ASSESSING STUDENTS' EVOLVING UNDERSTANDINGS ABOUT MATTER[1]

Summary

The Inquiry Project has been conducting a longitudinal investigation of students' thinking from ages 8-11 (Grades 3-5) with special emphasis on their evolving concepts of material, weight, volume, density and state of matter. In the project, we have been focusing on two general advances that will lay foundations for the atomic theory of matter: (1) a shift from *perception-centered* to *model-mediated thinking* about physical phenomena and (2) the development of *quantitative reasoning—reasoning about physical quantities in ways that highlight mathematical structure. Each of these advances benefits from students' growing familiarity with issues related to measurement and to drawing inferences from evidence.*

Here we illustrate how the assessment of student learning and conceptual change was crafted to suit our theoretical concerns and premises about the progression of learning. The clinical interviews at Grades 3 and 4, along with classroom observations, have helped us refine our models of student understanding and the roles of instruction in this development.

David W. Carraher, TERC
Carol L. Smith, University of Massachusetts at Boston
Marianne Wiser, Clark University
Analúcia D. Schliemann, Tufts University
Gabrielle Cayton-Hodges, Tufts University

June 10, 2009

[1] Paper prepared for the "Learning Progressions in Science" Conference, Iowa City, Iowa, June 24-26, 2009. The research was supported by MSP Grant #0628245 of the National Science Foundation's Division of Research on Learning in Formal and Informal Settings (DRL), in the Directorate of Human Resources. Special thanks to our partners in the Inquiry project, the Mason School (Roxbury, MA) and Forestdale School (Malden, MA).

What Kinds of Assessments Are Well Suited For a Learning Progression about Matter?

Some Initial Thoughts about Assessments in Learning Progressions

A Learning Progression project seeks to provide a compelling account of how students' knowledge in a given domain or conceptual field (Vergnaud, 1996) gradually develops (National Research Council, 2007). As in cognitive development research, it assumes the existence of overarching structures (schemes, models, themes) that cut across and underlie specific topics, tasks and classroom activities. However, it acknowledges that culture, and schooling in particular, play major roles in learning and cognitive development; for instance, many science concepts, techniques, and representational tools (alphabetic script, algebraic notation, graphs) are explicitly introduced in instruction. Researchers track student learning while attempting to understand the roles of certain characteristics of instruction (the scope, structure, and sequence of content; investigative practices; patterns of dialogue; cultivation of habits of mind). The general enterprise is predicated on the notion there are diverse ways of teaching and courses of learning in a given domain.

Research in Learning Progressions is far more than a quest to answer the question, "what works?". It is theoretical to the extent that it offers, explicitly and implicitly, accounts as to how student knowledge arises, develops, and adapts through interaction with other forms of knowledge, those other forms including knowledge (a) as understood from within the fields of science and mathematics (b) as conceptualized by teachers and (c) as presented in curricula.

There are several reasons why objective, "close-ended", paper-and-pencil measures of scientific understanding may be of limited use in learning progressions research with young students:

- Assessment tasks need to capture not only students' proficiency and achievement, but also their conceptualizations, beliefs, mental models, procedures, and ontological and epistemological assumptions. There is little reason to expect that researchers can anticipate the variety of views that students will hold or methods they would use if left to their own devices.[2]
- It is often desirable to assess conceptual understanding before students have been introduced to specialized vocabulary, mathematical formulas, units of measure, tabular and graphical representations, etc.
- Assessments must be especially attentive not only to desired end-states, but also to intermediate states of understanding where knowledge is largely unconsolidated.

[2] Where exhaustive research has already been undertaken, standard measures may of course be of use.

- We need to look beyond the vocabulary of the child and to the underlying concepts, which can often be missed in paper-and-pencil measures due to vocabulary constraints and misuse.

The developmental nature of Learning Progressions work recommends longitudinal over cross-sectional designs, but repeated measures approaches pose special challenges:

- Younger students may find that some tasks are too challenging or presume a familiarity with terminology and representations introduced only later in school; older students may find some questions geared to young students trivial or obvious
- Students may benefit from the fact that they had answered the same questions in interviews conducted earlier.

Such considerations led us, early on, to adopt clinical interview methods within a longitudinal design. As we now move to the particular focus of our work, we'll continue to emphasize the interplay between theoretical concerns and decisions regarding assessment methods.

The Inquiry Project

The Inquiry Project has been attempting to understand and promote learning among students in Grades 3 through 5 that will provide a solid foundation for later understanding about the atomic molecular theory of matter (Smith, Wiser, Anderson, & Krajcik, 2006; Wiser & Smith, 2008). Our work has been guided by the view that many of the difficulties students later have with the atomic molecular theory stem from limitations in their *macroscopic understandings* of matter and that developing these macroscopic understandings involves deep and broad conceptual change or reconceptualizations. In the present view, concepts consist of (i) certain *invariants*—objects, properties, and relations—relevant to (ii) certain sets of *situations* and conditions and expressed through (iii) conventional and personal *systems of symbolic representation*. Significant changes in one or more these components are indicators of reconceptualizations.

We propose that young school children can build a sound macroscopic understanding of matter and different materials through appropriate instruction in which the physical quantities of weight, volume, and density are differentiated and inter-related and for which students recognize solids, liquids, and gases as distinct forms of matter.

Such progress is viewed as resting on epistemological habits of mind and mathematical structures that are not content free (Carraher, 1996, 2001; Carraher & Schliemann, 1991; Schliemann & Carraher, 1992; Schliemann & Nunes, 1990; Vergnaud, 1996).

We especially call attention to two overarching and inter-related advances in students' thinking:

- A gradual shift from (a) *perception-centered thinking*, that is, understanding and explanation closely tied to perceptual judgment and appearances, to (b) *model-mediated thinking*, informed by views about matter and drawing upon a set of increasingly advanced, inter-related concepts and scientific habits of mind.

- The development of *quantitative reasoning and understanding of measurement* that students can use to make predictions, interpret, and explain relationships among physical quantities. Additive scalar quantities such as volume, mass, and length are gradually re-conceptualized as dense linear dimensions (Whitney, 1968a, 1968b) for which a metric supports additive comparisons (differences, e.g. "three cc more") and multiplicative comparisons (ratios, e.g. "three times the volume of") in terms of measures and their respective quanta. Students also learn to describe relationships between dimensions of different natures. Graphs and other representational systems will depict functional relations in a coordinate space.

Overall, in the Inquiry Project curriculum, five core concepts (matter, material, weight, volume, and density) are present in each unit, although some are fore grounded and taught explicitly while others are implicit. For example, weight and material are presented front and center in grade 3, while volume is introduced only briefly at the end of the unit; density and matter remain implicit.[3] One of the recurring props for hands-on activities is a set of "density cubes:" same size cubes made of different materials with different densities (pine, popular, oak, acrylic, aluminum, steel, copper). In grade 3, they are used to compare the merits of hefting vs. using a scale to measure their weights, and to systematize the set of properties that can be used to compare materials (one of them being the weight of the cubes). Neither the concept nor the word *density* is used but students' attention is drawn to the cubes all being the same size, and the teacher occasionally uses the term "heavy for its size." Thus density starts being scaffolded in third grade, linguistically, as a placeholder, and by being embodied in the set of equal size cubes. We see this kind of "implicit scaffold" as part of helping students to be ready for a new idea later on.

In the 4th grade, the focus shifts to a variety of Earth Materials, which includes liquids, granular materials, and solids. Volume and material are front and center, while weight recedes, as students investigate water displacement and discover it is volume, not weight that is relevant. Students explore multiple senses of volume and also begin to distinguish heavy and heavy for size as they compare samples of different materials that are all the same weight and observe their very different volumes.

The five concepts will once again be present in the fifth grade, with different saliences--greater foregrounding of matter (as students consider gases and the process of phase change) and weight (as they consider conservation of weight across phase change).

As already pointed out, to assess whether and how students make "progress" as they work through the Inquiry curriculum requires a longitudinal rather than cross-sectional design. Further, control groups are needed to distinguish the effects of differential

[3] Material and weight are fore grounded in grade 3 because they are already distinct concepts for children (and distinctively symbolized with words); hence these concepts provide valuable entry points for further learning and investigation. Because they are densely connected with other concepts, "moving" these concepts forward contributes to moving other concepts along. In contrast, other important concepts in the expert network, such as volume, density, and matter are not yet distinct concepts for young children, nor symbolized with distinct words; hence they become fore grounded only in later units.

instruction and development. A central hypothesis is that traditional instruction only leads to limited progress.

The clinical interview format we have chosen allows for staging problems for students to work on, to clarify the meaning of the child's responses, and is well-suited to age levels and the variety of understanding we sought to assess: explanation, prediction, inference, comparative judgment, devising measures, and reasoning (tacit as well as explicit).

We videotaped the interviews knowing some of the tasks would be too complex to reliably code in real time. Multiple raters could later carefully judge the children's responses while using coding procedures devised after the data collections.

