CBE—Life Sciences Education
Vol. 10, 329–333, Winter 2011

Feature
Approaches to Biology Teaching and Learning

Reconsidering "What Works"

Kimberly D. Tanner

Department of Biology and SEPAL: The Science Education Partnership Assessment Laboratory, San Francisco State University, San Francisco, CA 94132

It is a wonderfully exciting time to be involved in undergraduate biology education efforts. The publication of *Vision and Change: A Call to Action* has motivated many to redouble their efforts in engaging their colleagues and institutions in fundamentally changing the experiences of students in college and university biology classrooms around the country (American Association for the Advancement of Science [AAAS], 2011). In addition, there is a growing community of biologists embarking upon more analytical, scientific, and scholarly approaches to their teaching (Handelsman *et al*., 2004). Journals such as *CBE—Life Sciences Education* are building a body of evidence and understanding about the complex nature of biology teaching and learning. Most recently, the founding of SABER—the Society for the Advancement of Biology Education Research—is a further step toward fostering a thriving discipline-based biology education research community that encompasses the often-isolated subdisciplines of biology education.

Within this excitement and the new wave of biology education reform, though, there often appears to be a central focus on "solving the problems" of undergraduate biology education, and a desire to bring closure by determining the "best way to teach." Over the past several years, at conferences, on websites, and in scholarly conversations, I increasingly have heard common use of the phrase "what works" in referring to the primary goals of undergraduate biology education reform. However, trying to determine "what works" is problematic in many ways and belies the fundamental complexities of the teaching and learning process that have been acknowledged by scholars for thousands of years, from Socrates, to Piaget, to more recent authors and researchers. As our biology education community expands, it seems worthwhile to return to those lessons from the research literature that caution whether anything as simple as establishing "what works" is even possible in the messy landscape of teaching and learning. Language is powerful. Language can frame the issues in an emerging discipline, define the boundaries of a conversation, and imply the extent to which intellectual ideas are still open for discussion and examination.

Below are six perspectives on why the use of the phrase "what works" would seem to hinder, rather than foster, an evidence-based approach to biology education reform and scientific teaching. By articulating these challenges to the assumption that there exists anything as singular and clean as "what works," we can perhaps reconsider what we really mean when asserting "what works." In addition, we can perhaps refocus on what has been shown again and again to be the path toward effective teaching and learning: the development of reflective instructors who are analytical about their practice and who make iterative instructional decisions based on evidence from the students sitting right in front of them (e.g., Darling-Hammond, 1997a; Loucks-Horsley *et al*., 2003). Finally, a reconsideration of "what works" could encourage instructors, researchers, and policy makers to be careful with interpretations of evidence in emerging studies in biology education research, so that we continue to deepen our understanding of the nuances and complexities inherent in teaching and learning, as opposed to declaring our efforts done before we have even really started.

DOI: 10.1187/cbe.11-09-0085
Address correspondence to: Kimberly D. Tanner (kdtanner@sfsu.edu).

