Reconsidering Prior Knowledge

Daniel L. Schwartz, David Sears, & Jammie Chang

Stanford University

Corresponding Author:

Daniel L. Schwartz
School of Education
485 Lasuen Mall
Stanford, CA 94305-3096
Email: Daniel.Schwartz@stanford.edu
Phone: (650) 736-1514

An important goal of instruction is to prepare students for future learning. Educators hope that students will be prepared to learn from the next lesson, the next class, the next year, and opportunities beyond. Despite these hopes, assessment practices often overlook the goal of preparing people to learn, and this can make it difficult to judge the value of different learning experiences (Bransford & Schwartz, 1999). For example, televised interviews with recent Harvard graduates revealed serious misconceptions about the causes of the seasons. Under this assessment, their Ivy League educations seemed useless. This is a severe mismeasurement. If these students cared to learn about the cause of the seasons, they would be more prepared to do so than most young adults who never went to college. If the goal is to prepare people to learn, then it is important to design assessments that are sensitive to that goal. As we demonstrate below, assessments of preparation for future learning can reveal the hidden value of educational experiences that can look poor by many standard assessments.

In this chapter, we use preparation for future learning assessments to work backwards to identify the types of prior knowledge that prepare students to learn. Most constructivist inspired instruction attempts to make contact with students' prior knowledge so students will learn better. But what happens when students do not have the appropriate prior knowledge? We highlight how to develop a specific form of prior knowledge that many current models of learning and instruction do not address very well. The situation we emphasize occurs when people's prior knowledge is internally incommensurable. By incommensurable we mean that individuals cannot put information in the same rational system for combination or comparison. Incommensurables cannot be "measured" in the same terms. For example, young children who reason about the balance scale often see weight and distance as incommensurable. The notion of weight cannot be put on the distance dimension, and the notion of distance cannot be put on the weight dimension.

Incommensurable knowledge comes in many forms. With a contradiction, for example, the contradictory claims cannot both be rational, at least not according to one another. Here, we are particularly interested in cases where two or more measures of information cannot be related directly, and it is necessary to develop a higher-order representation to relate them. An example from the domain of statistics can clarify. Consider how beginners often find the variability of a sample. One common solution we have observed is they sum the distances between the data points. Given 3, 5, 7, they might subtract 3 from 5, 3 from 7, and 5 from 7. They then sum the values of the pair-wise distances (2 + 4 + 2 = 6). This works quite well, until one needs to compare samples of different sizes. If one sample has more data points than the other, it will appear to have greater variability because there are more pair-wise distances. To make different sample sizes commensurable, most variability formulas divide by the sample size 'n'. This finds the *average* distance between points and permits a comparison across samples of different size.

We believe that a key challenge for statistics instruction is that people's relevant prior knowledge often consists of pieces that are tacitly or explicitly incommensurable. For example, how can one compare high jump scores with sprint times? Time and distance are different measurement systems. In this chapter, we provide several examples and describe how our instructional model helps students appreciate and work to reconcile incommensurables. The knowledge generated by our instruction often includes a critical early form of prior knowledge. By itself, the knowledge does not support independent problem solving. However, as we demonstrate, this early knowledge has a large payoff because it prepares students to learn standard solutions quite well, which subsequently supports independent problem solving and flexible mastery.

EARLIER FORMS OF KNOWLEDGE

People learn by building on prior knowledge and abilities. This suggests it is important to design educational activities that are relevant to students' prior knowledge so they can treat lessons meaningfully. By the same token, students need the right prior knowledge to start with. If students do not have useful prior knowledge, then there is a strong risk that they will build new knowledge on a faulty foundation; they may develop misconceptions or brittle behavioral routines. Our instructional goal is to develop forms of prior knowledge that can support learning. We will label these as "earlier" forms of knowledge to differentiate them from the prior knowledge that students already have.

[Figure 1 about here]

To clarify the idea of earlier forms of knowledge, we borrow Vygotsky's (1934/1987) concept of the zone of proximal development. The top of Figure 1 shows that Vygotsky's construct comprises three zones. In the leftmost zone, the child cannot complete a performance, even with help. For example, a child may be incapable of adding two numbers. In the zone of proximal development, the child can complete a mature performance with help, for example, if an adult holds up fingers on both hands to help the child count the combined quantities. Finally, experiences in the zone of proximal development can help a child move to a zone of independent performance – the child can add two numbers without help.

A common instructional interpretation of Vygotsky emphasizes the latter two zones, as shown in the middle panel of Figure 1. Educators embrace the zone of proximal development for its instructional implications; it is important to consider what types of scaffolds, learning resources, and activities can best build on children's prior abilities to support learning and subsequent independent performance. By this interpretation, the latter zone of independent performance is most relevant to issues of assessment. To find out if children have learned, educators can examine how well they complete tasks independently. Bransford and Schwartz (1999) referred to these assessments of independent performance as *sequestered problem solving*, because students are shielded from learning resources and scaffolds that can contaminate a test of independent performance.

A second educational interpretation of Vygotsky, which we propose here, shifts the focus to the left, as shown in the bottom panel of Figure 1. In this view, there is more attention to instructional activities that can expand the zone of proximal development so a child can learn from available resources. In addition, the zone of proximal development itself becomes highly relevant to issues of assessment. Vygotsky directly spoke to the assessment implications:

> Like a gardener who in appraising species for yield would proceed incorrectly if he considered only the ripe fruit in the orchard and did not know how to evaluate the condition of the trees that had not yet produced mature fruit, the psychologist who is limited to ascertaining what has matured, leaving what is maturing aside, will never be able to obtain any kind of true and complete representation of the internal state of the whole development…" (1934/1987, p. 200)

As Vygotsky pointed out, assessments of a child's zone of proximal development are a valuable way to estimate a child's developmental trajectory. These types of assessments are often called dynamic assessments (Feuerstein, 1979), because they examine how well a child can move along a trajectory of learning when given help. Focusing on learning rather than child maturation, Bransford and Schwartz (1999) referred to these types of assessments as tests of *preparation for future learning* (PFL). They argued that sequestered assessments of independent problem solving can be misleading when the question of interest is really whether people are in a position to continue learning.

The left-shift in our focus on Vygotsky's formulation also changes instructional perspectives. Rather than solely focusing on scaffolds or learning resources that support competent performance, it is also important to consider how to develop earlier forms of knowledge that prepare people to take advantage of future possibilities for learning.[1] For example, Martin and Schwartz (2005) found that while certain learning materials did not affect independent performance, they did influence how prepared students were to learn from novel resources. In this study, they authors taught children to solve fraction addition problems using either pie wedges or tile pieces. Children in both conditions learned equally well to solve the problems using their base materials (pies or tiles). The children then received similar problems to the ones they had successfully learned to solve, but this time they had to solve them in their head without any materials. The students in both conditions did equally poorly. Thus, by this test of independent problem solving, the two methods of instruction performed the same. However, Martin and Schwartz also asked the children to retry the problems they could not do in their heads, this time using new physical materials (e.g., beans and cups, fraction bars, etc.). This PFL assessment, which included scaffolds in the form of new hands-on materials, revealed a different picture. The tile students learned to use the resources to solve 55% of the problems, whereas the pie students could only solve 35% of the problems. Thus the tile students were better prepared to learn how to use the new materials successfully, even though they could not solve these problems in their heads. Evidently, it is possible to develop forms of early knowledge that, although insufficient to support unaided performance, determine what learning resources will be within a student's zone of proximal development.

Developing earlier knowledge that can prepare people to learn is particularly important for statistics. The literature on judgment under uncertainty highlights the danger of not developing appropriate prior knowledge (Kahneman, Slovic, & Tversky, 1983). Statistics instruction can implicitly tap into inappropriate intuitions that have a surface similarity to statistics proper. People will use these intuitions to make sense of statistics instruction, but because the intuitions are not quite right, people will develop misconceptions.