Designing Tasks for Assessing Advances in Learning Over Long Periods

Research Design

Each student receives the same two-hour interview at multiple points in time, although sub-tasks are structured so that the questioning can be adjusted somewhat to the thinking of the student. Treatment students (those who received the Inquiry science curriculum for nine weeks in each of grades 3, 4, and 5) are interviewed at four moments over two and one-half years. Control students (students from the same school who had received the standard science classroom instruction in grades 3-5) are interviewed at three moments. To date we have completed just over 200 interviews collected at the following moments: before and after grade 3 for Treatment students, after grade 3 and after grade 4 for Control students.

The repeated measures design raises the possibility that performance may improve due to repeated exposure to the same questions. However this does not systematically benefit one of the student groups. Moreover, we avoided giving students feedback on the "correctness" of answers and, in a few cases (e.g., predictions about water displacement), we did not allow the student to test their predictions.

Interview Design

The interview consists of 10-multi-part tasks administered in two one-hour sessions. Various concepts were assessed (matter, amount of material, amount of matter, weight, volume, density, number, fraction, division, ratio, proportion), in multiple ways.

Each task focuses on different aspects of conceptual understanding and was carefully designed to be open to a certain degree of interpretation by students, so that we could ascertain what properties or relations were salient and important to the child in a given situation. This allowed our analysis to capture the child's meaning and to relate children's patterns of reasoning across different tasks. Each task also entails multiple concepts and involves different levels of difficulty. This allows us to examine whether there are systematic patterns in how new understandings emerge. Within a given task, questioning could also end when a student had reached her level of difficulty.

The 10 tasks assessed children's understanding of (1) the divisibility and granularity of clay; (2) judgments about (and measurements of) length, area and volume; (3) weight, size, heaviness of material, and density; (4) volume (water displacement); (5)

conservation of properties of clay after reshaping; (6) block rearrangements and measurement of weight; (7) sorting by matter/not matter as well as ideas about atoms and molecules (if children had heard of them); (8) granularity of number and length (9) transformations (beyond mechanical deformations) of materials; and (10) sweetness (ratio and proportion).

Illustrative Examples of Our Assessments

In this section, we describe four of our ten assessment tasks in detail. For each task, we discuss the important theoretical questions it was designed to address, give an overview of key findings based on the Inquiry Project data to date (i.e., data from our comparative longitudinal study), and raise questions we are still wondering about as our work unfolds.

1- The Properties of Tiny (Visible and Invisible) Things

Rationale for Particular Design of Questions

How do children think and reason about the properties of tiny things—things that are “infinitesimally small”? We use the term “infinitesimally small” in the sense of beneath the threshold of human sensory detection; this is somewhat different from the mathematical notion of “as close to zero as one wants to go.” Children’s thinking about these issues is potentially revealing about their thinking about a host of important issues. Do they have any (theoretical) beliefs about matter or weight that constrain their thinking? Are they willing to assume that there can be things outside their range of perceptual experience? If so, what do they think those things are like? Are they willing to grant that those things have any weight? Can they use the existence of those things to explain other things? What do they know about measurement? Can they use understanding of measures to extrapolate to scales beyond their immediate sensory experience? Moving children from perception-centered to model-mediated thinking is central to a learning progression headed toward the atomic-molecular theory of matter.

We assume children initially have no clear way of thinking about such tiny things and certainly do not grant special ontological status to these small things (i.e., using their properties to explain the properties of macroscopic objects); after all, their perceptual systems were designed to give them immediate information about intermediate size macroscopic objects. Further, we assume that from an epistemological point of view, children initially (quite reasonably) trust their senses to give them reliable information about the world.

However, as children engage in active theory building about the world and develop understanding of new tools, they have the opportunity to significantly change their thinking about these issues. Further their willingness to make inferences about physical properties of objects in the absence of (or in contradiction to) sensory information appears to constitute an important foundation for their burgeoning ability to develop models of unseen processes regarding macroscopic phenomena; it also sets the stage for models about phenomena occurring at orders of magnitude under human awareness.

Of course we could not ask children directly about limits and infinitesimals. Instead, the first part of Task 1 focuses on how children reason about the properties of: (a) a large ball-shaped lump of clay; (b) a tiny, yet still visible speck of clay; and (c) a piece of clay said to be so tiny that it could not be seen. At issue is whether children think that even tiny pieces of clay take up space, have weight, and maintain their existence and these properties with repeated division. A central aspect of our proposed learning progression is helping children construct a mental model of matter in which amount of stuff, size, and weight are explicitly inter-related. In such a model, children can mentally decompose and recompose an object into component parts that retain their identity and (some) properties (e.g., each piece has weight and takes up space) and imagine that the weight (or volume) of the whole is equal to the sum of the weights of the parts.

Children are first handed the large ball of clay (about the size of a squash ball) and asked: *Do you think this weighs something?* We chose a piece that has considerable weight (120 grams), so that children would readily agree that it does to set the stage for the next questions. Children then watch as a tiny piece (about 2 mm in diameter) is broken off. The piece is handed to them, so they have the opportunity to both see and feel the piece, and they are asked: *Do you think this weighs a tiny bit or nothing at all? How do you know? Do you think this piece takes up any space? How do you know?*

Children are next asked about the *possibility* of an invisible piece of clay: *Could there ever be a piece of clay so tiny that you couldn't see it?* If students agree that there could be such a piece, they are ten asked: *Would that tiny piece, so small you can't see it, take up any space? How do you know? What that tiny (invisible) piece have any weight? How do you know?*

Finally, all students are engaged in a thought experiment about the repeated divisibility of clay. Starting with the tiny 2 mm speck of clay, they are asked to imagine that we had special tools that allowed us to cut in half again and again. *If we keep cutting the piece in half, making it smaller and smaller, would we ever get to a point where there was nothing left, or would something always be there? Why do you think that?*

We have found these very simple questions are a powerful means of eliciting some of students most basic presuppositions about matter, weight, taking up space, and amounts with no technical vocabulary. Children can engage with these questions, and tell us quite interesting things.

Key findings about the Properties of Tiny Things

In keeping with past research, we found that many children do not initially think that tiny *visible* things take up any space or have weight; understanding that tiny things take up space is also easier than understanding they have weight. Figure 1 shows that at the beginning of grade 3, about 50% of the students in our sample said that the speck of clay weighed “nothing at all” rather than a tiny bit and denied that the speck “took up any space”. Another 40% of the students thought it took up space, but weighed nothing at all. Only 10% said that the speck both took up space and weighed a tiny bit.

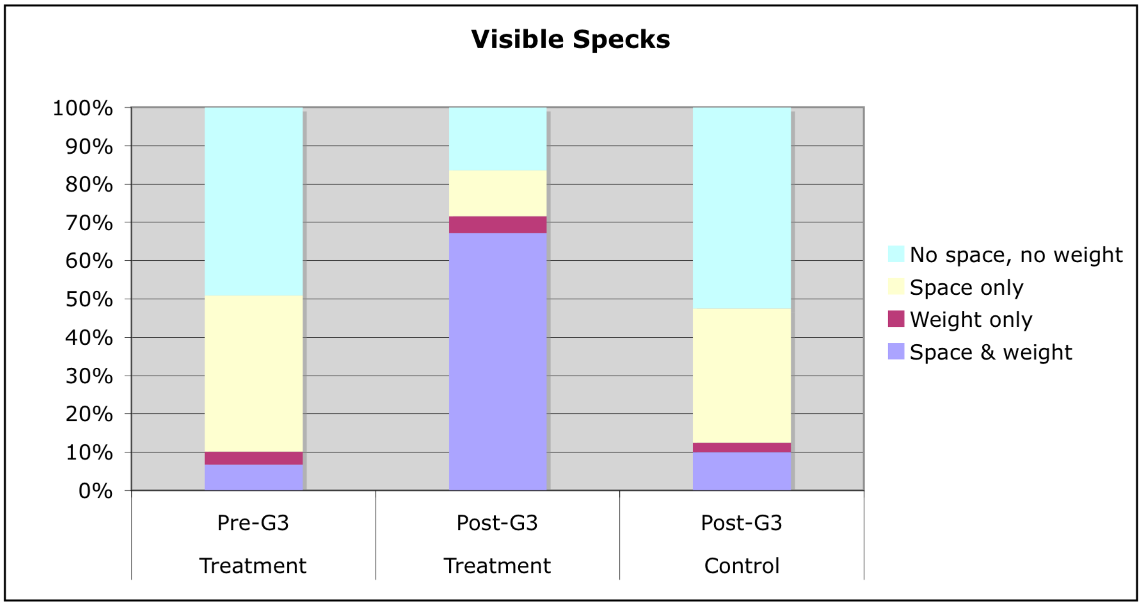


Figure 1: Comparison of the four patterns of judgment about the properties of tiny Visible Specks for the treatment students (grade 3 pre and post) and control students (interviewed at the end of grade 3). The four patterns are the visible speck: (a) takes up no space and has no weight; (b) takes up space, but has no weight; (c) has weight, but takes up no space; and (d) both takes up space and has weight.

