = explanation relates material characteristic(s) to strength |
| 7.2 – roof insulation test results | Insulation | Describing | 0 = Attempted from none up to 1/3 of page<br>1 = Attempted from 1/3 up to 2/3 of page<br>2 = Attempted from 2/3 up to all of page<br>0 = circled answers are not consistent with data tables<br>1 = some circled answers are consistent; some are inconsistent<br>2 = all circled answers are consistent interpretations of data table |
| 7.3 – roof summary chart; choosing roof material | Both | Describing | 0 = Attempted from none up to 1/3 of page<br>1 = Attempted from 1/3 up to 2/3 of page<br>2 = Attempted from 2/3 up to all of page<br>0 = circled choices are not consistent with data table, nor is rationale relevant to insulation or strength<br>1 = either circled choices are consistent interpretations of data table, or rationale is relevant to insulation or strength, but not both<br>2 = circled choices are consistent interpretations of data table, and rationale is relevant to insulation or strength |
| 8.1 – choosing and justifying wall and roof materials | Both | Selecting | 0 = Listed no choices or only one choice<br>1 = Stated only one choice with rationale, or stated both choices but only one rationale<br>2 = State both choices and rationales for both<br>0 = neither rationale mentions actual characteristics of material<br>1 = at least one rationale mentions actual characteristic(s) of |

| | | | |
|---|---|---|---|
| | | | material, but not relevant to house design goals<br>2 = at least one rationale mentions material characteristic(s) relevant to strength, stability, insulation, or sound absorption |
| 9.1 – explaining test for and importance of one teacher-selected property | Varies (property selected by teacher) | Testing, Selecting | (*No points for simply writing house name or property word*)<br>0 = No attempt at either "test for" or "important" response<br>1 = Attempt at either "test for" or "important" response<br>2 = Attempted both "test for" and "important" responses<br>(*Do not factor house name or house drawing into score*) |
| | | | 0 = neither A nor B is true:<br>1 = either A or B is true:<br>2 = both A and B are true:<br><br>A: "test for" response describes an action that could test for the property<br>B: "important" response relates material or object characteristics to a house design goal |

# Appendix E. List of All Descriptors Generated by Students

Preinstruction Interviews

Bamboo

1. Smooth on inside INT
2. Greenish INT
3. Brown INT
4. Won't break INT
5. Would easily break INT
6. Weak INT
7. Yellow on inside INT
8. Hard INT
9. Short - EXT
10. Skinny – EXT
11. Lightweight EXT
12. Hollow EXT
13. Long EXT
14. Shaped like a cylinder EXT
15. Rough at the top EXT
16. Wood MAT
17. Bamboo MAT
18. Line on it FEA
19. No words on it FEA
20. Has a dent FEA
21. Black at bottom FEA
22. Has parts that poke FEA
23. Two holes FEA
24. Like a log SIM
25. Like tree bark SIM

PVC

1. White – INT
2. Rough at the top, softer in the middle INT
3. Smooth INT
4. Strong INT
5. Thick INT
6. Hard INT
7. Won't break - INT
8. Flat on outside EXT
9. Round EXT
10. Stable EXT
11. Can see through it EXT
12. Hollow EXT
13. Heavier than wood - EXT
14. Shaped like a cylinder - EXT
15. Long - EXT
16. Fatter than wood - EXT
17. Taller than wood (gestured larger diameter) EXT
18. Metal MAT
19. Words on it FEA
20. Holes at the ends FEA

21. Numbers on it FEA
22. Like a railing SIM
23. Looks like sink pipe SIM
24. Commonly used FUNC
25. Makes noise when knocked - FUNC

<u>Aluminum Sheeting</u>

1. Smooth on both sides - INT
2. Reflective – INT
3. Cold - INT
4. Soft in middle, rough on sides - INT
5. Strong - INT
6. Silver – INT
7. Attracts sun and hot areas - INT
8. Hard – INT
9. Knock-able - INT
10. Gray – INT
11. Shines when held to light – INT
12. Stable – EXT
13. Bent in corners and back - EXT
14. Kinda heavy – EXT; Heavy - EXT
15. Bigger than plastic - EXT
16. Long - EXT
17. Shaped like a rectangle - EXT
18. Leaves fingerprints - EXT
19. Thick – EXT; Thicker - EXT
20. Flat – EXT; Mostly flat - EXT
21. Thin - EXT
22. Looks like a tray – SIM
23. Top looks like a skateboard - SIM
24. Lines on back – FEA
25. Sayings on back – FEA
26. Scratched - FEA
27. Can cook on it - FUNC
28. Makes stuff hot – FUNC
29. Metal – MAT
30. Has indentations - FEA
31. Curved part – FEA; One bent side - FEA
32. Dents on back – FEA; Top has dents - FEA
33. Looks like a journal - SIM
34. Aluminum foil on it - MAT
35. Has pin holes in it - FEA
36. Can cut edges off to make it fit in house – IRREL