For example, statistics often involves extrapolating from a sample to make an estimate of a population. The proportionality of the sample ratio and the estimated population ratio taps into a notion of similarity, and people will think about probabilistic situations using similarity rather than likelihood. We recall asking the following question of a psychology professor, "If you need a sample of 50 people to get a reliable estimate for a population of 1000, how large of a sample do you need if the population is 10,000?" The professor incorrectly replied that you would need a sample of 500 people (cf. Bar-Hillel, 1980). Though we are sure the professor had been taught the law of large numbers, this professor reasoned with similarity instead. One solution to this type of recurrent problem is to defuse misconceptions as they arise (e.g., Minstrell, 1989). For instance, we pointed out, to our colleague's chagrin, that the reliability of the sample is independent of the size of the population. Another solution is to develop the right early knowledge to start with so people will learn more deeply when they are first taught. In the following section, we describe the type of earlier knowledge that we think prepares people to learn well in the domain of statistics.

TROUBLES IN PRIOR KNOWLEDGE

There are many different forms of prior knowledge that can support subsequent learning. In general, there are two common approaches to the role of prior knowledge that can help people learn. One takes a conceptual focus and attempts to build upon students' prior intuitions and experiences. The other takes a procedural focus and aims for the mastery of component skills and the subsequent combination of those components. Both of these approaches are highly valuable, but here we highlight situations in which they appear insufficient. In particular, these methods have difficulty when prior knowledge holds incommensurables.

One Challenge of Building on Intuition and Everyday Experience

As stated before, people always build on prior knowledge to learn, and explicit efforts to leverage prior knowledge can be extremely valuable in instruction. A simple analogy can help clarify the structure of an otherwise foreign concept (atoms are like the solar system). This approach to instruction presupposes that the prior knowledge is already in place in the form of intuitions or grounded experiences. The challenge of instruction is to make contact with students' prior knowledge. If students cannot make sense of a lesson with their pre-existing knowledge, they will not understand.

There are cases, however, when people can encounter difficulty building correctly on prior intuitions because those intuitions are incommensurable in the learner's view. One example comes from an investigation of 12-year-olds' concepts of sampling. Schwartz, Goldman, Vye, Barron, and CTGV (1998) asked children to design sampling protocols for two different scenarios. In the fun booth scenario, the children had to develop a plan to sample 40 students to find out what type of fun booth the 400 students at the school would prefer at a school fair. In the gender scenario, they had to develop a plan to sample 40 students to estimate how many boys and girls there were at their school of 400 students. The children understood the idea of using a sample to infer a population. They also drew upon a variety of everyday concepts that exhibited family resemblances to the notion of sampling. Figure 2 provides a schematic summary of the different ways the children thought about taking a sample. The students had relevant prior knowledge, and the instructional challenge appeared to be how to connect this knowledge.

[Figure 2 about here]

The children, however, also exhibited a strange inconsistency. For the gender scenario, 87% of the children generated relatively random sampling methods, such as standing at the front door at the beginning of the day and counting the number of boys and girls among the first 40. In contrast, for the fun booth scenario, 87% of the children generated biased sampling methods, selecting their sample based on whether the people had attributes relevant to a particular fun booth. For example, the children said they would sample the baseball players to see if they would come to a booth that involved throwing balls to dunk a teacher. The context of sampling influenced how the children reasoned about taking a sample. They appeared to be shifting interpretations of the task to make sense of these two situations.

Further examination revealed the problem. The children thought of cause and chance as incommensurable. The gender scenario did not involve any potential causal relations between the people sampled and the outcome. The scenario is like sampling marbles from an urn, where nothing apparently causes the color of the marble. In cases that do not involve potential causal relations (e.g., dice, spinners), the children accepted random sampling methods. In contrast, the fun booth scenario brought up a *covariance assumption*. The children thought in terms of how the attributes of the sampled people would cause (or co-vary) with their likely answer. This led to the biased sampling methods, because the children restricted their samples to the segment of the population most likely to be interested in a particular fun booth.

The distinction between random and causal situations was quite strong for the children. The children thought that random sampling methods could not apply to causal situations. As one student stated, “You don’t just want to pick kids by chance, because you might get unlucky and end up getting kids who all like dunking booths.” In retrospect, we should have known. For mature statistical reasoning, randomness has patterns (e.g., a normal curve) and capitalizing on these patterns permits causal inferences. But, how would children of this age have a conception of how randomness could be used to help infer causality when cause and randomness are seen as opposites (Piaget & Inhelder, 1975)?

In a subsequent study, Schwartz et al. (1998) tried to build on the children's prior conceptions. The authors had some successes, but overall they did not succeed as well as they would have liked. A new group of children worked for two weeks developing a plan to solve a Jasper Adventure that involved issues of sampling (CTGV, 1997). The Jasper Adventure, called the Big Splash, presented students with a 20-minute video of a child who needed to develop a plan for a fair. The Jasper Series is an example of an excellent instructional model that attempts to build on children's prior knowledge. It develops a problem in an easily understood visual format that taps into children's own prior experiences. Even so, the children were never able to reconcile issues of cause and chance satisfactorily, and instead, they used another pre-existing concept that worked in many sampling situations shown in the video. They began to think of sampling in terms of "fairness." A fair sample gave everybody an equal chance of participating. Unfortunately, the concept of fairness is not adequate. On a posttest, the children liked the idea of leaving a survey on a table. They felt that any students who were interested could fill out the survey, so it was fair. This method of sampling leads to a self-selection bias; only those students interested in a dunking booth might choose to fill out the survey. Building on the children's prior knowledge was insufficient. The children never reconciled cause and chance, and as a result, they relied on prior concepts that were adequate for some situations, but manifested misconceptions in others.

Challenges to the Mastery of Skills

In contrast to building from prior intuitions, a mastery approach tries to develop the relevant prior knowledge. Mastery-focused instruction has learners memorize or master component skills that will later become part of a more complex performance. For example, an excellent way to learn to play a piano piece is to first learn the left hand of the first phrase, then the right hand, and finally put the two hands together before moving on to the next phrase. Much success has been achieved with this method of instruction. One of the primary methods of teaching children to read is to build component skills through phonics. In math, it is important for children to have mastered single-digit subtraction before they take on multi-digit subtraction. Efficiency in component skills can free up cognitive resources for learning new materials and can prevent errors in execution. In these cases, the earlier knowledge that prepares people to learn also supports miniature independent performances, though these performances are restricted to tasks that fall short of the ultimate learning goal. Overall, the mastery model emphasizes independent performance as both the input and output of instruction.

An emphasis on independent performance is exceptionally valuable when the ultimate context of application comprises situations of high frequency and stability. Reading, for example, occurs over relatively conventional forms of text. Educators can count on these stable textual contexts when designing instruction (words are grouped horizontally, reading goes left to right, periods signal the end of an idea unit). The same is true of many apprenticeships and jobs that involve stable practices and few opportunities for mobility. It is also true for the recurrent problem formats that occur in highly stylized curricula, like those of intelligent tutors. In these cases, the goal is not to prepare people to continue learning given changing circumstances, but rather, the appropriate goal is to build a set of self-contained independent performances.

Efficient mastery of facts, skills, and concepts is important in all domains, including contexts that require innovation (Schwartz, Bransford, & Sears, 2005). Yet, there are perils associated with a unitary emphasis on mastery. One limitation is that mastery approaches may not always develop people's zone of proximal development as well as they might. People may master knowledge for one form of independent performance, but this not increase their abilities to learn in new situations. Except for high stability and high frequency application contexts, it is hard to guarantee that instruction for one form of independent performance will fit the type of performance needed for a new situation.

