Standards & Assessments: The Role of Depth versus Breadth in Student Success

Dr. Marc S. Schwartz
University of Texas, Arlington

Dr. Theo Dawson
Lectica, Inc

Introductiona Standards vs. the Archer's Bull'sa Eye

Standards matter. Yet as true as this statement is, it belies a more complex story, nuanced by context, and influenced by important variables. Imagine such a story—you are standing in front of a target, the red bull's eye plainly in sight. Bold, black concentric rings surround the unblinking eye; your bow is in hand. Today you are using a heavier and stiffer arrow than you have in the past; the coach said it was more accurate. The stakes are unusually high because you are representing your school at the state finals. Even without additional details, this scenario conveys a rich story about the human experience. No one has to explain the goal and its i mportance. Soon you will coordinate numerous actions and thoughts in a dynamic environment to launch the arrow to its destination. Seconds later there will be clear, unmistakable feedback. Success in this moment is the result of knowing what is required, having practiced enough to master the necessary actions, and having the opportunity to continually adjust those actions in light of the current context.

Archery is a useful metaphor for learning and suggests an analogous narrative for s tandards in education. Yet, we wonder how often educational standards actually serve as a clear call for action? How often do students and teachers have opportuni ties to interpret and adjust those actions to improve success? What kind of support do standards provide for learning?

We persist in setting educational standards because we believe they serve as clear, measurable goals that can improve educational outcomes. Yet close to two decades into the process of creating standards, their impact on student learning is unclear. Part of the reason may be a disconnect between what scientists know about how students learn and the combined effects of the standards we set and the tests that are designed to measure conformance with these standards.

Goald Actiond Feedback: The foundational "feedback loop" of life and learning

What any archer understands is that success depends on practice. Scientists agree on a general model of learning represented here as the goal-action-feedback loop (although they offer a more sophisticated account of this process). This loop contains four essential variables: goal, action, feedback, and *repeat*. When integrated dynamically they help the archer focus on the goal, choose an appropriate action, interpret the outcome, and decide how to correct or adjust the current available pool of actions and ideas to better reach the goal.

Piaget characterized the feedback loop as our ongoing attempt to achieve balance between assimilation and accommodation.[i] Later, Powers arrived at a similar view, articulated in a model he called Perceptual Control Theory.[ii] The idea is simple to describe, even if difficult to execute in educational settings: We select and execute the actions that best fit our prediction of how to achieve a desired goal.[iii] More recently, neuroscientists have observed a similar feedback loop. Neurons in the neocortex act in cohorts where repeated stimulation of the network in a variety of contexts fine-tunes the emergence of specific patterns.[iv] [v] The evidence is clear. Whether viewing the human experience from the perspective of neurons or behavior, success depends on practice.

Depth vs. Breadth: Balancing time and focus to inform goals & action

Practice is necessary for learning. This is well understood by most educators. However, the time required to cycle through the feedback loop to consistently and accurately attain a target is limited in the classroom by the sheer number of goals teachers are required to address. Therein lies the tension between *quantity* and *quality* that is at the heart of a centuries old educational debate about depth versus breadth.

The archer's success depends on time, practice, coaching, and the opportunity to learn from mistakes. This is no less true for students learning Newton's laws or any teacher attempting to teach them. When students have an opportunity to cycle repeatedly through the learning feedback loop, they are more likely to achieve a deep understanding of the subject matter. But there is so much to know, so much material to cover.

During the last 20 years, the balance between depth and breadth has shifted toward breadth. Our educational standards are sweeping. Our standardized tests follow suit.

---

[i] Piaget, J. (1983). Piaget's theory. In P. M. Mussen (Ed.), Handbook of Child Psychology (Vol. 1, pp. 103–128). New York: Wiley.

[ii] Powers, W. (1973). Perceptual Control Theory. Hawthorne, NY: Aldine DeGruyter.