The fact that these problems arise for *visible pieces* is instructive, as it suggests problems emerge well before microscopic scale. Indeed, the pattern of responding about invisible pieces is very similar (see Figure 2): for young third graders: 60% either denied the existence of an invisible piece or said the invisible piece took up no space and had no weight, 25% said it took up space and had no weight, and 7% said it took up space and had weight.

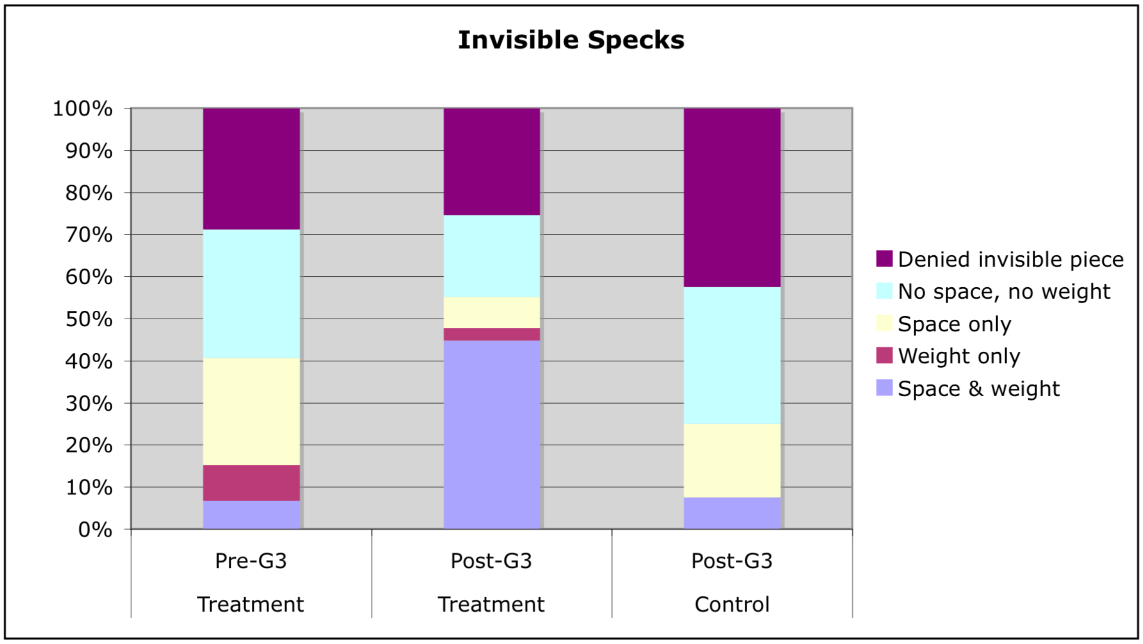


Figure 2. Comparison of the five patterns of judgment on the Invisible Specks questions for the treatment students (grade 3 pre and post) and control students (interviewed at the end of grade 3). The five patterns are—(a) denies that there can be a piece too small to see; or agrees there can be a tiny invisible piece and judges that it (b) takes up no space and has no weight; (c) takes up space, but has no weight; (d) has weight, but takes up no space; and (e) both takes up space and has weight.

In general, we found that the majority of young third graders did not have as much difficulty imagining that there might exist pieces too small to see as imagining that they would take up space and have weight. Seventy per cent of the pre-treatment 3rd graders were willing to grant that there could be pieces of clay too small to see, and about half thought you could keep cutting pieces smaller and smaller without getting to nothing (see Figure 3). The greater problem was that they just didn't think these very small pieces weighed anything (and many also denied they took up any space, although again this was easier). Of course, some students deny there would be invisible pieces (e.g., arguing that if you can't see it it's not there, that you can always see things like clay that are colorful, or that they have very good eyesight), but most of these students had already had difficulty with the questions about the properties of visible stuff.

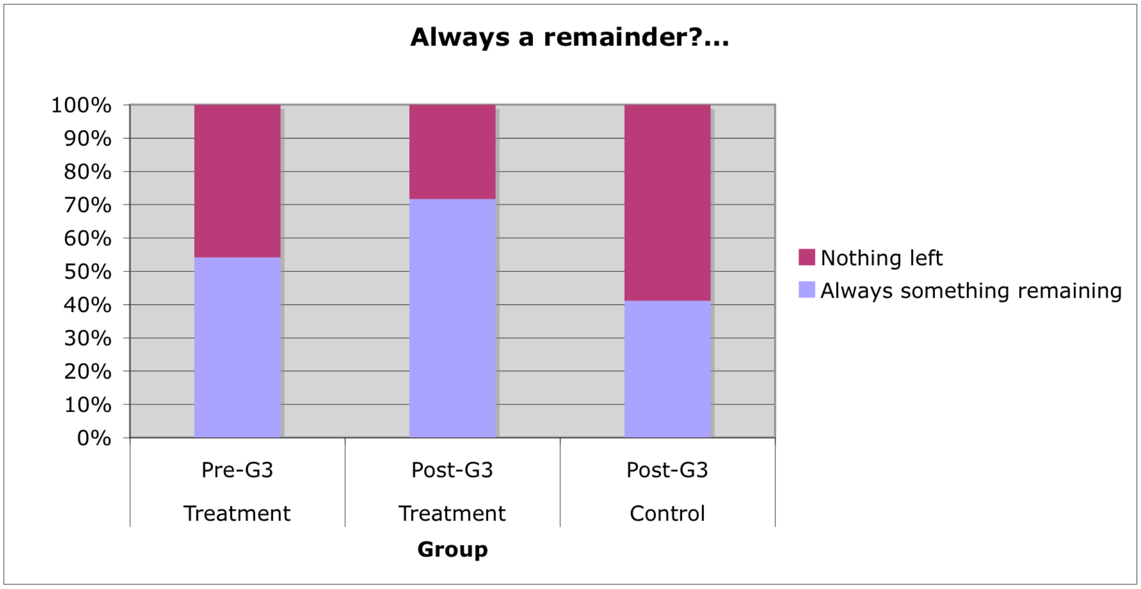


Figure 3. Comparison of student answers to the thought experiment about repeated division of visible speck for the treatment students (grade 3 pre and post) and control students (interviewed at the end of grade 3). Students were asked to choose whether there would always be something remaining or whether there would be a point where nothing was left.

Why do so many deny that small pieces they can see and hold in their hand take up any space or have weight? Their justifications suggest there are probably a variety of inter-related reasons that relate not only to their understanding of taking up space and having weight but also what it means to be "some amount" versus "nothing.

Consider first their arguments for weight. One reason they judge the piece weighs nothing at all is their reliance on "felt weight"—children frequently noted "it doesn't feel like anything in my hand", "it feels like a feather", and "I can't even feel it." One child even explicitly argued: "If it weighed something, I would feel it." Another reason involves the fact that they relate "having weight" to "being heavy". Thus, children also frequently argued—it weighs nothing because "it's not heavy", "it's too light to weigh something" or even "it's pretty light, it weighs zero". These justifications are instructive because they suggest for young children "light" is qualitatively opposed to "being heavy" rather than meaning "something has a little weight". Still another commonly invoked reason was the small size of the object. Children frequently noted the speck was simply "too small" or "too tiny to weigh anything." What is interesting about these answers is that they reveal children have some generalizations relating size and weight, but that they think things get to a point where they exist but weigh nothing at all. Indeed, when children were asked about the invisible piece having weight (after denying the small speck had weight), they sometimes argued: if the larger piece doesn't have weight, then the smaller one won't either. Children sometimes even invoked knowledge of measurement and scales in arguing the speck weighed nothing: "On a scale it won't move", "It doesn't feel like any pound", or even "It's really, really small and weighs a lot less than any amount." This shows how learning about measurement doesn't automatically undermine their belief that small things weigh nothing—it fact they use the evidence from scales to support it!

There is a similar flavor to their arguments about why the speck doesn't take up space. Most commonly children simply assert the speck is simply too small to take up any space. For example: "It's too small…too tiny", "So little it takes up nothing", or simply "it's not that big". Underlying this is the notion that things have to be big enough to take up space; so rather than thinking of taking up space as a continuum of amounts, there seems to be a cut-off value needed to qualify. Children also frequently refer to the size of the surrounding space or container or the fact that the speck can fit anywhere: "It's so small, if you put it on this table, there's a huge big space around it." "It's small and you can store it anywhere." Because when I think of taking up space, it's usually like a bunch of things around it that takes up space. "Really small, can put it anywhere, wouldn't see it, loose it. But this is like really small, so it doesn't take up space. " They may also be impressed by its relative lack of significance: "Because something that's big won't even notice that it is there."