For example, Schwartz and Bransford (1998) asked college students in one experimental condition to write a summary of a chapter on classic studies of memory. In a second condition, students received simplified data sets from the studies, but they never saw the chapter. Their task was to graph what they thought constituted the important patterns in the data. On a subsequent true-false test, the summarize students did much better than the graphing students. They had a better mastery of the facts. However, this mastery was limited to simple retrieval performances, and it did not help them learn. Evidence to this point came from a second part of the study. The students from both conditions later heard a common lecture that reviewed the studies and their implications for broader human behavior. On a test a week later, the students received a description of a new experiment and had to predict the results. Students who had completed the graphing activities produced twice as many correct predictions as the summarize students with no increase in wrong predictions. The graphing students had been more prepared to learn from the lecture and then use this learning in a new situation. We know the graphing students learned from the lecture, because other graphing students who did not receive the lecture did quite badly on the prediction problem. One explanation of these results is that the graphing activity, which did not yield mastery on the true-false memory test, still developed important forms of early knowledge that prepared students to learn deeply from the lecture and apply this knowledge flexibly. A second explanation is that the summarize students thought they had already mastered the chapter, and therefore they did not listen as carefully. In either case, activities that overly emphasize mastery for specific forms of performance may not be ideal for situations that depend on learning related ideas that take a slightly but significantly different form.

A second example of the strengths and limitations of mastery for future learning comes from a study by Schwartz and Martin (2004). They taught roughly 200 high school students how to compute variability using the mean deviation and to understand why the mean deviation divides by 'n'. After instruction, the students received a test that examined their mastery, their "insight," and their preparation for future learning. The mastery asked them to compute the mean deviation for a set of numbers. The insight item showed the students the mean deviation formula and asked why it divides by 'n'. Their answers to this latter question indicated whether they appreciated that dividing by 'n' provides a way to compare samples of different sizes. The PFL item asked students to determine which of two basketball players was more consistent in their points-per-minute scoring. They had to determine the variability (covariance) of bivariate data (points and minutes played). These students had only learned about univariate data, so working with bivariate data was something new and would require learning how to put two incommensurable dimensions, points and minutes played, within a single structure.

Figure 3 shows the percentages of students who learned to solve the bivariate problem during the test. The results are broken out by whether the students showed mastery and/or insight into the mean deviation formula. The notable finding is that mastery of the mean deviation formula by itself was not optimal for helping students learn. The students also needed insight on issues of incommensurability (i.e., why it is important to divide by 'n'). At the same time, insight alone was not optimal either. Students needed both forms of knowledge to help them solve this very difficult transfer problem that required going beyond what they had learned before. This is correlational evidence – later we provide relevant experimental evidence. For now, the take home message is that mastery is important, but there are other forms of earlier knowledge that prepare people to learn. In particular, explicit early knowledge about incommensurables can be helpful in the domain of statistics.

[Figure 3 about here]

Inconsistency in Prior Knowledge for Learning

To summarize, there is a large family of learning theories that presuppose commensurable prior knowledge or intuitions. However, in the study of children's notions of sampling, we found that the children had incommensurable prior knowledge, and that replacing these concepts with yet a third piece of knowledge (fairness) did not lead to a mature understanding. Modern associative theories of learning propose that learning is characterized by the aggregation of component prior knowledge. This mastery view of learning works well, until people have implicitly inconsistent knowledge that cannot be easily associated. Though it may be possible to teach people the right performance, the underlying inconsistencies may remain and yield misconceptions or interfere with their abilities to learn subsequently.

A different view of learning, that does not presuppose consistent prior knowledge, explicitly emphasizes the reconciliation of inconsistent or incommensurable knowledge. Chi, De Leeuw, Chiu, and LaVancher (1994), for example, found that one benefit of self-explanation is that it helps learners work through inconsistencies in their knowledge. Piaget, in particular, described development in these terms. Children discover that their beliefs are internally inconsistent, which causes a disequilibrium. Through a process of reflexive abstraction that resolves the disequilibrium, the children develop more mature knowledge. Piaget, however, did not write very much about instruction. He also did not consider situations where implicit contradictions and incommensurables do not cause disequilibrium, which we think is often the case, especially in statistical reasoning. We have attempted to develop a model of instruction that directly helps students see incommensurables and try to handle them. In turn, this early knowledge prepares students to learn the canonical solutions invented by experts.

LEARNING AS RECONCILIATION

The significance of reconciling incommensurables for learning was first developed in detail by Plato to address the learning paradox. The learning paradox asks how people can possibly learn something if they do not already know it in some form, thus raising the very question of whether it is possible to develop new knowledge, or whether all knowledge growth needs to presuppose complete prior knowledge. Accepting the paradox leads to a view of learning that emphasizes the refinement of prior knowledge (e.g., innate concepts, forms, language modules, intuitions, and so forth). Meno formulates two aspects of the paradox specifically:

> But how will you look for something when you don't in the least know what it is? How on earth are you going to set up something you don't know as the object of your search? To put it another way, even if you come right up against it, how will you know that what you have found is the thing you didn't know? (Meno, 80.d; Plato, 1961).

The first half of the paradox asks how people can look for knowledge if they do not already know what they are looking for. Plato's solution is that incommensurables alert people to the need for learning. As we mentioned, the term incommensurable refers to the situation where multiple elements cannot be accommodated by the same rational system. An example comes from an isosceles right triangle. If we define the length of the sides as one unit, the hypotenuse yields an irrational number. Alternatively, if we define the hypotenuse length as the unit, the sides will yield an irrational number. Thus, a failure in a measurement system can let one know where to look for new knowledge. It causes the disequilibrium that begins the search for resolution.

The second half of the paradox asks how people can recognize that they have found knowledge if they did not know it already. The solution to this half of the paradox is that people will know they have found new knowledge when the incommensurables can be explained by the same system. So, in the case of the right triangle, the resolution comes by squaring the sides, which can then be measured in the same unit system. One knows they have "found knowledge" because it explains previously incommensurable elements.

There are three elements of this account that we take to heart in our instruction. First, it is important to help students sufficiently differentiate the elements of a problem so they notice there are incommensurables. Second, we provide students with the impetus and means to work on reconciling the incommensurables into a higher-order structure. Finally, we provide students with the canonical solution. It is unrealistic to expect students to develop conventional solutions that took trained experts many years to invent. The first two elements of our instruction create the earlier knowledge that prepares the students to deeply understand the "ideal" solution when it is presented to them.

In the Meno, Socrates demonstrates the resolution to the learning paradox by leading Meno's slave through a geometry problem similar to the right triangle example. The slave never resolves the incommensurables on his own. Ultimately, Socrates provides the solution to the slave, who debatably, is in a position to recognize the significance of the solution ("Yes, you are right Socrates"). In other words, the process of noticing incommensurables and trying to reconcile them sets the stage for appreciating the solution when it appears – whether through personal discovery, the grace of the gods, or in the more typical case, by being told. Thus, we adopt the aims of the Socratic method, but we do not adopt the Socratic method of question asking, which only the most talented of people can wield successfully.

DEVELOPING EARLIER FORMS OF KNOWLEDGE

In our instructional model, students develop an appreciation of incommensurables at the same time as they attempt to reconcile them. This simultaneity is important, because students co-evolve their ability to identify the incommensurable features with their ability to handle them. The two processes fuel one another. However, for discussion, we can distinguish them analytically.

Noticing Incommensurables through Contrasting Cases

The first component of our instruction is built into the instructional materials and helps students perceive the features that enter into an incommensurable relation. When people enter a new domain, their prior knowledge can be too vague to support the appreciation of lurking incommensurables. Therefore, we help them develop more precise early knowledge. To do this, we rely on contrasting cases. Contrasting cases are a powerful way to help people notice specific features they might otherwise gloss (Bransford, Franks, Vye, & Sherwood, 1989; Gibson, & Gibson, 1955; Marton & Booth, 1997). For instance, Howard Gardner (1982) describes an art exhibit that juxtaposed original paintings and forgeries. At first people could not tell the difference, but through careful comparison, they began to notice the features that differentiated the original. Similarly, in our materials, students learn to discern relevant features by comparing data sets. Contrasting cases of small data sets that highlight key quantitative distinctions can help students differentiate important quantitative features.