[iii] Powers, W. (1998). Making Sense of Behavior—The Meaning of Control. New Canaan, CT: Benchmark Publications Inc.

[iv] Mountcastle, V. B. (1998). Perceptual Neuroscience: The Cerebral Cortex. Cambridge, MA: Harvard University Press.

[v] Hawkins, J. and S. Blakeslee (2004). On Intelligence. New York: Times Books.

Not only does this trend fly in the face of everything we know about learning[vi], recent research provides hard evidence that students may be paying a heavy price.

Standards- Historical Perspective

Educators in the United States have a long history of discussing and documenting the student's struggle to understand science. Much of the early analysis of the problem and the recommendations that followed can be traced to the introduction of textbooks, the earliest of which appeared at the end of the 19th century. Jackman, in 1894, noted with flair, "The mechanics of the lever is only so much 'stuff and nonsense' to him, until he finds that this knowledge will render possible an economy of his energy, and thereby immensely prolong his mortal existence." [vii]

While a century later the prose is less poetic, it is no less poignant. In the 1980's the American Association for the Advancement of Science (AAAS) brought scientists and educators together to map out a new vision for science education in *Science for All Americans*.[viii] During this same decade Texas was also seeking strategies to better support learning. In response to a 1981 legislative mandate, the Texas State Board of Education approved a curriculum-reform package intended to increase high-school graduation requirements and set statewide standards for the "essential elements" of 13 subjects at various grade levels.[ix] These early standards were the precursors to today's more demanding K-12 state standards, "Texas Essential Knowledge and Skills."

At the national level, similar standards-based reform movements were also underway. G*oals 2000: Educate America Act* was signed into law in 1994, providing states with the resources to improve graduation rates and support the development of state standards. The AAAS continued to reshape the philosophy and pedagogy of science instruction with publications such as *Benchmarks for Science Literacy* and

---

[vi] National Research Council (2000). How People Learn: Brain, Mind, Experience, and School. Washington, DC: National Academy Press.

[vii] Jackman, W. (1894). Nature Study for the Common Schools. New York: Henry Holt. p 6.

[viii] American Association for the Advancement of Science (AAAS). (1990). Project 2061: Science for all Americans. New York: Oxford University Press.

[ix] Bridgman, A (1984). Texas board approves 'essential elements for schools' curricula. Education Week (Feb. 22, 1984).

The Atlas of Science Literacy.[x] [xi] These later publications and others like the *National Science Education Standards* provided a synthesis of the central understandings achieved in science.[xii]

Today, the national standards movement continues with the development of common core standards in Language Arts and Mathematics. The *Common Core State Standards Initiative* is a state-led effort coordinated by the National Governors Association Center for Best Practices and the Council of Chief State School Officers.[xiii] In 2009 Texas released for public comment the college-readiness standards.[xiv] In 2010 the Texas Education Association plans to release K-12 standards. Both sets of standards, the College and Career Readiness and the Common Core State Standards, will complete a set that should guide teachers from kindergarten into graduate school.

What counts as a standard still faces the ongoing challenge of finding a negotiated balance between competing views. But this dialogue is important. Shepard and colleagues at the National Academy of Education noted, "For many states, the content standards adopted a decade ago represented that state's first effort at trying to develop some kind of curriculum framework. For most, the process was highly political—as well it should be in a democratic society. But without previous experience and access to coherent curricula representing particular curricular perspectives, the political solution of adding in everyone's favorite content area topic created overly-full, encyclopedic standards in some states, or vague, general statements in others."[xv]

Trying to incorporate the philosophy of standards in the classroom is as much a challenge to teachers as developing the standards is for policy makers and for students to understand those standards. As noted in, "*What is the Influence of the*

---

[x] American Association for the Advancement of Science (AAAS). (1993). Benchmarks for Science Literacy. New York: Oxford University Press.

[xi] American Association for the Advancement of Science (2001). Atlas of Science Literacy. Washington, DC: American Association for the Advancement of Science.