During Grade 3, the treatment students made quick and dramatic advances in the tendency to attribute weight and space to tiny yet visible objects as well as to invisibly tiny pieces of matter. In contrast, the control students, who had their regular science curriculum, did not. This suggests that young children are certainly ready to engage with these issues, but they need instructional support to do so—these ideas don't simply "come for free" with development. (Further data, of course, will be needed to see what happens with the control students in 4^{th} and 5^{th} grade as they advance with the math and science curriculum. Do they "catch up"? Data from previous research suggest that many students arrive at middle school still thinking that small things weigh nothing at all.)

The 3^{rd} grade Inquiry Curriculum provided direct support for thinking about this issue in multiple ways (see Wiser, Smith, Asbell-Clarke, & Doubler, 2009 for more details). Briefly, children used their hands to order a series of objects by weight, compared their "felt weight" order with the weight obtained from using a balance scale, and explicitly discussed the strengths and limits of their senses. They also learned to use measurements to compare the weights of objects and to use weight line representations. They explored the additivity of weight in the 10-10-10-10 challenge (constructing new objects from 10 gm of Styrofoam, 10 gm of Aluminum, 10 grams of wood, and 10 grams of Plasticene) and predicting the weight of the diverse objects created by class members. Finally, they used their weight line representations to explicitly considered what happens to the weight of a 4gm object (clay piece) that is repeatedly divided in half. Would you ever get to zero? They also put the pieces together to confirm that they summed to 4 gm.
The data suggest these discussions were highly effective. As Figure 1 shows, the percent of students arguing the visible speck of clay has weight and takes up space jumped from 10 % to almost 70%. Further an examination of children's justification revealed how children were able to forge new links among existing ideas, which gave them a new

(more theoretical) perspective for thinking and reasoning about weight. Significantly, a number of coordinated changes occurred in children's thinking.

One of the important changes was coming to link "the existence of something" with "having weight" and "taking up space". Thus, nearly all the treatment children who argued that the visible speck took up space and had weight made arguments linking existence, being something, and taking up space and having weight. For example, "If it didn't take up space, then it wouldn't really be there." "Everything in the world takes up space." "Everything has weight, even a grain of sand." "Anything that takes up space has weight."

Some also made explicit links between "seeing something and feeling" something and its taking up weight; having some size. For example, "because like it's a ball and if you see it and feel it, then it takes up space." Recall before that students were denying that things they could see or feel had to take up space or have weight.

Still another change was coming to "explain" to themselves why it's small size, or lack of felt weight didn't matter. Students often noted, "It's still something, it still weighs something, although I might not know it because my hand is not sensitive enough." "It might be very small…it might be half an ounce but you could still measure it." Scientists can measure it with "sensitive scales." "Just because it's small doesn't mean it doesn't take up space."

There are a number of reasons we think children's coming to believe that small visible things take up space and have weight is an important watershed idea.

First, it provides an important conceptual foundation for more deeply understanding the logic of weight and volume measurement. Measurement after all involves additive composition; it also allows extrapolation to scales that go beyond sensory experiences. Believing everything (no matter how small) takes up space and has weight provides necessary theoretical support for children's use of measurement in model building.

Second, it provides a foundation for starting to explain *why* things weigh what they do--which in turn opens the door for developing an explicit (and differentiated) concept of density. Prior work has suggested that differentiating weight and density goes hand in hand with coming to believe that all matter has weight. In our work we will be examining this issue by exploring the linkages between children's patterns of responding about the properties of visible things and (a) their explanations of why things weigh what they do and (b) their differentiation of weight and density (both probed in Task 3, see last section of this paper).

Third, it provides a foundation for thinking about the properties of microscopic pieces. Significantly, we found that most (approximately two-thirds) of children who came to believe that tiny visible pieces take up space and have weight were willing to credit those properties to tiny (invisible) pieces as well. Further, by the end of the 3rd grade, 90% of the treatment students who thought little pieces weighed something and took up space also argued that matter did not disappear with repeated division. All this seems very promising.

Of course, some children may deny things weigh anything or take up space for trivial reasons (e.g., perhaps they just misunderstand question; or colloquially treat "nothing" as equivalent to a "small amount). Alternatively, some of those who do say it weighs something may not truly *understand* the implications of what they are saying (e.g., perhaps they are just memorizing a "slogan" used in class). These are critical issues as the deeper goal of assessment is not just to report patterns of judgment in particular tasks, but to make inferences about changes in children's underlying beliefs and patterns of conceptual organization. We would argue that ultimately we can never make inferences about patterns of conceptual organization from one task; rather assessments must also consider the child's pattern of reasoning across tasks. This is another reason that questioning the same child in a variety of ways on a broad set of tasks is central to developing and evaluating learning progressions. We turn now to look at what we learn when children are questioned in different ways.

2- Conceptualizing Weight, Length, and Number as Dimensions

The Rationale for the Design of the Questions about Granularity

We just mentioned the initial reluctance of young children (i) to acknowledge the existence of objects too small to be perceived as well as (ii) to ascribe the qualities of space and weight to objects when they themselves could not detect these qualities. If we may loosely refer to this case as involving the recognition of *infinitesimal objects*, then we now will consider the challenge of assessing whether students acknowledge *infinitesimal changes in value*.

Clearly, such things cannot be asked directly of young students. But we can make headway by framing the matter in terms of dimensions.

To understand what it means to conceive a physical quantity as a *dimension* (more precisely, a one-dimensional or linear space), it is useful to consider the closely related case of the number line.

A number line is visual depiction of how numbers are ordered; there is a one to one correspondence assumed between each real number and its respective location on the idealized line.[4] The model of the real numbers is only partially revealed through number line diagrams. In actuality, the real number line model includes *operations* on real numbers, most prominently, addition, subtraction, multiplication and division. Arithmetical operations and fundamental properties of real numbers, known as the Field

[4] In a drawn number line, points necessarily have size; in the idealized model of the number line, mathematical points do not have size. There are other differences between a drawn number line—which is a signifier, a symbolic token—and the ideal number line model—the signified or referent. It is even prudent to think of the number line as a symbolic representational system rather than an isolated symbol.

Axioms, can be represented dynamically in the model as actions on line segments corresponding to intervals[5].

When a teacher first introduces a number line in an elementary mathematics classroom, she and her students will be having a discussion about rather different things. Her young students will almost invariably believe that the number line contains only the numbers they know: the counting numbers[6], 1, 2, 3... They believe that no numbers exist between consecutive counting numbers. From our point of view, they believe the number line to be "sparsely populated". Furthermore, young students will fail to understand how subtraction and division relate to operations on the number line. And, until and unless they receive instruction on the matter, they will not understand how a fractional operator (e.g $\times \frac{2}{3}$) relates to multiplication and division by integers. Their expanded knowledge of number systems will go hand in hand with adjustments in their understanding of the number line. The number line also provides a foundation for two-dimensional and higher-order coordinate spaces of use for visually representing functional dependency.

A physical quantity can be thought of as a *dimension* along which the amount or intensity of an attribute (weight, volume, brightness, distance…) can be spatially ordered. The dimension has a metric if there is an agreed upon way to assign values to locations and distances (intervals) on the continuum.

Because young students readily make comparative judgments of quantities and numbers (one object is heavier, wider, warmer… than the other ; 6 is bigger than 2.), we know they treat them as (ordered) magnitudes. But how fine are the smallest distinctions they acknowledge? How many numbers are there in a given interval? How finely grained are their notions of length and weight? At issue are students' presumptions about the granularity of quantities and numbers.

Such questions happen to bear directly on the two issues mentioned above; namely, the shift away from an over-reliance on explanations based on perceptual judgment and the view of measurable quantities as dimensions having properties akin to those of real numbers. In Grade 4, students observed the effect of adding a small object to a glass of water: the water level rose appreciably. However, when asked to consider whether a rock cast into a lake would cause the water level to rise, they were greatly divided. At issue is whether the students are disposed to acknowledge *imperceptible changes in magnitude.*

We created three tasks to probe children’s ideas about the granularity 'or continuity' of attributes thus addressing the second shift (about the establishment of a metric for quantities).

Overview of Key Findings About Granularity

In the Granularity of Weight subtask), children are given two (very different) size clay balls (one is about four times the diameter of the other) with noticeably different felt

---

[5] The additive axioms (associativity, commutativity, identity, and inverse) are straightforward; the multiplicative ones are not. E.g. a multiplicative inverse requires that there be a metric.

[6] The counting numbers are also known as the natural numbers or whole numbers (although some authors consider zero to be a whole number).

weights and are asked if they could fashion a ball that has a weight falling in between the weights of these two (unmeasured) balls. Once they answer this (sometimes with a little assistance from the interviewer about what "in between" means) they are asked: Could there be other weights that are in between? About how many weights could there be between the weight of this (large) clay ball and this (much smaller) clay ball?