[Figure 4 about here]

Figure 4 provides one example of how we do this for the topic of variability. Each grid shows the result of a test using a different baseball-pitching machine. The black circles represent where a pitch landed when aimed at the target X. The grids create a set of contrasting cases that alert learners to important features of distributions. For example, when most students come to variability, they over-assimilate the concept of variability to their well-developed notion of accuracy. Variability is viewed as a lack of accuracy, rather than deviations around the mean. The pitching grids specifically include an example where all the pitches are extremely close together, yet they are far from the target. This helps the students notice that variability and lack of accuracy are distinguishable properties. The contrasts between the four machines also draw attention to issues of sample size, outliers, and density.

Working Towards a Higher Order Structure

Helping students differentiate key quantitative features is an important component of our instructional model, but it is not sufficient for developing the types of earlier knowledge that we think prepare people to learn. Well-differentiated knowledge may only yield a set of discrete observations. But, the goal of most statistics instruction is for students to understand the structure that can accommodate all the features. For example, the grids include a number of incommensurable properties that models of variance have been designed to handle. For instance, sample size and distance from the mean are incommensurable features. One does not naturally measure distance in terms of sample size or vice versa. Variance formulas provide a structure so that sample size and distance can be put in the same rational system. For example, with the pitching grids, one can divide the total distances between the pitches by the sample size to get the average distance between pitches. We think students need to recognize the incommensurable that dividing by 'n' resolves so elegantly.

We doubt that simply pointing out that two components are incommensurable is going to work very well, because students will still not know how to reconcile them. It is important to have students work to reconcile the incommensurable elements into a single structure. The recognition of incommensurable features without the effort or means to reconcile them can cause people to fall back to one dimension or the other when put in a situation that requires both. One example of this behavior comes from the children who flipped between random and covariant sampling methods above. Another example comes from work conducted with children learning about the balance scale. Schwartz, Martin, and Pfaffman (2005) provided 9-10 year-old children with a series of balance scale problems that varied the distances and number of weights on either side of the fulcrum. For each problem, the children chose what they thought would happen (tilt left, balance, tilt right), and they justified their prediction. They then received feedback by seeing the true behavior of the scale, and they were asked to re-justify the result if they had predicted incorrectly. Half of the children were told to use words and half were told to "invent math" to justify their choices. The children told to use words performed like average children of this age, using one dimension, for example "more weight," to justify their choice. If their choice was wrong, they simply switched to the other dimension ("more distance"). Only 19% of the children included both distance and weight in the same justification, and the children did not appreciably improve in their ability to predict the behavior of a balance scale.

The children who were told to invent math showed a much different pattern. They explored different types of mathematical structures across problems including addition, subtraction, and multiplication. All told, 68% of these children included both dimensions of information in their justifications. The "invent math" children tried to find a way to reconcile the different dimensions of information, and ultimately, they learned to solve balance scale problems at the rate of adults.

One take home message from the balance scale study is the following: To help students work towards reconciling dimensions, it is important to provide tools that can help them make headway. The "word" children in the balance scale study knew they were flipping back and forth between dimensions, but they had no way to do anything about it. Young children lack the cultural structures and artifacts that can help organize the complexity of incommensurables into a higher-order structure. One cultural structure that is particularly good for handling complexity is mathematics. Thus, when we have students work with statistical notions, we do not ask them to simply notice things qualitatively. We encourage them to use math to help find a way to manage the high complexity of distributions.

To return to the pitching grid example (Figure 4), we included a task structure that ensured the students would recognize and attempt to reconcile the different quantitative features. The students' task was to develop a mathematical procedure that computes a single value to help shoppers compare consistency between the machines, much as an appliance receives an energy index. By providing the students the goal of comparison, they notice contrasts in the data from each machine, and they work to innovate a common measurement structure that incorporates all of these dimensions in making comparative decisions.

Delivering a Canonical Solution Once Students are Prepared to Learn

Finally, per our three component instructional model, we assume that noticing and working to reconcile incommensurables creates the earlier knowledge that prepares students to learn the higher-order structure created by experts. So, even though no students invented an ideal solution to the pitching grid problem, they developed the earlier knowledge that creates a time for telling (Schwartz & Bransford, 1998). For example, given a lecture, the students can appreciate the significance of what the lecture has to offer. Lectures and other forms of direct instruction are not inherently less valuable than hands-on activities. Their value depends on the knowledge that students can bring to the exposition, and we assume that the first two components of our instructional model help create the relevant earlier knowledge.

This assumption appears to be borne out. For example, Schwartz and Martin (2004) provided 9th-grade students with a five-minute lecture on the mean deviation after they completed the pitching grid activity (plus one more). The authors' assumption was that the lecture would be within the students' zone of proximal development, and they would learn well and master the procedure. Pre- and posttests suggested this was the case. On posttest, 86% of the students could compute a measure of variance compared to 5% on pretest, which is impressive given that the students heard the complicated lecture in a very brief presentation and only practiced for 10 minutes. On a delayed posttest a year later, and without any intervening practice, the students still performed at 57%. This is well above the 0% the authors found in a sample of university students who had completed a semester of college statistics within the past year. It is important to note, however, that asking the college students to compute a measure of variance was like asking them the cause of the seasons – it was a potential mismeasurement of what they had learned. We simply present the college benchmark to show how well the 9th-graders had been prepared to learn from the lecture, and at the end of the next section, we provide a better measure of the college students' understanding.

A DEMONSTRATION OF INSTRUCTION THAT PREPARES PEOPLE TO LEARN

The preceding claim that our instruction prepared students to learn from the brief lecture is only a plausible inference. The students had limited exposure to the conventional solution, and they showed abilities quite above what we had previously observed among college students. Nevertheless, we did not compare how well these students learned from the lecture versus students who had not completed our preparation activities. Therefore, it is important to develop evidence in a more compelling research design. In this section, we describe that evidence.

The experiment arose from a concern that most current assessments of knowledge focus on mastery or independent performance. These assessments use tests of sequestered problem solving, where students have no chance to learn to solve problems during the test. We were worried that this assessment paradigm had created a self-reinforcing loop. Educators use methods of procedural and mnemonic instruction that support mastery as measured by these sequestered tests, and therefore, the tests generate continued feedback that tells educators their methods of instruction are effective. Our concerns were two-fold. One was that sequestered tests tend to measure mastery, but we suppose that a goal of secondary instruction is to increase students' zone of proximal development so they can continue to learn. Our second concern was that these sequestered tests were likely to miss the benefits of activities that engage students in innovating their own solutions. Students often do not generate the canonical solutions that sequestered tests evaluate, and therefore, they perform badly. However, we thought well-designed innovation activities may prepare students to learn efficient solutions once they appear. Therefore, we thought it was important to work on a form of dynamic assessment that included resources for learning during the test and could evaluate students' preparation for future learning.

The instructional content of the experiment tried to help students learn how to reconcile yet another incommensurable common to statistics: normalizing data using standardized scores. Standardized scores make seemingly incommensurable data comparable (e.g., comparing athletes from different sports or eras), and they are one key measure in inferential statistics. The experiment began on the last day of two weeks of instruction (Schwartz & Martin, 2004). During the prior two weeks all the students had completed a number of “noticing and reconciliation” tasks and received direct instruction on canonical solutions in the domain of descriptive statistics. For the experiment on learning standardized scores, students from multiple classes were divided into two different instructional treatments: a tell-and-copy condition and an invention condition.

Both conditions began with a front of the class presentation on the problem of comparing grades from students in different classes. The presentation showed how an 85% could be an average grade for one teacher’s class, but it could be an outstanding performance in a second teacher’s class. Students quickly observed that it did not seem fair for students in each class to receive the same grade. The teacher explained that the problem is to find a way to determine who did better when one cannot simply use percentage correct. The teacher explained that “grading on a curve” was a way to solve this problem, and that it did not simply mean shifting the scale so everybody does better (a common misconception among students).