[xii] National Research Council (1996). National Science Education Standards. Washington, DC: National Academy Press.

[xiii] www.corestandards.org

[xiv] Texas Higher Education Coordinating Board & Texas Education Association (2009). Texas College and Career Readiness Standards. Austin, TX: University of Texas Press. http://www.thecb.state.tx.us/

[xv] Shepard, L., Hannaway, J., and Baker E. (eds.) (2009). National Academy of Education, Education Policy White Papers Project: Standards, Assessment, and Accountability.

National Science Standards?" teachers have dramatically different views about what the standards mean and how they should guide pedagogy.[xvi]

A Cognitive and Neuroscience Perspective

Earlier we introduced Perceptual Control Theory, a cognitive model that describes learning in terms of a feedback loop that informs individual progress on the way to reaching specific goals.[xvii] In this model, an effective goal focuses attention and empowers the learner to choose specific strategies of action. The actions chosen are those the learner predicts will lead to success. Recognizing what counts as success creates the context that helps students analyze the feedback generated by their actions. From the perspective of this model, educational outcomes that do not empower students to identify and organize actions will function more like wishes than goals.

A more recent cognitive model that builds on the work of Piaget and Powers is Fischer's dynamic skill theory, which suggests that the complexity in our thinking is shaped by the interaction between changing contexts and progress in brain development.[xviii] Improvement in cognitive performance reflects more than the incremental accumulation of facts and ideas. Knowledge is organized hierarchically, with complexity increasing in stable, measurable levels. Each level subsumes the one below, and each level is qualitatively more complex than the level that serves as its foundation. Each successive level can be thought of as a more complex and integrated way of thinking. Learning within a level requires the integration of new knowledge with existing knowledge. Moving from one level to the next requires the elaboration of knowledge at the current level to a point where the system reorganizes and a new way of thinking emerges. Typically, four such reorganizations take place during PK-12.

Neither of these models suggests that we should abandon breadth for depth—both are required for development—but they do suggest that there is likely to be an optimal balance between accumulating and integrating knowledge, one that sends

---

[xvi] National Research Council. (2003). What Is the Influence of the National Science Education Standards? Reviewing the Evidence, A Workshop Summary. Karen S. Hollweg and David Hill. Steering Committee on Taking Stock of the National Science Education Standards: The Research, Committee on Science Education K-12, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

[xvii] Powers (1973); (1998)

[xviii] Fischer, K. W., and T.R. Bidell (2006). Dynamic development of action and thought. In W. Damon & R. M. Learner (Eds.), Handbook of child psychology: Theoretical models of human development (Vol. 1, 6th ed., pp. 313–399). New York: Wiley.

students into the world with the breadth and depth they need to participate fully in society.

Depth vs. Breadth- Supporting the learning feedback loop

Both cognitive and neurological models suggest that depth of study would pay dividends, if only to provide the time to practice juggling and coordinating the core ideas implicit in each science standard. In the only large-scale study that tests this hypothesis, Dr. Schwartz and his colleagues surveyed over 8,000 freshmen at 55 randomly selected post-secondary institutions in the United States at the beginning of their first college course in biology, physics or chemistry.[xix] The course professor collected the surveys, and added the student's course grade at the end of the semester. The survey questioned a number of variables regarding the student's high school experience, but pertinent to our hypothesis was the amount of time students spent studying any one topic in high school science. The findings make a compelling case for providing students with time to go deeper into individual topics.

Students who reported that their high school teachers spent two or more weeks investigating a single topic performed, on average, about one to two points higher in their college science course than students who did not. While this increase does not sound like much, Schwartz and his colleagues also measured the impact of students taking an additional year of the same subject in high school.[xx] Depending on the discipline, an additional year of study predicted a two to four point increase in the college course grade. Spending two or more weeks on a single topic in an introductory high school course was equivalent to roughly a quarter of a year of additional study in chemistry, a half-year in chemistry, and almost two thirds of a year in physics.