In keeping with their initial reliance on felt weight (which has little granularity), the vast majority of children (approximately 4 out of 5 children) in Grade 3 assert there are only very few weights in between the examples (generally only 1, 2, or 3). Some children (about 10-15%) begin to suspect there might be more than a dozen, but still are imagining a limited number (15, 20, 50, etc.). But only exceptionally do young students (approximately 2-3% of the students) suggest that the number of possible distinct weights in the interval is enormous ("lots and lots", "infinite", "goes on and on"). We are interested in tracing how children begin to revise their estimates upward (e.g., maybe 20, hundreds, thousands, lots), ultimately extrapolating there are "as many as you want", "an infinite number." There were no differences between treatment and control students in their answers on these questions.

The <u>Granularity of Length Task</u> probes students' understanding about the linear density[7] of lengths: how many cases lie between the two extremes? The student is shown two lines differing in length. One is approximately 5 inches long; the other is approximately 7 inches long, but no measures are given and no notches or partitions are displayed. The student is asked to venture an opinion about the number of lines of distinct lengths that would be longer than the shorter and shorter than the longer of the two sample lines.

Once again, the vast majority of treatment and control students (over 3 of 4) in Grade 3 believe there are very few distinct lengths that could possibly exist between the two examples. Clearly, these students do not think of length as a continuum or even a highly packed dimension. Rather, there are relatively few lengths in a given interval. Somewhat to our surprise, there was no indication that length was more finely grained than weight.

The <u>Granularity of Number Task</u> has students judge how many *numbers* fall between two integers. Approximately 9 of 10 Grade 3 students, from both the treatment and control groups, only note the counting numbers spontaneously. Many more students acknowledge (sometimes after a prod from the interviewer) that there are fractions amidst the counting numbers. But almost without exception they treat these as a few isolated cases, for example, the case of "halves" or "fourths". Again, there were no differences between treatment and control students.

By obtaining information about the granularity of weight, length, and number, we hope to understand how children establish the idea of physical and mathematical dimensions (or, loosely speaking, continua). This is interesting for a number of reasons, not the least of which is the fact that, in the world of mathematics, the rational numbers introduce density but not continuity (achieved by the real numbers), measurements are confined to rational numbers (with varying degrees of precision and accuracy), matter itself ceases to be understood as continuous, becoming quantized (consisting, instead, of molecules, and

[7] This is density in a mathematical sense, as when one speaks of the density of rational numbers on the number line.

atoms). Such distinctions will be important as students begin to think of quantities as continua; they will also be important for thinking about the nature of matter at very small scales.

Clearly the post treatment grade 3 students find it much easier to conceive of infinitesimal objects (Task 1) than to conceive of infinitesimal differences in value; control students have little appreciation of either. This finding suggests these two understandings of infinitesimals are psychologically distinct and that somewhat different kinds of conversations are needed to develop each insight. Limitations in students' understanding may reflect that these students are just beginning to learn about fractions, and that conversations about these important issues have not occurred in the context of their math instruction, rather than that there are absolute limits in what children these ages are capable of understanding about these issues.

Additional Considerations: Children's Understanding of Division and Divisibility

There is a large body of literature attesting to the elusiveness of division, not only for students, but also for mathematics researchers. In pure mathematics, *division* is an operation involving two numbers, e.g. 7÷3. A *partition* is something quite different, namely the expression of a counting number as the sum of positive integers. With partition, the addends (parts) need not be of the same size; consequently, there are multiple ways of partitioning positive integers (except perhaps for the numbers 1 and 2).

Outside of pure mathematics, additional considerations come into play by virtue of the fact that both physical quantities and pure numbers are involved. In everyday language, "to divide (e.g, a pie) into four pieces" describes a partition, the result of which is is four slices; that is, the complete pie remains. "To divide (a pie) by four" presumably describes a partitive division; the result is one piece!

To find out how many times one extensive quantity "fits into" another extensive quantity of the same nature—e.g. how many dresses can be cut from a particular bolt of cloth—is known as quotition. Quotition evokes the process of measurement: determining the number of times a unit of measure (a dress) fits into or measures a dividend quantity.

In addition, one encounters in science cases in which a quantity of one kind is "divided" by a quantity of another kind. For example one may divide a measure of distance by a measure of time to produce a measure of average speed. The result, a composed quantity, is referred to as an *intensive quantity*.

Insofar as such cases involve very distinct conceptualizations of division and partition, we need to treat them as distinct.

We decided to assess students' understanding of division of numbers in order to compare it to the questions we were asking about the repeated halving of a piece of clay. Would students treat these cases as essentially the same or different? Would repeated having of a piece of clay eventually terminate, leaving no clay remaining? Would the repeated division of a number by 2 always leave a non-zero result or would the process stop at zero?

We were somewhat surprised to learn that division of a number by another number was not something we could assume third grade students to be familiar with. At the beginning of Grade 3, fewer than 1 in 5 students correctly identified the result of (a) eight divided by four and (b) one half of one. Only one of every two students realized that one-half of two equals one.

By the end of grade three, a much greater portion of students seemed to know what is meant by dividing by two (although many students still did not realize that half of 1 is ½). However, the treatment as well as the control students overwhelmingly believed that the repeated division would terminate, reaching zero. (For many it terminates rather quickly after one of two divisions.) It is unclear why they think so, but we will be examining several alternatives as we continue to analyze further questions in the section on number. Were they thinking of division as some people think of Zeno's Paradox? Is it that the notion of limit is abstract?

So far, the data indicate that for many students, the problem is that they are confusing division with repeated subtraction, underscoring the fact that division is not yet well understood in the context of number and is psychologically quite different from understanding of repeated division of physical quantity (assessed in Task 1). Perhaps learning about fractions will make the interminability of repeated division more accessible; after all, each successive division merely doubles the size of the denominator. Significantly, some (of the few) students who showed insight about the repeated division of number, talked about "cutting" and alluded to the similarity of the two problems (dividing physical quantity and dividing number.) In any case, we will follow with eagerness developments in their thinking and we hope to relate their increased awareness about number with their notions about the repeated divisibility of matter. Ultimately, numerical division and division of matter part ways: the former never ends, but the later does. One wonders about the interaction between these opposing points of view at different points of learning and development.

3- Understanding Volume

Rationale for Design of Questions About Volume[8]

Our principal concern was to assess students' understanding of volume before they were familiar with the term, *volume*, or with formulas for calculating volume, such as $V = l \times w \times h$. We decided to create a task that would allow us to identify the invariants (properties and relations) students attended in comparing the size of objects. For this purpose it was sufficient that they investigate a pair of objects, make a judgment of comparative size, and justify that judgment.

---

[8] The full task probed students use of a variety of tools to measure the lengths of two ribbons, to compare the "space covered" of two differently shaped cards, and "the space taken up" or filled by two blocks. The task also involved "faulty" measuring instruments (a broken ruler) and different size paper clips in order to assess their deeper conceptual understanding of length measure (the role of iteration of fixed units; starting from 0, etc.) A full analysis will allow us to compare students understanding of length, area, and volume.

Eventually we wanted to see whether they thought about volume as a single dimension. But this presumed that they were actually attending to volume in their judgments. We decided first to test this assumption. What were the underlying physical properties young students attended to in judging volume?

We also wanted to know whether students knew or could devise a systematic way of attributing values to the volume of each object. Of course, if they knew a formula, they could make use of it. But if they didn't, they could use any of several objects on the interview table to help measure the volume.

In the <u>Volume Sub-Task</u> children are asked to compare the amount of space *filled or taken up by* two solid blocks. In the task (see <u>Figure 4</u>), the wooden block measures 2”x3”x3” whereas the blue foam block measures 2”x8”x1”. [The students were not told these measures. However they were allowed to inspect the blocks and use any of various objects to help in their judgments. The interviewer asks the child to compare the amount of space each object takes up all together, using sweeping gestures to emphasize 3-dimensions rather than 2.]

So it happened that the longer (foam) block was actually smaller in terms of volume. In addition, the block with the greater surface area had a smaller volume. These constraints were built in so as to make it unlikely that the student gave the right answer for the wrong reason. We were able to identify the properties students attended to with a high degree of inter-judge reliability anyway.

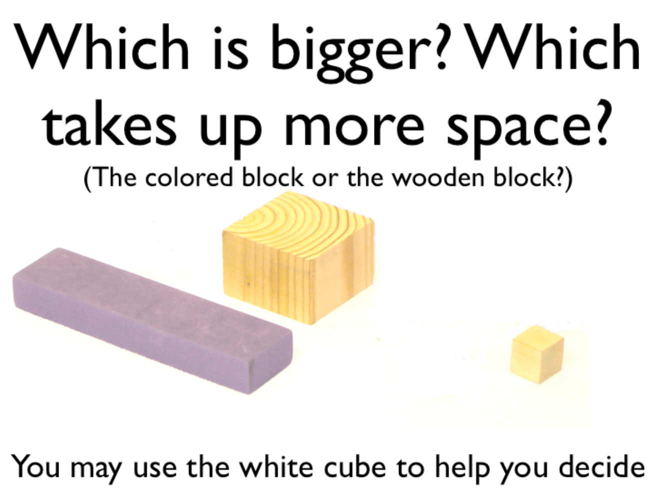


Figure 4: Principal question in the volume task

We first asked the student to determine which block "took more space" (There was little sense in asking third grade students which block had the greater volume). As needed, we waved our hands in a region next to a block in an attempt to indicate we were focusing on 3-space rather than space of one of the block’s surfaces. After giving judgment, the student was encouraged to measure the size of the wooden and foam blocks ("How big are they?") using any of various objects on the table: a clipped measuring tape, paper clips of two different sizes, tiles and small cube (approximately 1" on edge). If a student didn't spontaneously use the small cube to measure the target blocks, s/he was invited to do so. We will examine here how they used the small cubes for measuring volume.