In the tell-and-copy condition, the teacher continued on and showed a visual solution to the problem. This solution involved finding the mean deviation and then marking out deviation regions on a histogram to determine which region a student falls into. This is the visual equivalent of finding how many Z-scores or deviation units an individual is from the mean. In the invention condition, students did not receive any instruction for how to solve the problem of comparing scores from different distributions. Their lesson simply ended with the challenge of comparing scores from different distributions.

Afterwards, the students in both treatments received raw data sets that required comparing individuals from different distributions to see who did better. For example, in one activity, the students had to decide if Bill broke the high-jump world record more than Joe broke the long-jump record, given scores for the top high-jumps and long-jumps that year. In the tell-and-copy treatment, students worked with data to use the procedure they had just learned. If necessary, the students received corrective feedback whenever the teacher walked by their desk. The goal was to help them develop mastery of the visual procedure. In the invention condition, students had to innovate their own way to solve this problem of comparing incommensurable performances. They did not receive any feedback. The goal was for them to further notice the problem of incommensurables and work to reconcile it. By hypothesis, this would create the earlier form of knowledge that would prepare them to learn.

To compare whether the invention or the tell-and-copy conditions better prepared the students to learn, the study included a second factor that controlled whether the students received a worked-out example embedded in a posttest given a few days later. The worked example showed how to compute standardized scores given descriptive measures only (e.g., means, variances), rather than using raw data as in the classroom activities. The example showed how Cheryl determined if she was better at the high dive or low dive. The students had to follow the example to determine if Jack was better at high jump or javelin. Half of the students from each instructional condition received this worked example as part of their posttest and nearly every student who received the worked example followed it correctly, demonstrating excellent mastery. The question was whether the students would learn what this worked example had to offer, or whether they would simply see it as a "plug and chug" problem. To detect if students learned from the worked example, there was a target transfer problem several pages later in everybody's posttest. Like the worked example, the target transfer problem required calculating standardized scores from descriptive measures rather than raw data, but used a different context (e.g., comparing homerun hitters from different eras).

[Figure 5 about here]

Figure 5 shows the percentages of students who solved the target transfer problem. Students who invented their own methods for standardizing data learned from the worked example embedded in the test and spontaneously transferred this learning to solve a novel problem. They performed better on the transfer test than inventing students who did not receive the embedded resource. They also performed better than students in the tell-and-copy condition, who exhibited no benefit when they had the embedded resource. A closer look at Figure 5 helps to locate the effect more precisely. An answer to the transfer problem was considered correct when a student gave a quantitatively precise answer or when a student gave a good qualitative or visual answer. Figure 5 shows that the inventing students used the quantitative technique provided in the worked example more than three times more often than the other conditions.

The study provides two bits of information. The first is that noticing and working to reconcile incommensurables can develop the earlier knowledge that prepares people to learn mature canonical solutions. Interestingly, the invent students were able to transfer across the different surface forms of the initial classroom instruction, the worked example, and the target transfer problem. Across this study and a subsequent replication, there was no evidence that surface form or problem topic influenced transfer. Evidently, there are other ways to promote transfer besides having people master a full-blown procedural schema by abstracting from multiple cases or repetitively applying it in multiple contexts (e.g. Gick & Holyoak, 1983). In particular, working to reconcile incommensurables appears to help students learn the high-order structure that makes problem solving possible and it enables them to see that structure in new situations.

The second bit of information is that dynamic assessments that include resources for learning during a test can be sensitive measures of earlier forms of understanding. These forms of understanding are insufficient for independent problem solving – the students who did not receive the embedded worked example did not do very well. Even so, these earlier forms of knowledge can support the acquisition of more mature forms of understanding. Such PFL tests indicate what is within the students' zone of proximal development, and they help reveal the value of instructional methods that are overlooked by tests of independent problem solving. The two forms of instruction in our study, inventing measures versus tell-and-copy, would have appeared equivalent had we not included the resource item from which students could learn.

The value of PFL assessments is further amplified by considering college students. Above, we reported that college students who had taken a full semester of statistics utterly failed to compute a measure of variability. By this measure of mastery, their semester of statistics was useless. However, this is a mismeasurement. Figure 6 presents a different story. We gave these same college students the assessment that we gave to the 9^{th}-graders in the same testing format. We also included a sample of college students who had not taken any statistics. Although the college students performed more poorly than the 9^{th}-graders, they were able to take advantage of the worked example in the test to learn the quantitative technique, and students who had taken a semester of statistics were most likely to learn from this resource. So, even though the college students failed to show mastery on basic statistical procedures, they were able to learn given a worked example. We assume this is a consequence of their college experiences and the statistics course, although it is always possible that the ability to learn well from worked examples is what got them into college in the first place. In either case, PFL assessments can reveal important forms of knowledge that tests of brute memory or sequestered mastery can miss.

[Figure 6 about here]

CONCLUSIONS

In this chapter, we have tried to reconsider the role of prior knowledge in learning. The importance of prior knowledge in learning is well-documented, but existing models tend to either presuppose the prior knowledge (intuitions) or to ignore prior knowledge that does not support independent problem solving (mastery). We think that it can be problematic to leave prior knowledge untouched in situations where prior knowledge is inconsistent. Intuitions sometimes contain tacit conflicts that, if left unresolved, become shaky foundations for further learning. For example, we described 12-year-olds whose notion of chance and cause as opposites made it difficult for them to understand random sampling. Mastery of independent problem solving routines, while a critical part of building expertise, can yield inflexible knowledge that cannot adapt to variations in problem context and new learning situations. High school students who mastered the mean deviation formula but did not have insight into its properties (why the formula divides by 'n') were less able to adapt their knowledge to a bivariate problem. We suspect that methods emphasizing intuition or mastery are insufficient because they do not help students notice and reconcile incommensurables in their prior knowledge, especially in a domain like statistics where incommensurables abound.

Borrowing from Piaget's notion of disequilibrium and one of Plato's resolutions to the learning paradox, we proposed a different approach to prior knowledge. We shifted the focus of instruction to help learners develop "early" knowledge that they did not already have, and we shifted assessment procedures away from independent performances towards evaluating a learner's zone of proximal development. Our proposal is that noticing and reconciling incommensurables prepares students to understand canonical solutions when given an opportunity to learn, and the best way to assess this readiness to learn is to use measures of preparation for future learning.

Our instructional model (Designs for Knowledge Evolution, Schwartz, Martin, & Nasir, 2005) incorporates three key elements. First, we use contrasting cases to draw students' attention to key features that produce the incommensurables, such as different sample sizes in the pitching grid task. Second, we provide students a task structure and set of cultural tools (e.g., mathematics, graphing) that help them try to reconcile incommensurable elements. For example, we showed that young children learned about balance when they tried to apply math to understand the behavior of the balance scale. Third, we teach students the expert solutions that they are now in better position to understand at a deep structural level.

To bring these points together, we described an experiment with 9th graders that showed the benefit of inventing compared to tell-and-copy activities. Students who tried to innovate a standardization procedure for comparing incommensurable data sets were better prepared to learn from a worked example in a posttest and to spontaneously transfer this knowledge to a target problem later in the test. Notably, without a learning resource in the test, the inventing students looked the same as students who received the more mastery-oriented instruction. The invention activity created an earlier form of knowledge that did not support independent performance but did provide the knowledge necessary to learn.

Knowing which forms of early knowledge are useful and how to develop them seems like an important goal of the learning sciences that has been neglected by both the developmental and cognitive science literatures. The developmental literature emphasizes maturation over knowledge, so the question is rarely what specific instructional conditions can develop early readiness. This is Vygotsky's allure to educators. His work provides a developmental perspective that emphasizes the value of particular circumstances for learning – circumstances that educators can sometimes control. The cognitive science literature and the derivative educational literature have largely examined the forms of knowledge and concepts that support mature and expert performance. Perhaps because they are hard to assess with standard instruments, cognitive science has paid less attention to the role of earlier forms of knowledge. However, we think that assessing early knowledge with PFL instruments can be accomplished with some simple modifications to standard procedures.