More surprising was the relationship between breadth of study and the final college grade. Students in high school science courses that focused exclusively on breadth, received, on average, final grades that were about one point lower than those who had experienced depth. In biology this correlation represents a loss of half a year of study. This would be akin to ending the school year (for biology students) in the middle of March.

Researchers undertook this study because they were concerned about an increasing trend for high school teachers to focus on breadth over depth in their efforts to cover standards and prepare students for high stakes standardized tests. The results suggest that while standards highlight the outcomes we desire for students, it is less clear that they are helping teachers plan for these outcomes or prepare students to

---

[xix] Schwartz, M. S., Sadler, P. M., Sonnert, G., & Tai, R. H. (2009). Depth versus breadth: How content coverage in high school science courses relates to later success in college science coursework. Science Education, 93, 798–826.

[xx] Schwartz, M. S. et al. (2009).

realize them. A growing literature suggests that the current experience of students in schools does not help students create the rich understandings summarized in the standards.[xxi] [xxii] In fact many students, including many whose test scores qualify them for admission to our best colleges, leave school with ideas that are very different from those we intend to teach.[xxiii] [xxiv] If students are to profit from standards, then teaching must support the kind of learning that allows students to build accurate and robust understandings. To accomplish this outcome, teachers and students need time to unpack the complexity embedded in key concepts and skills, and teachers need a way of assessing student progress toward the target that helps them reinforce the feedback loop. Assessments that serve this role not only tell teachers how far students are from the target, but also what they are most likely to benefit from learning next on their way to the target. These are important challenges that educators, researchers, and policy makers must face together.

Current Dilemmas

The educational system faces several dilemmas. First, as a consequence of sweeping state standards and the high stakes testing being done to satisfy the requirements of No Child Left Behind (NCLB), many teachers report that they have tipped the breadth-depth balance toward breadth.[xxv] The standards, along with the tests designed to evaluate their implementation, have become a new educational problem. There is simply not enough time to encounter, much less uncover the complex patterns embedded in the many important concepts and skills science has generated.[xxvi]

In the Texas Essential Knowledge and Skills 2010, for example, the Integrated Physics and Chemistry Course includes 13 process skills and 27 science concepts. Students are required to *know, organize, describe, recognize, demonstrate, investigate*

[xxi] National Research Council. (2003).

[xxii] National Research Council (2000).

[xxiii] Halloun I., and D. Hestenes (1985). The initial knowledge state of college physics students. American Journal of Physics, 53, 1043-1055.
1056–65

[xxiv] Brown, D. (1990). Students' concept of force: the importance of understanding Newton's third law. Physics Education, 24, 353-358.

[xxv] National Research Council. (2003).

[xxvi] Champagne, A. B., Gunstone, R. F., & Klopfer, L. E. (1985). Effecting changes in cognitive structures among physics students. In L. West & A. Pines (Eds.), Cognitive structure and conceptual change (pp. 163–188). Orlando, FL: Academic Press.

and *analyze* a number of relevant course topics.[xxvii] Each of these verbs places a different cognitive demand on children. Intuitively, the reader might note that the embedded difficulty increases with each successive verb, placing an expanding demand on teachers and curricula to include lessons that support these more sophisticated outcomes. Additionally some of these skills depend on the mastery of other skills (e.g., "Investigate the law of conservation of energy.") and some represent more than one skill (e.g., demonstrate that moving electric charges produce magnetic forces and moving magnets produce electric forces)."