Overview of Key Findings on Volume

Approximately one half of the students at the beginning of Grade 3 focus their attention on lengths or perimeters (see Figure 5). Roughly one third of the students attended to some measure of the area—either the size of a single face or the sum of the surface area of several faces.

Only exceptionally do third grade students without specific instruction about volume use the blocks to build a replica of the object and then count the number of blocks.

It is only fair to recognize that perimeters, surface areas, "footprints", and lengths are all spatial magnitudes that correspond in some sense to how much "space an object uses". Volume is a particular kind of spatial magnitude that needs to be compared to and differentiated from the others.

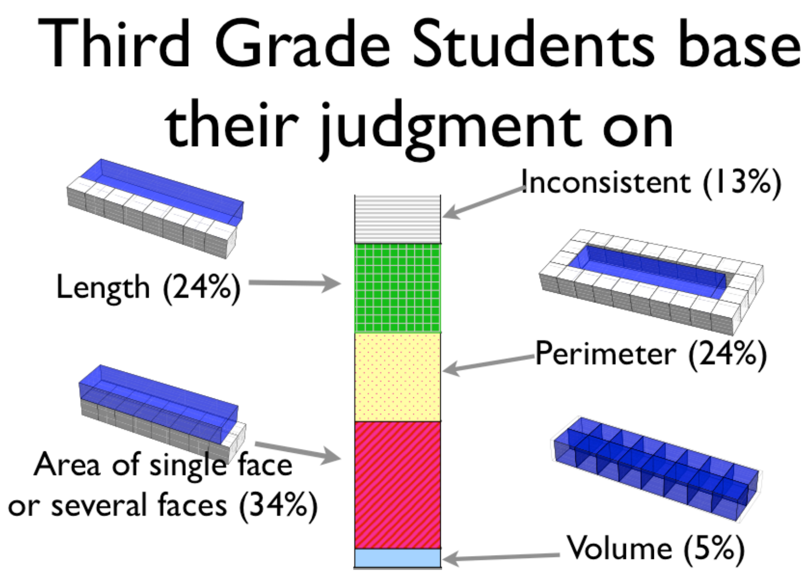


Figure 5: Properties (invariants) that incoming Grade 3 students attended to in judging the size of objects.

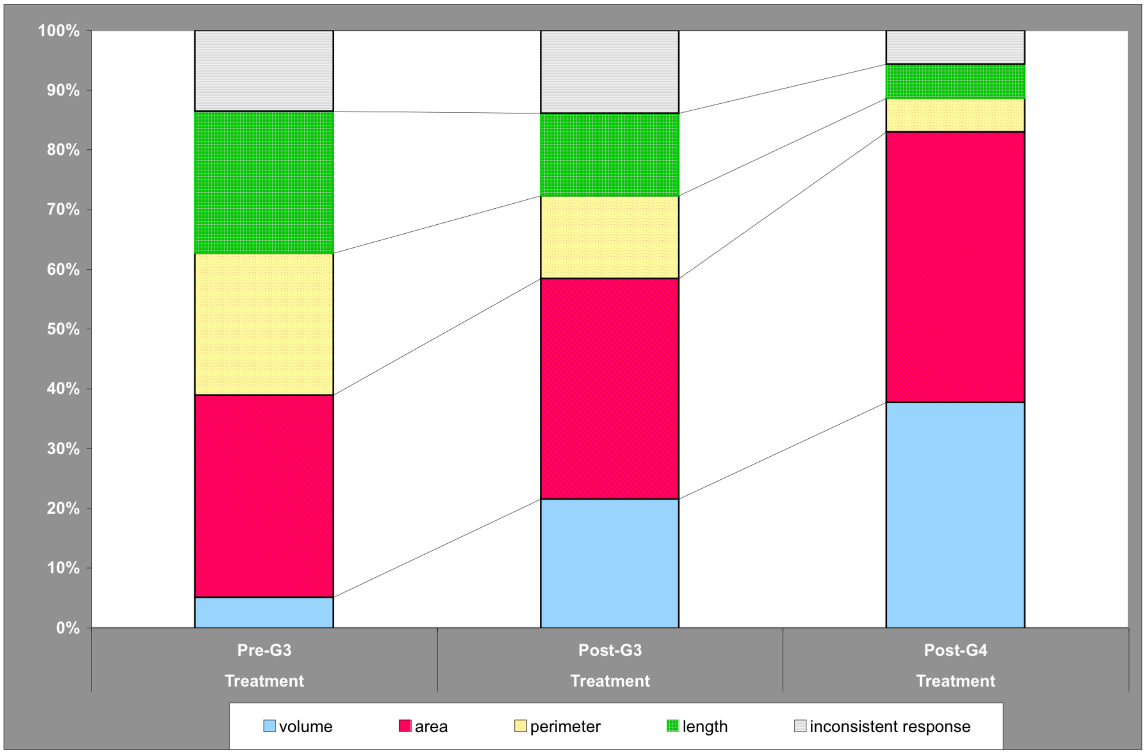


Figure 6: Invariants attended to by students in judging size of 3-dimensional objects

The treatment students make considerably better progress than control students in attending to volume (compare Figures 6 & 7). So there is clear evidence of impact of participation in the Inquiry Project lessons.

What are we to make of the fact that by the end of grade four, less than half of the treatment students have a solid grasp of volume? When we place their results against those of the control group, we begin to suspect that this is progress indeed. We are beginning to suspect that the students who have move away from one-dimensional judgments of size (length and perimeter) but moved to area by the end of fourth grade (especially those who are doing surface area) are making significant progress nonetheless. And we will be curious to see whether their progress continues in grade 5.

In order to get a rich description of their spatial development, we plan to trace how children's attention to different spatial relations and properties interacts with their learning to measure those properties, their learning of differentiated vocabulary for those properties (assessed in the last part of the Length, Area and Volume Task), and their understanding that volume (rather than weight) is relevant to the Water Displacement.

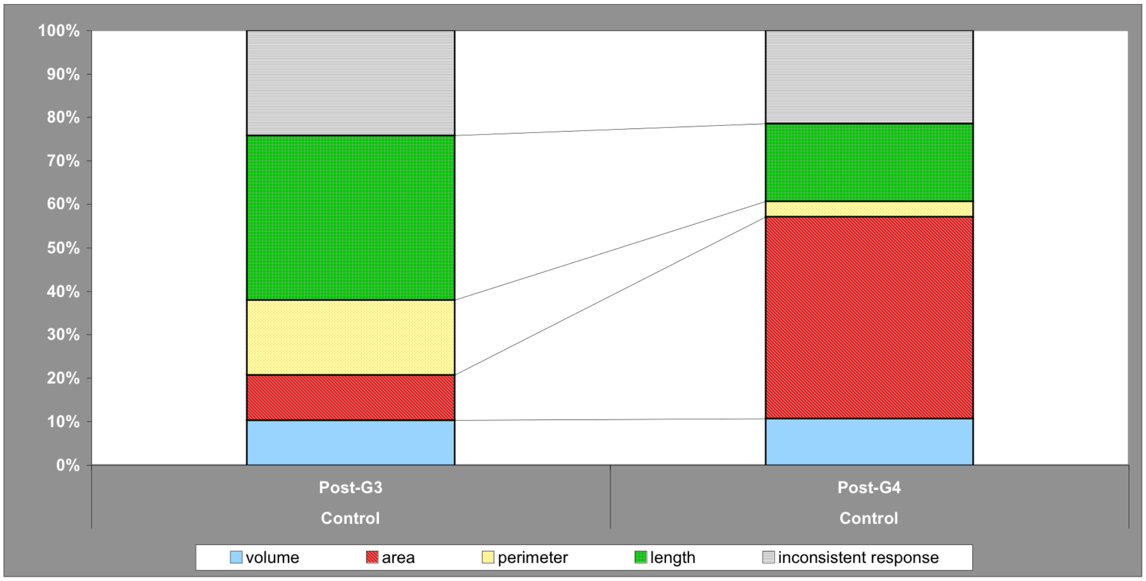


Figure 7: At the end of Grade 4, the Control subjects have not shown any greater awareness of volume. They have, however, reduced their reliance on length and perimeter..