For example, a standard transfer paradigm has students learn by one of two methods and then measures their success on a target transfer problem. In our final experiment with 9th-graders learning how to standardize scores, we made a simple modification that created a "double transfer" paradigm. Students needed to spontaneously "transfer in" what they learned from the class activities to learn from the worked example embedded in the posttest. Then, they needed to spontaneously "transfer out" this new learning to solve the target transfer problem later in the test. In contrast to most transfer experiments, the manipulation of interest was not how we presented the target procedure itself (which was the same worked example across conditions), but rather, how we differentially prepared students to learn from the worked example. We propose this "double transfer" instrumentation is a more ecologically valid measure of transfer, where the transfer of one's earlier knowledge determines what one learns, and what one learns determines what is transferred to solve a subsequent problem.

We suspect there are other ways to assess preparation for future learning, and expanding our repertoire of these instruments will enable us to further reconsider the possibilities for developing the earlier knowledge that prepares people to learn. We have chosen one form of early knowledge here that we think is particularly important for learning higher-order knowledge, because many higher-order concepts in the sciences solve the problem of how to relate ideas that are otherwise incommensurable (e.g., mass and velocity). And as we have tried to indicate, statistics is filled with incommensurables that catch people unaware. Nisbett, Krantz, Jepson, Kunda (1983) demonstrated that people can learn statistical concepts with relatively simple instruction when those concepts work in accordance with people's prior knowledge. In light of that, we could limit our instructional aspirations to those topics for which people have good prior knowledge. Or, we could develop the right early knowledge to start with.

Acknowledgments

We thank Janet Go for her last minute heroics. This material is based upon work supported by the National Science Foundation under Grants No. BCS-021458 & REC-0196238. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Bar-Hillel, M. (1980). What features make samples seem representative? Journal of Experimental Psychology: Human Perception and Performance, 6, 515-550.

Bransford, J. D. & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad and P. D. Pearson (Eds.), Review of Research in Education, 24, 61-100. Washington, D.C.: American Educational Research Association.

Bransford, J. D., Franks, J. J., Vye, N. J. & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds.), Similarity and analogical reasoning (pp. 470-497). NY: Cambridge University Press.

Chi, M.T. H., de Leeuw, N., Chiu, M-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18, 439-477.

CTGV [Cognition and Technology Group at Vanderbilt]. (1997). The Jasper Project: Lessons in curriculum, instruction, assessment, and professional development. Mahwah, NJ: Lawrence Erlbaum Associates.

Feuerstein, R. (1979). The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques. Baltimore: University Park Press.

Gardner, H. (1982). Art, mind, and brain: A cognitive approach to creativity. New York: Basic Books.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment. Psychological Review, 62, 32-51.

Gick, M. L., & Holyoak, K. J. (1983). Schema Induction and Analogical Transfer. Cognitive Psychology, 15, 1-38.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1983). Judgment under uncertainty: heuristic and biases. NY: Cambridge University Press.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of the fraction concept. Cognitive Science, 29, 587-625.

Marton, F., & Booth, S. (1997). Learning and awareness. Mahwah, NJ: Erlbaum.

Minstrell, J. A. (1989). Teaching science for understanding. In L. B. Resnick & L. E. Klopfer (Eds.), Toward the thinking curriculum: Current cognitive research (pp. 129-149). Alexandria, VA: ASCD

Nisbett, R. E., Krantz, D., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. Psychological Review, 90, 339-363.

Piaget, J., & Inhelder, B. (1975). The origin of the idea of chance in children. (L. Leake, Jr., P. Burrell, & H. D. Fischbein, Trans.) NY: Norton.

Plato (1961). Collected Dialogs of Plato Including the Letters. (E. Hamilton & H. Cairns, Eds.) Princeton, NJ: Princeton University Press.

Schwartz, D. L., Goldman, S. R., Vye N. J., Barron, B. J., & CTGV. (1998). Aligning everyday and mathematical reasoning: The case of sampling assumptions. In S. Lajoie (Ed.) Reflections on Statistics: Agendas for learning, teaching and assessment in K-12 (pp. 233-274). Mahwah, NJ: Erlbaum.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for learning: The hidden efficiency of original student production in statistics instruction. Cognition &

Instruction, 22, 129-184.

Schwartz, D. L., Martin, T., & Nasir, N. (2005). Designs for knowledge evolution: Towards a prescriptive theory for integrating first- and second-hand knowledge. In P. Gardenfors & P. Johansson (Eds.), Cognition, Education, and Communication Technology (pp. 21-54). Mahwah, NJ: Erlbaum.

Schwartz, D. L., Martin, T., & Pfaffman, J. (2005). How mathematics propels the development of physical knowledge. Journal of Cognition and Development, 6, 65-88.

Schwartz, D. L., Bransford, J. D. (1998) A Time for Telling. Cognition & Instruction, 16, 475-522.

Schwartz, D. L., Bransford, J. D., & Sears, D. A. (2005). Efficiency and innovation in transfer. In J. Mestre (Ed.), Transfer of learning from a modern multidisciplinary perspective (pp. 1-52). Greenwich, CT: Information Age Publishing.

Vygotsky, L. S. (1987). The collected works of L. S. Vygotsky. (R. Rieber & A. Carton Eds.) NY: Plenum.

Footnotes

1. To our knowledge, the instructional possibility of developing earlier knowledge to expand children's zone of proximal development is not something Vygotsky addressed. Vygotsky was working on the proposal that culturally-derived ("scientific") ideas can cause development, and therefore, he wanted to show that the introduction of these ideas in the zone of proximal development pushed a child's development forward. "*Instruction is only useful when it moves ahead of development.* When it does, it *impells [sic] or wakens a whole series of functions that are in a stage of maturation lying in the zone of proximal development.*" (p. 212, Vygotsky, 1934/1987, italics in original).

Figure Captions

Figure 1. Vygotsky's formulation of the Zone of Proximal Development and two different interpretations of where to look for instructional implications.

Figure 2. Twelve-year-old children have many intuitions that exhibit family resemblances to sampling, but none works quite right. (Schwartz et al., 1998, adapted with permission).

Figure 3. Students' abilities to learn how to compare covariance without instruction. Results are broken out by student mastery of a formula for computing univariate variability and their insight on how the univariate formula resolves problems of making different sample sizes commensurable.

Figure 4. Contrasting cases for noticing and dealing with incommensurables in variability. Students invent a reliability index. (Schwartz & Martin, 2004, adapted with permission).

Figure 5. Design of assessment experiment and results for 9th-graders.

Figure 6. Performance of college students with and without college statistics who received the same test as the 9th-graders.