A second important dilemma is the mismatch between what students are prepared to do versus what education leaders want them to do. Standards tend to be based more on the preferences of discipline experts than solid evidence about how and when students best learn particular skills and concepts. This has led to the specification of standards that are out of line with evidence about what students are capable of learning as well as the sequences through which concepts and skills are actually learned.[xxviii]

An enormous body of peer-reviewed research continues to show that students build increasingly adequate and more complex skills and concepts over time, requiring multiple exposures to these skills and concepts.[xxix] [xxx] [xxxi]This is especially true for the highly abstract ideas prized by scientists and educators (which form the basis of science standards). Educators know this intuitively, and the trade journals as well as the research literature continue to offer evidence to this end. But despite the conclusions reached through 100 years of research in the cognitive sciences, the educational system has unintentionally set goals that drive pedagogy in a direction that supports neither learning nor good teaching. While our metaphorical archer's success is still judged by his ability to drive the arrow into the bull's eye, he must now accomplish this task in less time, with less practice.

The legacy of research offers basic principals that can serve as strategic guides for classroom and policy decisions. Students need time to unpack the complicated

---

[xxvii] Texas Education Association (2010). Integrated Physics and Chemistry. *Texas Evaluation of Knowledge and Skills* http://ritter.tea.state.tx.us/rules/tac/chapter112/ch112c.html#112.38

[xxviii] Schwartz, M.S. (2009). Cognitive development and learning: Analyzing the building of skills in Classrooms. Mind, Brain and Education. 3,4 198-208.

[xxix] Case, R. (1998). The development of conceptual structures. In W. Damon (Series Ed.) & D. Kuhn & R. S. Siegler (Vol. Eds.). *Handbook of child psychology: Vol. 2. Cognition, perception and language* (5th ed., pp. 745–800). New York: Wiley.

[xxx] Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. Oxford: Oxford University Press.

[xxxi] Fischer, K. W., and T.R. Bidell (2006).

patterns in the world that are embedded in standards. Teachers, in turn, need the time to create supportive environments that offer students structure and guidance for the "unpacking" enterprise. Furthermore, learning is not a linear process where students accrue more and more ideas like new words on new pages in a dictionary. Learning is a process of re-organizing ideas into increasingly complex cognitive structures. If we rush students through this process, depriving them of the time to build a solid foundation at each level, their cognitive structures will become increasingly inadequate and future learning will suffer. Finally, not only is time an essential variable to student and teacher success, but also the way time is used. Over 100 years of learning science has revealed that organizing knowledge into more complex structures is the basic paradigm for building useable knowledge, yet curricula are often aligned with a pre-scientific notion of learning as the simple accumulation of knowledge. As the TIMSS authors note, countries that excel in science teaching are more likely to employ curricula that acknowledge the hierarchical structure of knowledge than countries—like the United States—that focus more on rote learning.[xxxii]

A third dilemma, from the student's perspective, is that the standards don't appear as relevant patterns. Without the opportunity to experience the context from which relevant patterns are sought and gained, the patterns (or in this case the concepts we want students to understand) presented in class appear more like a disconnected ensemble of items to memorize. The situation is somewhat like being handed a new ball, but not seeing how to coordinate it with the balls we are currently juggling. Given the way that the brain stores memories, new ideas presented in such a way are temporary distractions, stored briefly in memory, and dispensed with after serving a short term function—like taking an end of chapter or course exam.[xxxiii]

Curricula that focus on knowledge quantity often fail to provide students with the time to integrate their current views with those of scientists. In students' mental "laboratories", where the ideas of science are often in competition with their existing ideas, research has consistently shown that existing ideas most often prevail.[xxxiv] This research, once called *misconception* research because of its focus on students' misconceptions, also revealed a misconception about learning—that it is the simple accumulation of knowledge. If learning was just about adding new knowledge,

---

[xxxii] Schmidt, W.H., Wang, H.C., & McKnight, C.C. (2005). Curriculum coherence: An examination of U.S. mathematics and science content standards from an international perspective. *Journal of Curriculum Studies*, *37*(5), 525-559.

[xxxiii] Klingberg, T. (2009). The Overflowing Brain: Information Overload and the Limits of Working Memory. New York: Oxford University Press.