4- Weight, Size, and "Heaviness" of Materials

Rationale and Design of Questions

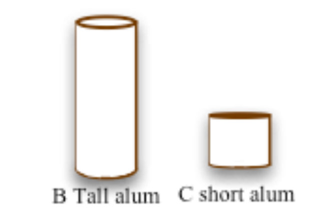
Are bigger objects always heavier than smaller objects? Are heavier objects always larger than lighter objects? Do children realize the weight of an object also depends on the kind of material it is made of? If so, do they make a principled differentiation between the weight of an object and the density of material?

As children move from both perceptually grounded judgments of matter and physical quantities to explicit models of matter and develop reasoning skills related to dimensional analysis, they should be increasingly able to explain the weight of objects as a function of their volume and the density of their materials; further, they should "invent" or assimilate new words or phrases for symbolizing these distinct concepts and use them productively in novel problem solving.

Task 3 of our interview was designed to probe these understandings by engaging children with puzzles about a set of real world objects without requiring that they be familiar with any technical vocabulary (no use of the terms volume and density). In this way, we should be able to see signs of the early development of a concept of density and to study the interaction among multiple elements (reflecting both informal and more formal knowledge about matter and physical quantities) in its construction. There were three main subparts of this task, which each probed for different facets of student understanding.

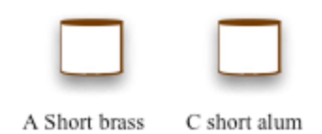
Explaining why larger is not always heavier

Part 1 presents children with a covered set of four objects, and probes their spontaneous *explanations* for four contrasting pairs of objects that have the same or different sizes and weights. (The objects are covered in white contact paper, so they cannot see what they are made of. No mention is made that the objects are made of different materials.) In the first pair, the larger is (predictably) heavier; in the last three pairs, it is not. We are looking at the evolution of children's models of matter that allow them to confront and explain these puzzling situations.



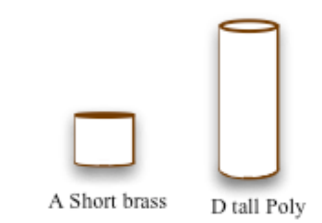
Problem 1

In Problem 1, children are shown a large object (B) and a much smaller one (C) and are asked: (a) *Are these the same size or different*? (b) *Are they the same weight or different*? After eliciting their judgments that B is both bigger than C and heavier, and letting them check that B is heavier by putting it on the balance scale, they are asked to explain why they think this might be so. Most students easily note that one is heavier because it is bigger.



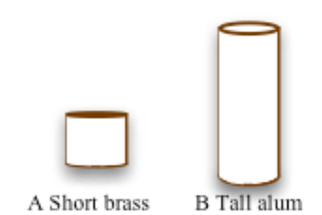
Problem 2

In Problem 2, children are shown two same size objects (A and C) that are quite different in weight (A is much heavier). After asking whether the objects are the same or different size and weights, children are then asked to explain the following puzzle: *How come they are the same size but don't weigh the same?* If children suggest that one is empty and the other full, they are told: *In fact they are both solid pieces. Can you think of any another reason they might have different weights?*



Problem 3

In Problem 3, children are shown two different size objects (A and D), where the smaller is much heavier. After verifying that children agree A is smaller and heavier, they are asked: *How can the smaller one be heavier?* Again if children focus on one being empty, they are told they are solid and asked if they can think of an additional reason.



Problem 4

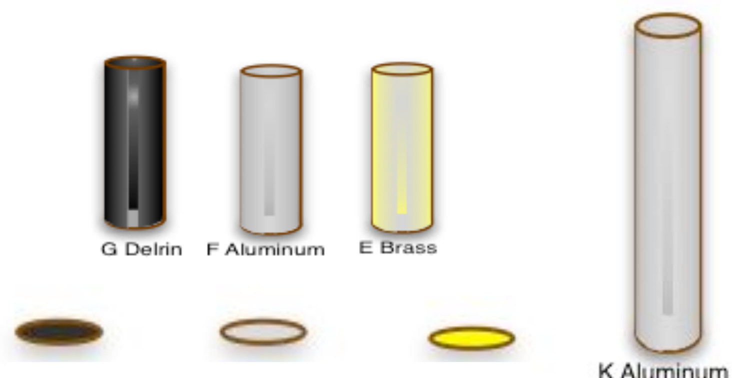
Finally, in Problem 4, children are shown two different size objects (A and B) that turn out to have the *same weight*. After predicting what will happen on the balance scale and then checking their prediction, children are given the final puzzle: *How can they be such different sizes yet weigh the same?*

At issue is whether children have explicit beliefs linking kind of material and size to weight in their matter knowledge network. Also, we want to know whether they can coordinate these variables in explaining why a smaller object can be the same weight as a larger object, or whether they mistakenly conclude that the objects in problem 4 are made of the same material because they are same weight.

Distinguishing "heavier" from "heavier kind of material"

Part 2 of this task then brings out a new set of (uncovered) objects of different sizes made of three different materials. Children are told these objects are made of Delrin (a kind of hard plastic), aluminum, and brass. Children are then involved in making a series of 5 judgments about whether or not one object (in a pair) is made of a heavier kind of material that assess whether they rely on direct perceptual comparisons or engage in more model-driven inference.

To start, children are handed two same size comparisons made of different materials: Objects E and F (same size objects made of aluminum and brass) and then Objects G and F (both same size objects made of aluminum and delrin) and are asked: *Is one of these made of a* ***heavier kind of material****, or not?* If yes: *Which one?* Children have no difficulty concluding that object E (brass) is made of a heavier material than F (aluminum) and that object F (aluminum) is made of a heavier material than G (delrin) as these objects are both heavier and denser.



I Delrin H Aluminum J Brass

Figure 8: Objects and materials used in Part 2 of Task 3.

The last three pairs are the critical comparisons that give us an indication of whether children are making an intuitive distinction between weight and density. In all cases, children have to use the prior information from the same size comparisons and ignore salient perceptual information about the weight of the pairs in reasoning about whether one of the objects is made of a heavier kind of material. These three comparisons are: (1) Object H and Object I, little tiny (same size) slivers of aluminum and delrin that both feel very "light"; (2) Objects J and K: a little (light) sliver of brass and a big (heavy) cylinder of aluminum; and (3) Object H and K (a little (light) sliver of aluminum and a big (heavy) cylinder of aluminum.

Inferring material composition from weight and size

Finally, Part 3 of this task involves children in novel inference making and problem solving (inferring the materials the covered cylinders could be made of). The four covered cylinders are brought out again, along with some extra copies of B, D, A, C. Children are asked to figure out whether it could be made of one of the three materials (delrin, aluminum, and brass) or whether it is made of something else. A balance scale and ruler are available for them to use if they choose.

This task involves making inferences about material composition from information about weight and size. It assesses children's ability to derive inferences and to imagine conditions that would allow the testing of ideas, compensating for differences in values along one variable to compare along another. Of special interest is whether they stack the three small A or B cylinders in order to make a cylinder the same size as the uncovered cylinders E, F, and G and whether they explicitly talk about ratios of quantities.

In fact, B is the same size and weight as F (aluminum). D is lighter than G (Delrin), and is in fact made of something else (a different kind of plastic). A is the same weight as F (aluminum) but one-third the size; if children think to stack the "extra copies" of A, they can determine that 3 A's match E in size and weight. C is not equal in weight to any of the objects; however, 3 C's match F in height and weight.

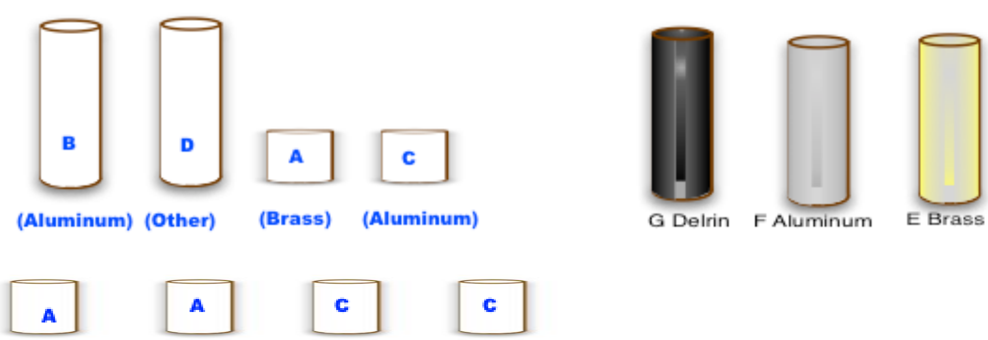


Figure 9. Objects used in Part 3 of Task 3.

At issue is what features of the objects children attend to in making their inferences about what material each is made of and what tools they use: Do they focus simply on weight, or consider the relative size and weight of the objects? Do they not only lift the objects, but use the scale? (It is hard to distinguish the weights of G and D, but G is clearly heavier if they put them on the balance scale? Do they think to use the stacking strategy? How do they explain and justify their choices?