Figures

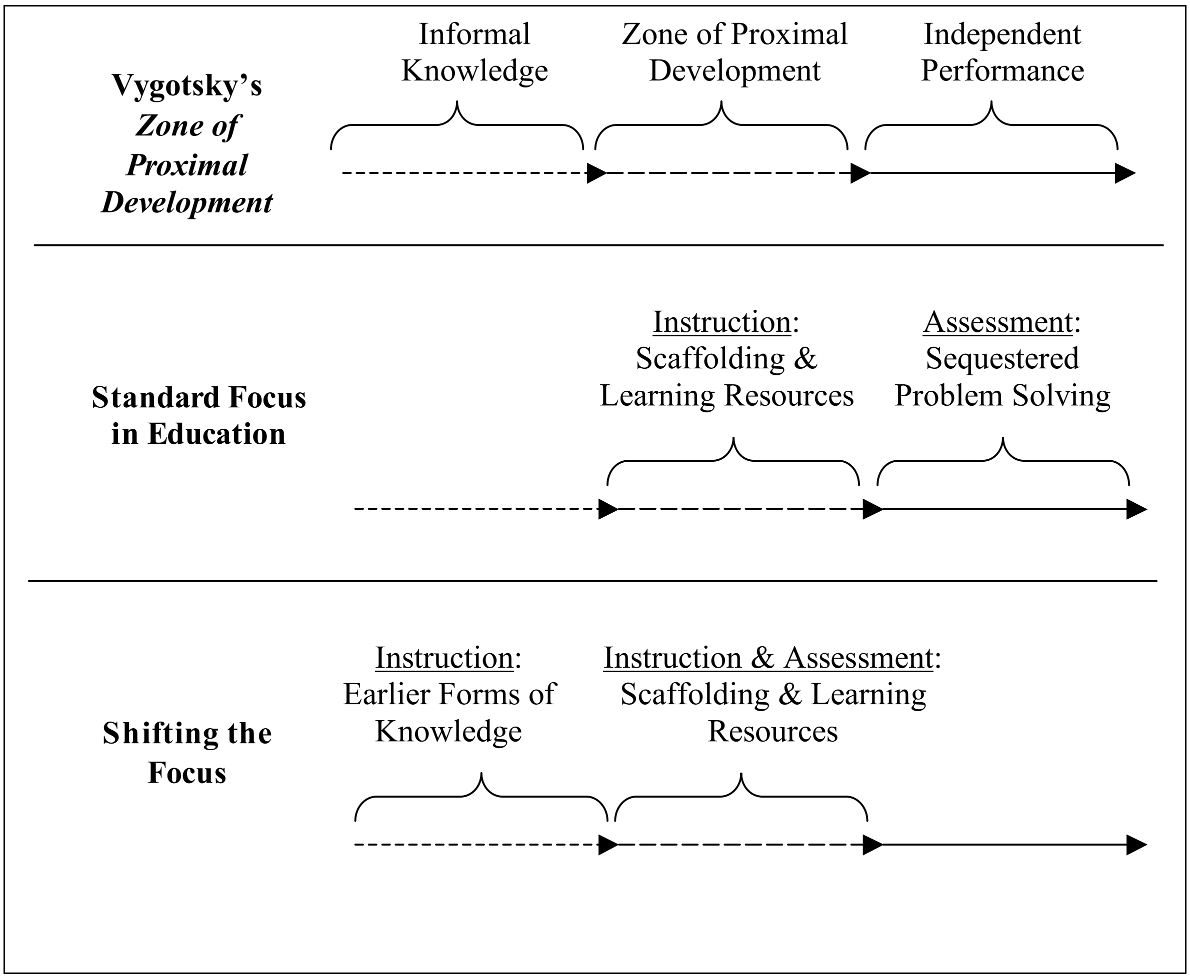

Figure 1

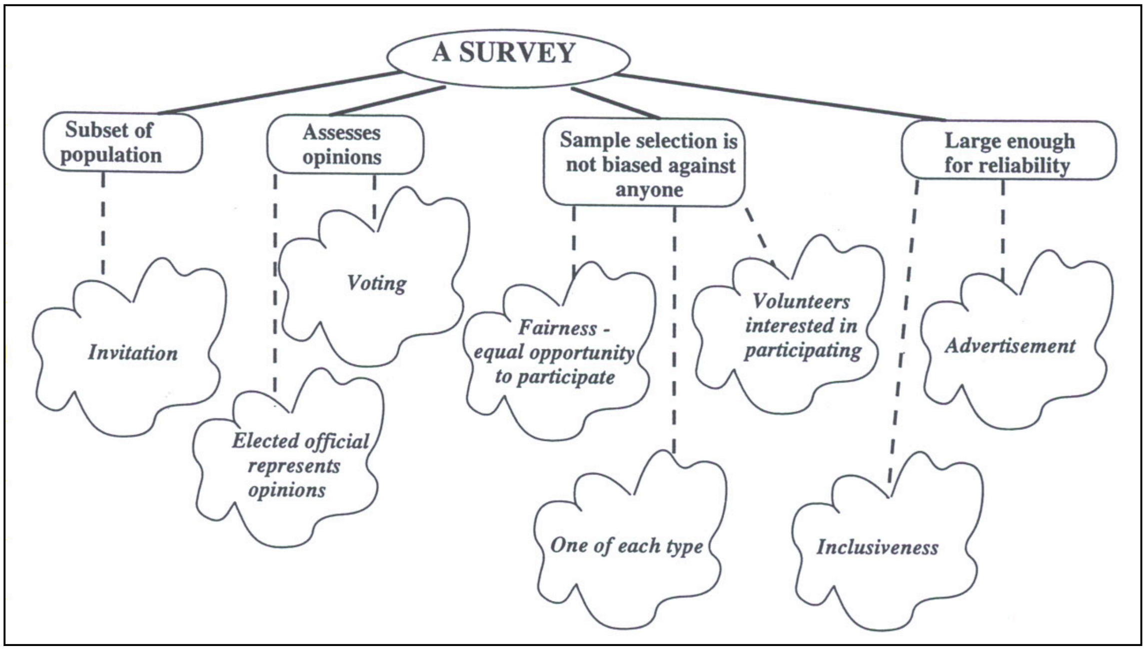
A SURVEY
Subset of population
Assesses opinions
Sample selection is not biased against anyone
Large enough for reliability
Invitation
Voting
Elected official represents opinions
Fairness - equal opportunity to participate
One of each type
Volunteers interested in participating
Inclusiveness
Advertisement

Figure 2.

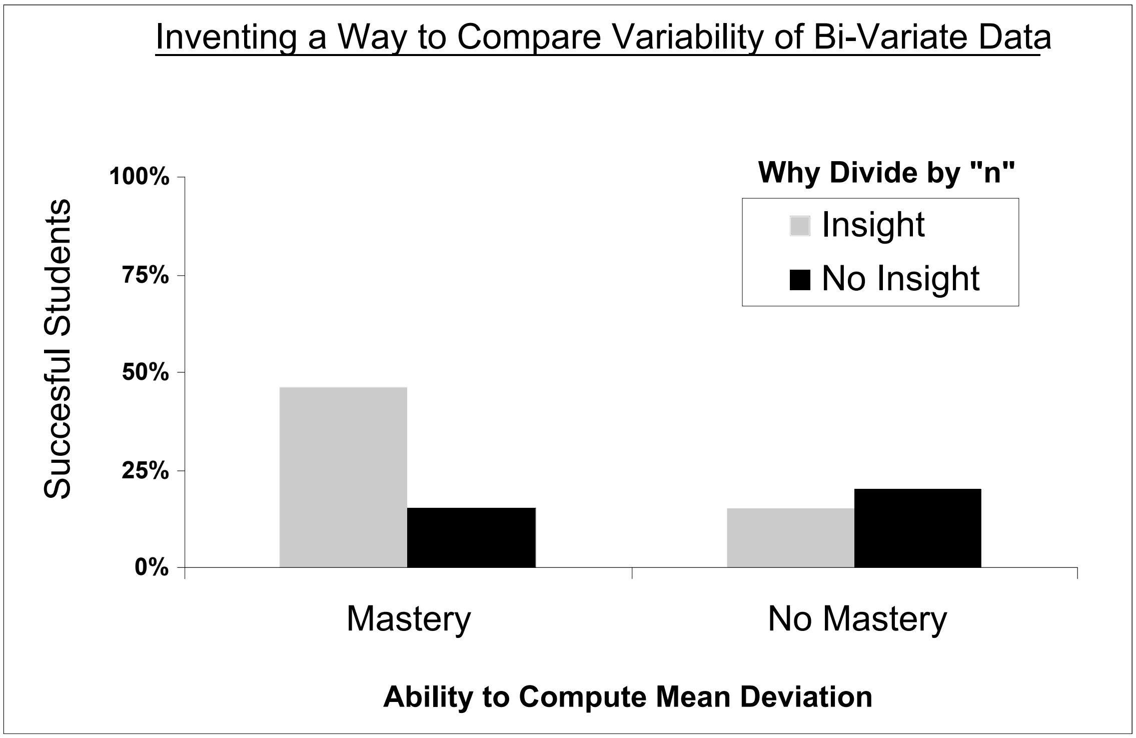
Inventing a Way to Compare Variability of Bi-Variate Data
Succesful Students
100%
75%
50%
25%
0%
Why Divide by "n"
Insight
No Insight
Mastery
No Mastery
Ability to Compute Mean Deviation