[xxxiv] Pfunt, H., & Duit, R. (1994). Bibliography: Student's alternative frameworks and science education (4th ed.). Keil, Germany: Institute for Science Education at the University of Keil.

students would have little difficulty accommodating ideas that are different from their existing ideas. They would simply replace them. But learning is not so straightforward. Before students enter school they have already spent a lifetime making sense of their experiences. They are already little scientists, identifying patterns that allow them to make predictions about the world. In school, lessons often require a profound restructuring of their thinking. For example, Newton's notion of force (that a force, once applied to an object, sets it permanently in motion unless it encounters a new force) competes directly with their consistent experience that to keep something moving one must continually add a force. The Newtonian perspective is further complicated when physics teachers or textbooks, in an attempt to simplify instruction, ask students to ignore friction—the very thing that explains their real-world experience with moving objects. Faced with such a complicated world, students who are not given the time to confront their personal beliefs and compare them with competing beliefs are unlikely to build robust understandings.

Finally, what we value and what we are measuring are not in harmony. Although we want students to develop rich, robust understandings and good reasoning skills, our standardized assessments focus primarily on the ability to recall facts and apply formulas. The research evidence shows that getting items right on most multiple choice tests neither equates with understanding nor tells us much about students' thinking skills.[xxxv] Facts and formulas are important, but they clearly are not sufficient if we take current knowledge about learning and the brain seriously. The danger that Schwartz and his colleagues highlighted is, "… high-stakes examinations that require recall of unrelated bits of scientific knowledge in the form of facts and isolated constructs will increase the likelihood that teachers will adjust their teaching methodologies to address these objectives. The more often discrete facts appear on high-stakes examinations, the more often we imagine that teachers will feel pressured to focus on breadth of knowledge."[xxxvi]

Historically the testing and research communities have failed to deliver reliable, standardized metrics that measure "rich" or "deep" thinking in educationally relevant contexts. Until recently, we have had to choose between low-reliability, non-standardized, subjective educative assessments in which students show their thinking and understanding, and high reliability, standardized, objective, multiple-choice assessments in which students show the content of their knowledge but do not reveal much about their thinking or depth of understanding.

[xxxv] Sadler, P. M. (2000). The relevance of multiple choice tests in assessing science understanding. In J. J. Mintzes, J. H. Wandersee & J. D. Novak (Eds.), Assessing Science Understanding: A human Constructivist View (pp. 249-278). San Diego, CA: Academic Press.

[xxxvi] Schwartz et al. (2009). p 821,

For high stakes assessments, we primarily employ the latter. This means the assessments that dictate who will graduate, who will qualify for college, who is an adequate teacher, or which district will receive funding, leave out skills and knowledge that are critical for success in adulthood. Moreover, these tests serve only one purpose—selection. They are not diagnostic. The scores tell us little about how we can help students learn or how we can improve instruction. To fill this gap, we need to focus on the role of assessments in the goal-action-feedback loop.

Assessment—closing the learning feedback loop

In its most powerful form testing would act as an extended conversation between student and instructor that continually increases the student's ability to achieve educational goals that require higher levels of cognitive performance. The challenge for educators and those creating exams has been in developing assessments of this nature. In the early 90's there was a great deal of excitement about the possibility of creating diagnostic tests by incorporating open-ended items or creating multiple choice items that were informed by current research on "misconceptions." Unfortunately, methods for coding open-ended items did not produce levels of reliability that would justify their use in high stakes contexts, and very few large-scale assessments were created that made use of the misconceptions research.

Given limits in technology through most of the 20th century, testing companies addressed the general need for assessment primarily with multiple-choice questions. While these assessments were sufficient to measure changes in content knowledge, they were not well suited to measure changes in conceptual understanding.[xxxvii] [xxxviii] [xxxix]

Dawson and her colleagues have confronted this challenge by developing a new approach, employing well-developed theories from the cognitive sciences and advances in technology to improve test design. This approach was guided by six requirements:

1. Using a robust theory of learning and development to inform test development
2. Correlating levels of empirically verified and well-understood learning sequences with test scores

---

[xxxvii] Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research, 75, 27–61*.