<u>Foam Sheeting</u>

1. White - INT
2. Soft - INT
3. Warm at bottom, not as warm at top - INT
4. Weak – INT
5. Breakable – INT
6. Has a fine temperature - INT
7. Hard –INT
8. Bends more easily than metal – INT
9. Light – EXT
10. Rough on edges - EXT

11. Front side is smooth - EXT
12. Shaped like a rectangle – EXT
13. Long - EXT
14. Smaller than metal – EXT
15. Thick - EXT
16. Thin – EXT
17. Flat, mostly – EXT
18. Plastic – MAT
19. Foam - MAT
20. Floats when you make it fall – FUNC
21. Can use it as a fan - FUNC
22. Makes noise when you wiggle it – FUNC
23. Dots on it – FEA
24. Round bumps on edge – FEA
25. Soft edges (referring to curved edges) – FEA
26. Has a square - FEA
27. Has words on it – FEA
28. Warm air goes through it - INNAC
29. Won't burn down - INNAC
30. Can help animals – IRREL

## Postinstruction Interviews

Bamboo

1. Brownish INT
2. Green – INT
3. Hard – INT
4. Would snap INT
5. Yellow INT
6. Strong INT
7. Can chip off - EXT
8. Round EXT
9. Fits inside PVC EXT
10. Partially hollow EXT
11. Can't see through it - EXT
12. Long - EXT
13. Cylinder - EXT
14. Skinny EXT
15. Short EXT
16. Thin EXT
17. Wood MAT
18. Bamboo MAT
19. Rough on edges FEA
20. Lines on it FEA
21. Dent on bottom FEA
22. No words on it FEA
23. Marks on it FEA
24. A little broken FEA
25. Things sticking out FEA
26. Rolls FUNC
27. Like a log SIM
28. Like a tube SIM
29. Like tree bark SIM

PVC

1. White – INT

2. Doesn't break when bent - INT
3. Strong - INT
4. Hard – INT
5. Soft INT
6. Shiny INT
7. Smooth all around INT
8. Like a circle EXT
9. Two flat surfaces EXT
10. Skinny EXT
11. Open on the end EXT
12. Clean EXT
13. Thick EXT
14. Stable EXT
15. Can see through it EXT
16. Long - EXT
17. Cylinder - EXT
18. Has words on it FEA
19. Two big holes FEA
20. Circles at ends FEA
21. Nothing sticks out FEA
22. Like a log SIM
23. Like a telescope SIM
24. Like a pirate thing SIM
25. Rolls FUNC
26. Could hurt someone FUNC

<u>Aluminum Sheeting</u>

1. Silver – INT
2. Bounces light like a CD – INT
3. Cold – INT
4. Gray – INT
5. Can see yourself it in – INT
6. Doesn't bend much – INT
7. Cannot break - INT
8. Shiny – INT; Shiny – INT
9. More hard - INT
10. More smooth - INT
11. Has four corners - EXT
12. Has two different surfaces - EXT
13. Rectangular prism - EXT
14. Flat – EXT; Flat - EXT
15. Leaves fingerprints – EXT
16. Clean - EXT
17. Like tinfoil around edges – MAT
18. Metal – MAT
19. Looks like a tray – SIM
20. Like a ramp at the top - SIM
21. From cafeteria - FUNC
22. Can be a fan – FUNC
23. Can cook on it – FUNC
24. Soft in one place – FEA
25. Has lines on back – FEA
26. Says words on back – FEA
27. Has scratches – FEA

28. Bent a little bit in one place – FEA
29. Flat flap - FEA
30. Dents – FEA
31. Has a nail thing in the edge – FEA
32. Easy to melt - INACC

Foam Sheeting

1. White – INT
2. Hard – INT
3. Soft in middle - INT
4. Warm – INT
5. Breakable - INT
6. Kind of smooth – INT
7. Rough along side - EXT
8. Light - EXT
9. Rectangular prism shape - EXT
10. Flat – EXT; Flat – EXT
11. Skinny - EXT
12. Thick - EXT
13. Thin on edge - EXT
14. Big – EXT
15. Foam – MAT
16. Has bumps on side – FEA
17. Has scratches – FEA
18. Has edges - FEA
19. Has dots - FEA
20. Has lump - FEA
21. Has square on it – FEA
22. Has numbers and words – FEA
23. Dented – FEA
24. Can be used as fan - FUNC
25. Float - FUNC
26. From the cafeteria – FUNC
27. Is a material - IRREL

Appendix F. Mean Scores on Selected Workbook Segments

| Measure | Mean | (SD) | Possible |
|---|---|---|---|
| Total Workbook Score | 37 | (11) | 60 |
| Workbook Strength Tasks Score | 23 | (4.9) | 38 |
| Workbook Insulation Tasks Score | 17 | (8.4) | 34 |
| Workbook "Select" Score | 29 | (8.6) | 48 |
| Workbook "Measure/Test" Score | 11 | (3.1) | 16 |
| Workbook "Describe" Score | 11 | (3.2) | 16 |