Figure 3

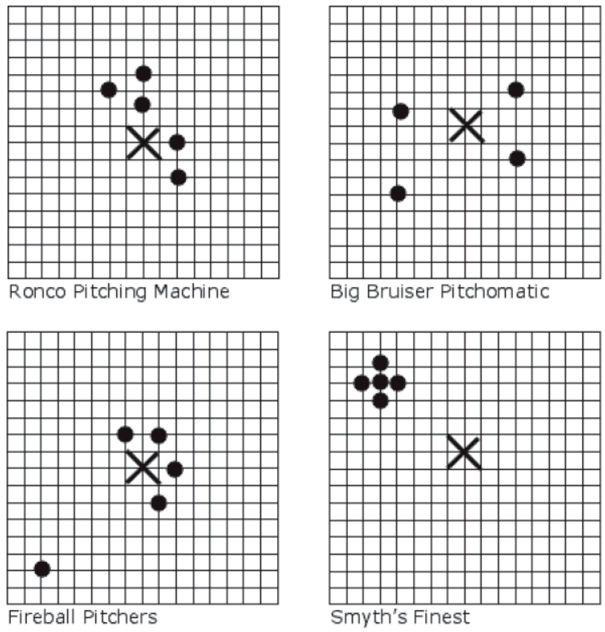
Ronco Pitching Machine
Big Bruiser Pitchomatic
Fireball Pitchers
Smyth's Finest

Figure 4.

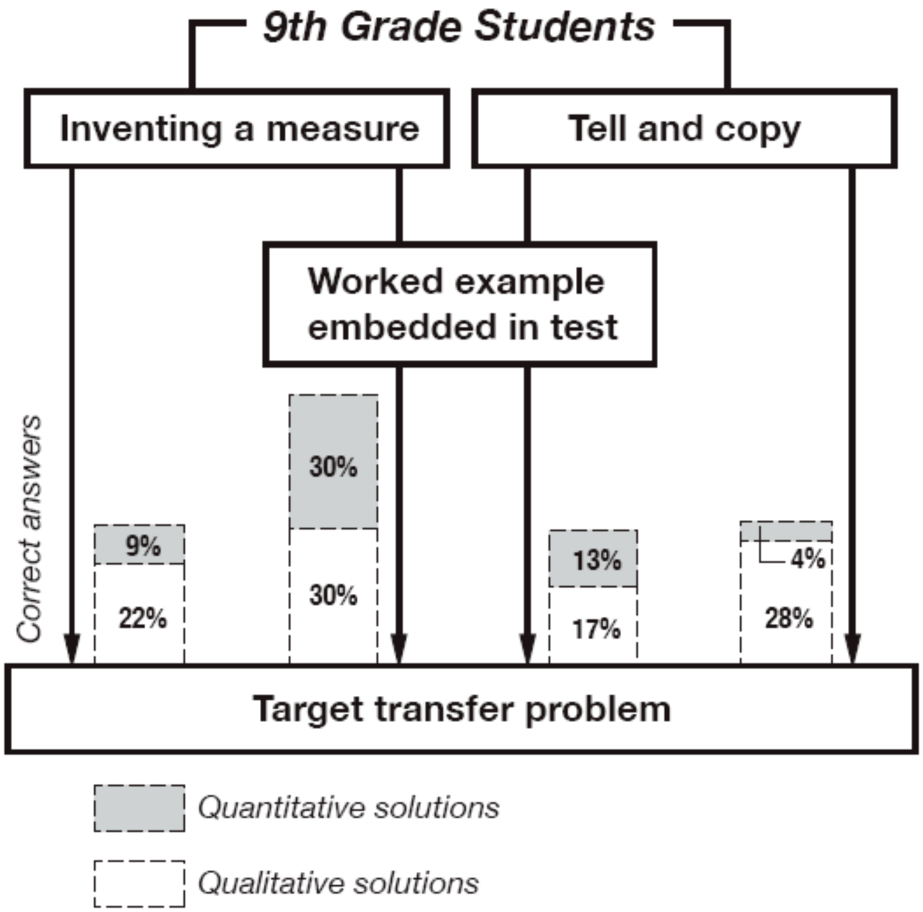
9th Grade Students
Inventing a measure
Tell and copy
Worked example embedded in test
Correct answers
9%
22%
30%
30%
13%
17%
4%
28%
Target transfer problem
Quantitative solutions
Qualitative solutions

Figure 5.

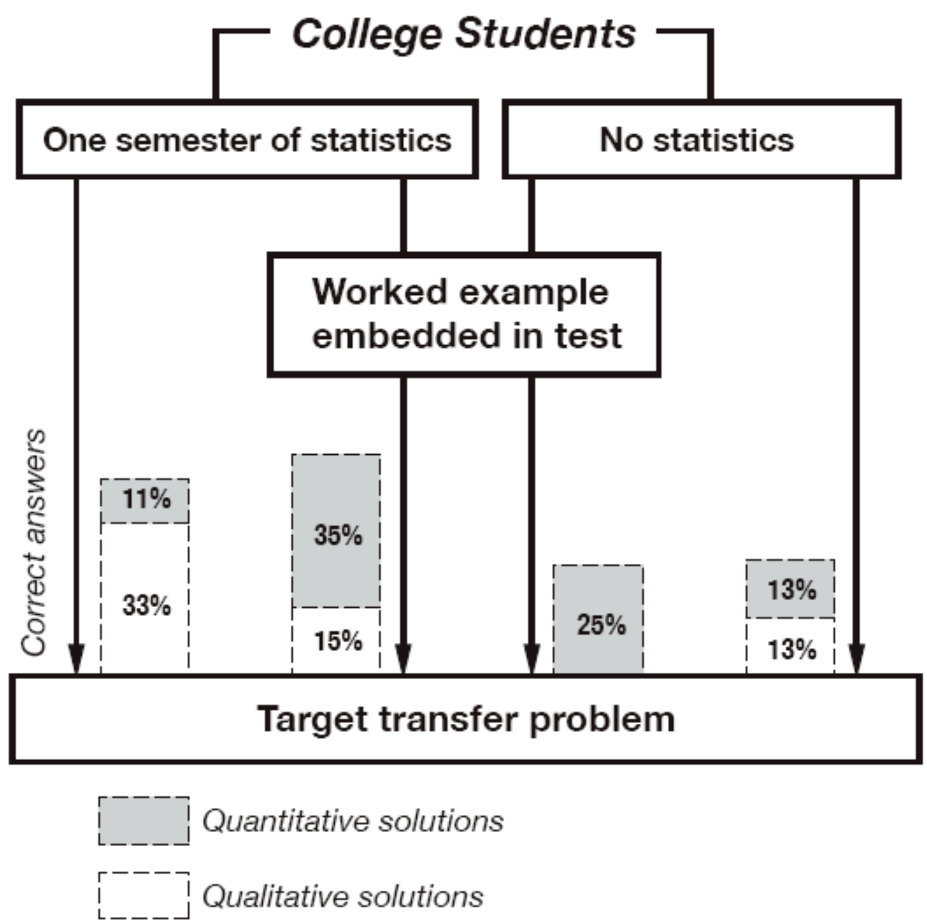
College Students
One semester of statistics
No statistics
Worked example embedded in test
Correct answers
11%
33%
35%
15%
25%
13%
13%
Target transfer problem
Quantitative solutions
Qualitative solutions

Figure 6.