[xxxviii] Schoenfeld, A. (Ed.). (2007). *Assessing mathematical proficiency (Vol. 53)*. New York: Cambridge University Press.

[xxxix] Agung, S., & Schwartz, M. S. (2007). Students' understanding of conservation of matter, stoichiometry and balancing equations in Indonesia. *International Journal of Science Education, 29(13),* 1679-1702

3. Requiring students to make connections by thinking through problems and producing arguments
4. Contributing to teacher development by providing teachers with a steady stream of useful information about student learning
5. Tracking student learning over time
6. Calibrating all tests to the same standardized scale

As it turned out, fulfilling the sixth requirement was the key to fulfilling the other five. In general, keeping track of student growth could be accomplished by tracking change in the levels of cognitive complexity students used to address classroom problems. Skill theory, introduced earlier, provided the foundation for a set of methods that allow researchers to document these changes over time and in varying contexts.

A hierarchical view (Skill Theory & Lectical Analysis)

During the last century, many researchers described learning sequences for a wide range of concepts in several knowledge areas. Other researchers, noticing similarities across these sequences, worked to define a generalized sequence. By the end of the last century, the level definitions for the most developed of these generalized sequences, including those of Piaget, Fischer, and other researchers were very similar, suggesting convergence on a single, measurable, developmental dimension. Fischer's Dynamic Skill Scale provided the initial foundation for a generalized developmental scoring system, which came to be known as the Lectical Assessment System (LAS).[xl]

The LAS is a well-validated, accurate, and reliable developmental scoring system that detects approximately 20 distinct "*phases*" of performance between kindergarten and grade 12. *Levels* along the scale represent qualitative changes—changes in the way students reason. Phases represent progress—in the form of increasing elaboration—within levels. The LAS can distinguish degrees of skill as effectively as the best designed standardized tests—detecting from 5 to 7 distinct phases within a single age cohort.

Learning sequences

[xl] Fischer, K. W., and T.R. Bidell (2006).

The Lectical Assessment System is used to study and describe detailed learning sequences for specific concepts.[xli] [xlii] Once a learning sequence is well understood, it can form the basis for assessments, called DiscoTests. These are scored with low inference, evidence-based coding rubrics that students and teachers can easily use. The rubrics consist of a set of menus containing statements that are modeled upon the answers of real students performing in each relevant phase. The number of pull-down menus for a given question corresponds to the number of sub-concepts or skills targeted by the question. Coders simply select the code that is most like a response in the assessment being coded. In addition to using the coding rubrics, both teachers and students can make comments about other aspects of a performance, including things like quality of argumentation, use of terminology, etc.

Rubric selections are translated into scores that inform reports containing targeted feedback and learning suggestions. These reports include:

1. A phase score, calculated on the basis of the rubric selections, and displayed on a developmental "ruler"
2. A description of the way students performing in a given phase generally think about the concepts and skills targeted in the assessment
3. A map tracing student performance on the test subject over time
4. Suggestions for individual or group activities that are designed to help students build the concepts and skills of the next phase
5. Student and teacher comments; and, for teachers
6. A description of what the student is most likely to benefit from learning next, often with specific suggestions for students on somewhat different pathways. (For example, a student can perform better on one cluster of sub-concepts or skills than on another, revealing where possible strengths and weaknesses lie.)

DiscoTest- Illustrations in science

Once the research required to describe a set of learning sequences has been conducted, DiscoTests are made available for classroom use. We are currently working on the development of a wide range of assessments on topics as varied as moral and reflective judgment, decision-making, and self-understanding; however, the assessments on which we have made the most progress deal with three "big

---

[xli] Dawson, T. L., and Z. Stein (2008). Cycles of research and application in education: Learning pathways for energy concepts. Mind, Brain, & Education, 2(2), 90-103.