Overview of Key Findings

Our analyses of these data are still in progress. To date we have only completed some analyses of the third grade results. Results confirmed that at the start of grade 3, students were generally at "floor" in distinguishing weight and density throughout Task 3. We look forward to being able to analyze how their understandings progress over time, how their understanding on the different parts of the task relate to each other, and to understandings shown in other tasks.

More specifically, in part 1, the majority (about 80%) of grade 3 pre-treatment students understood bigger objects are often heavier, and appealed to the size differences of objects B and C in explaining why B was heavier. However, they generally did NOT consider that the kind of material an object was made of affected its weight (in the cases where the objects were the same size but different weights, or the smaller was heavier). Instead, some appealed to what might be "inside" the objects, whether one might be hollow or filled with more objects—which seemed more "object" level than material level explanations. Many others offered no deeper explanation at all. Clearly, a pre-requisite for developing a concept of density is being aware of the variation among materials.

In part 2, grade 3 pretreatment children were generally not consistently distinguishing heavy and heavier kind of material. Only 16% of the students were systematically correct in their judgments across the 3 critical problems.

Finally in part 3, grade 3 pretreatment children generally reasoned based on direct comparisons of the weights of objects, rather than consideration of both weight and size. Children often simply compared the objects in their hands, without using the balance scale. The fact that children were often "oblivious" to the importance of size was often quite striking. For example, for children who put Object A (the small brass) on the scale and saw that it balanced with the Object F (the aluminum cylinder that was three times taller), many concluded Object A was made of aluminum. Stacking of objects as a strategy was noticeably absent!

Through use of the nine density cubes, the 3rd grade Inquiry curricular unit provided these children with a wealth of opportunities to learn about different materials and to see the relevance of kind of material for weight differences. Given that the unit did not stress the difference between weight and density, however, (children worked with same size comparisons, in which weight and density are correlated) or fully develop a concept of volume, we did not expect children would make that much progress on fully differentiating weight and density. We did expect that they would improve in understanding that the kind of material was relevant to explaining weight differences. This is exactly what we found.

More specifically there was striking improvement for the post-treatment 3rd graders in referring to differences in material or heaviness of material in their explanations on Problems 2, 3, and 4; there was improvement in the use of specific material kind names. Interestingly, lots of the materials they thought the cylinder might be made of were names of the materials they had investigated in their classroom. The 3rd grade post-treatment group were much better than the 3rd grade control in this regard (who were similar to the 3rd grade pretreatment students). Although it logical to assume this increase in focus on kind of material was a result of the curriculum, it might also be due to the fact that the children had the same interview twice, rather than a genuine improvement. Hence the need to analyze the 4th grade controls. Children in the treatment group also tended to over-generalize and used "material" as a justification for why a larger object was heavier that a small one, without considering that size matters, and to say that objects that were the same weight were made of the same material. Both are evidence that as children form initial links between weight and material, they use their existing "weight" concept rather than a distinct and differentiated notion of heavy for size.

There was some improvement from pre to post treatment in problems that called for differentiation of heavy and heavy kind of material (Problem 4, Part 1 and Part 2), but still very little stacking in Part 3. No clear differences between treatment and control emerge on any of these problems.

It is possible, however, that children in the third grade were making progress (by beginning to link weight and material) that would be "cashed in" in 4th grade when they more fully considered volume and heavy for size. We haven't yet formally analyzed this aspect of data yet; but preliminary inspection suggests that there is a lot more "stacking" among the 4th grade treatment children when making inferences about what the objects were made of in Part 3 (rather than direct perceptual comparisons of weight); there are also many more students who are now consistently using model-based reasoning in making inferences about which object is made of a heavier kind of material (Part 2) (and ignoring direct perceptual information) than there had been in 3rd grade. It will be of interest to see how their performance of 4th grade treatment students compares with 4th grade controls.

Concluding Thoughts

The present paper illustrates how the our assessments were designed *to suit certain characteristics of the population* (young students initially unfamiliar with technical terminology, formulas, and conventions regarding measurement), *to monitor intra-individual change over three years* (repeated measures within a quasi experimental design) and *to capture a variety of conceptualizations*, several of which could not be anticipated at the outset (clinical interview methods). Assessments about "The Properties of Tiny Things", "Conceptualizing Weight, Length, and Number as Dimensions," Volume, and "Weight, Size, Heaviness of Materials, and Density of Objects" were created to provide information about (1) a shift from *perception-centered* to *model-mediated thinking* about physical phenomena and (2) the development of *quantitative reasoning—reasoning about physical quantities in ways that highlight mathematical structure.*

The data from Grades 3 and 4 have provided information about treatment effects as well as fresh insights into students' reasoning processes.

During Grade 3, the treatment students made noteworthy advances in the tendency to attribute weight and space to tiny yet visible objects as well as to invisibly tiny pieces of matter. Their inferences about physical objects in the absence of or in contradiction to sensory information sets the stage for models about phenomena occurring under the threshold of human awareness.

However, most third and fourth grade students seem reluctant to acknowledge the existence of invisibly small *differences* or *changes* in values (infinitesimal differences). At the end of Grade 3, almost without exception, treatment and control students viewed length, weight, volume and number as sparsely populated continua (e.g. shown two sample lines, of 5 and 7 inches in length, they commonly asserted that there existed less than a handful of lines that would possibly have distinct lengths falling between the lengths of the sample lines).

The Volume task revealed that young learners are inclined to attend to a variety of invariants when asked to judge the size of objects. Volume is a three-dimensional property. Many young learners tend to judge size in terms of one (length or perimeter) or two dimensions (area of one or more faces). But, in another sense, volume is a one-dimensional magnitude that shares features with other additive scalar measures. Over a wide range of conditions, it exhibits operations (addition, subtraction, multiplication, and division) and properties (the Field Axioms) akin to the set of real numbers. However, there are a sufficient number of nuances that distinguish reasoning about numbers from reasoning about physical quantities (such as volume) so as to merit further careful investigation.

We have only just begun to analyze the data from the "Weight, Size, Heaviness of Materials, and Density of Objects" Task. Those data, as well as data from six other tasks, will serve to clarify how students begin to draw inferences about relations among measurable quantities. In such cases the structures that emerge are not conferred directly by the mathematical properties of the measures. Rather, they relate to functional dependencies that inhere to the scientific phenomena themselves. As such they will bring us ever closer to understanding how students build models of matter.

References

Carraher, D. W. (1996). Learning about fractions. In L. P. Steffe, P. Nesher, G. Goldin, P. Cobb & B. Greer (Eds.), *Theories of Mathematical Learning*. Hillsdale, NJ: Erlbaum.

Carraher, D. W. (2001). Proportional reasoning. In L. S. Grinstein & S. J. Lipsey (Eds.), *Encyclopedia of Mathematics Education*. New York: Routledge-Falmer.

Carraher, D.W. & Schliemann, A.D. (1991). Children's understanding of fractions as expressions of relative magnitude. *Proceedings of the XV International Conference Psychology of Mathematics Education*. Assisi, Italy.

National Research Council (2007). *Taking science to school: Learning and teaching science in grades K-8.* Washington, D.C.: National Academies Press.

Schliemann, A.D. & Carraher, D.W. (1992). Proportional reasoning in and out of school. In P. Light & G. Butterworth (Eds.) *Context and Cognition*. Hemel Hempstead, Harvester-Wheatsheaf, 47-73.

Schliemann, A. D., & Nunes, T. (1990). A situated schema of proportionality. *British Journal of Developmental Psychology*(8), 259-269.

Smith, C. L., Wiser, M., Anderson, C. W., & 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: Interdisciplinary Research and Perspectives, 14*(1-2), 1-98.

Vergnaud, G. (1996). The theory of conceptual fields. In L. P. Steffe, P. Nesher, P. Cobb, G. Goldin & B. Greer (Eds.), *Theories of mathematical learning* (pp. 219-239). Hillsdale, NJ: Erlbaum.

Whitney, H. (1968a). The Mathematics of Physical Quantities: Part I: Mathematical models for measurement. *The American Mathematical Monthly, 75*(2), 115-138.

Whitney, H. (1968b). The Mathematics of Physical Quantities: Part II: Quantity Structures and Dimensional Analysis. *The American Mathematical Monthly, 75*(3), 227-256.

Wiser, M., & Smith, C. L. (2008). Teaching about matter in grades K-8: When should the atomic-molecular theory be introduced? . In S. Vosniadou (Ed.), *International handbook of research on conceptual change.* Hillsdale, NJ: Erlbaum.

Wiser, M., Smith, C.L., Asbell-Clarke, J., & Doubler, S. (2009). *Developing and Refining a Learning Progression for Matter: The Inquiry Project: Grades 3-5.* Paper presented at AERA Learning Progressions for Matter Symposium, San Diego, CA, April 14, 2009.