[xlii] Dawson-Tunik, T. L. (2004). "A good education is…" The development of evaluative thought across the life-span. Genetic, Social, and General Psychology Monographs, 130(1), 4-112.

ideas" in science, including the physics of energy, the nature of matter, and conservation of matter.[xliii]

Like all DiscoTests, the assessments are made up of 5-7 open-ended questions, none of which has a single correct answer. The selection of questions models our current understanding of how knowledge develops, both historically and within individuals. The questions shift the focus from getting the "right" answer, to formulating increasingly adequate answers. They require students to make connections between ideas and show these connections in written arguments. Making connections and writing about them improves student writing and thinking, increases the depth of students' understanding of course content, and provides teachers with information that can be used to diagnose specific strengths and weaknesses.

The use of a single, non-arbitrary standardized scale—*the skill scale*—provides DiscoTest scores with a number of properties that other test scores do not have. For example, all DiscoTests are scored on the same scale. All scores have a specific meaning in terms of the skill scale and in terms of a particular topic. Student progress meaningfully can be followed over time within and across subjects. And because DiscoTests are formative assessments that are made to be embedded in curricula, students can be tested frequently, permitting teachers, parents, schools, and districts to keep a close eye on student learning.

Conclusions

Integrating new knowledge requires recognizing how new patterns encountered in the world map onto patterns we already recognize. Seeking patterns is the basic algorithm of the neocortex. We learn from matching patterns in our environment to relevant memories of patterns. Closing this gap requires multiple opportunities to experience and work with the complex nature of the ideas we want children to understand.

The standards are humanity's legacy to the next generation, a product of dedicated and tireless work in identifying the patterns in nature important to humankind. The desire to pass along these treasures to our children is an understandable aspiration, but no longer a realizable goal. For centuries it was possible to provide the next generation with the whole of human understanding, but as information currently doubles in the space of years instead of centuries, the task is not possible, practicable, or even useful. A dangerous trajectory is increasing the number of standards as the fund of knowledge increases. One can imagine a future with the absurd outcome of teachers needing to address and students needing to learn several standards a day.

[xliii] These can be viewed at http://www.discotest.org/visitors/_selectassess_visitor.php Choose the "Energy Teaser" to view a test, a rubric, and a report. The DiscoTests site is currently undergoing renovations so there may be missing links or undiagnosed bugs.

Educators and policy makers need to prevent the standards from becoming a wish list instead of a list from which educators can develop meaningful goals. We recommend setting limits on the number of standards for which teachers and students are responsible. This number can be negotiable, but it should provide a target that teachers and students can reasonably work toward.

We further recommend that policy should support teachers in integrating "*depth of focus"* in classroom pedagogy to allow students to identify, share and discuss the patterns they discover. This work provides the necessary foundation for later comparison to the patterns recognized and used by scientists.

To complete the feedback loop, students and teachers need assessments that complement available standardized tests. Current tests effectively measure what students know. New assessment tools are now available that measure the degree of complexity in student thinking. We introduced DiscoTests as an assessment tool that emerged after decades of research in cognitive science and testing in schools. This tool serves teachers and students by enriching the learning process and supporting teaching, and has the added advantage of allowing researchers to continually refine the learning sequences and in turn improve assessments and feedback.

And thus we complete the feedback loop of this paper. We set out to highlight and demonstrate the importance of goals, and the impact of current state and national standards on our collective success in helping students understand the content in the standards. They represent a goldmine in possible outcomes for students, but neither students nor teachers can focus on every nugget. Research supports the importance of focused work to complement the breadth of exposure to the rich legacy embedded in the standards. And finally, any meaningful growth in what and how students understand is the outcome of repeated assessment of their progress. We need to evaluate not only what students know, but also how students know, because the richest part of being human is the ability to build on our legacy. Ultimately this is the goal we have for all our children.

Bibliography