"WHAT WORKS" IS INCONGRUENT WITH THE NATURE OF SCIENCE

Most simply, the phrase "what works" seems incongruent with the nature of any evidence-based, scientific enterprise. The phrase implies the speaker has solutions that will be immediately applicable in all contexts. Additionally, the phrase hosts a somewhat sinister implication that further investigation into the issue at hand—in this case effective biology teaching and learning—is likely unnecessary and unwarranted. As scientists by training, I would suppose that most involved in biology education reform would find this stance and this language perplexing if they encountered it in their basic, biological research context. "What works" is not a common phrase in the vernacular of scientists, who work in a land where details are everything. The experimental methods that work in one model system often require tweaking or complete rethinking to be used in a different organism. In fact, we are entering the age of personalized medicine, a practice that differs for each individual based on molecular diagnostics and bioinformatics. At one point in the history of biology, it was postulated that we would discover a singular cure for cancer. An important insight into cancer treatment in the last decade, however, is the realization that each cancer is unique, and that knowledge of the molecular origins of a tumor are needed to predict the most likely effective treatment for an individual's cancer. It is also fundamental to the nature of science that all explanatory paradigms are subject to skeptical review, and that inquiry into a question is never "done" or "complete" (Lederman, 2007). This is, in fact, one of the most common misconceptions about science that we are collectively trying to alter among our students at all educational levels, kindergarten through college and beyond (AAAS, 1989, 2011; National Research Council, 1996; Lederman, 2007). It is, in fact, from reexamination of a well-accepted idea in another context that an explanatory story in the sciences can begin to fall apart. The concept of viruses and bacteria as the primary causative agents of human disease impeded the consideration of novel mechanisms, namely misfolded proteins, that might be at work as agents of human disease (Prusiner, 1991). As a second example, a publication appeared just last year in one of the most prestigious science journals in our country questioning the fundamental portrayal of the atomic building blocks of life that appears regularly in our textbooks (Wolfe-Simon *et al*., 2011). It is fundamental to the nature of science that issues as basic as these are still under discussion. In fact, philosophers of science have pointed out that this aspect of the nature of science—that everything can always be questioned and reexamined—is essential for scientific revolutions and paradigm shifts in human scientific thought (Kuhn, 1962). Since it would be rare, if not unheard of, to express a sentiment similar to "what works" in our scientific investigations of the natural world, why should our evidence-based investigations and views about the issues in the teaching and learning of biology be any different?

"WHAT WORKS" IGNORES INDIVIDUAL STUDENTS AND THEIR BRAINS AS KEY VARIABLES

> In education, we used to study large groups of individuals. That's like studying a mountain chain with averages. You miss all the peaks and the valleys. Now, we study individuals—Kathleen Fisher, as quoted in Schnepps (1997)

The phrase "what works" often appears in reference to the efficacy of a particular teaching strategy for students. Many current studies of biology teaching and learning appear to operate on the unfounded assumption of uniformity among students, measuring the effectiveness of teaching outcomes with minimal or no attention to the individuality of learners or the particular demographics of the students involved in the study. This is akin to assuming the uniformity of a strain of an organism in experimental procedures, and many of us in biology education were trained using those very assumptions in our basic biological research efforts. However, this assumption appears to be poorly justified, given the heterogeneous populations in our classrooms, which needless to say are occupied by humans that bear no resemblance to an inbred line of laboratory-based model organisms. Research in a variety of social science fields is revealing that "what works" for students with different demographic characteristics—including gender, language background, levels of family education, ethnic identity, and a myriad of other characteristics—is simply not the same (e.g., Delpit, 1985; Steele and Aronson, 1995; Brown, 2004; Johnson, 2007). These research findings *do not* suggest that there will exist certain pedagogies that are optimal for students with certain demographic characteristics. However, they do suggest that assessments of the effectiveness of a teaching strategy must take into consideration the nature of the student population under study. Increased attention to the effectiveness of different teaching approaches with different populations of students at different institutions appears to be an emerging and driving question in biology education research, one that is entirely warranted by the existing research literature.

In this context, it may prove useful to remember that learning is fundamentally a biological process occurring in the individualized brains of our students and to consider the biological origins of that individuality and variability. Even with comparable demographic characteristics, students come to classrooms with variable levels of understanding and experience in the areas of biology we are attempting to teach. While there are common mechanisms for learning, memory, and brain plasticity among humans, the substrate upon which those mechanisms act in individuals sitting side by side in a classroom is likely far from similar. The synaptic architecture of all of the connections between neurons in an individual's brain is a perpetually changing state that results from a combination of early development; nonstop, experience-dependent change; and the unique set of genetically determined protein receptors mediating attention, motivation, and memory in that individual (for an overview, see Eshel [2007]). If we start with all of these individual brains—individual in terms of both architecture and information previously stored within—why would we logically expect any singular approach to teaching and learning to be effective in all cases? Indeed, what if there really is no one approach to undergraduate biology teaching that "works" for all students in all brain states? What if what really "works" is to approach teaching from a diversified pedagogical perspective and to teach in such a way as to maximize access to biological learning for the greatest number of individuals? What if the right way to teach is *not any singular way*, but rather the use of a variety of teaching techniques intertwined to benefit a range of learners and their experiences in a heterogeneous classroom? What if the closest we can get to "what works" is to teach using *all* of the available techniques and not just one? While this may not be a surprising conclusion, it is worth reminding ourselves that diversity in pedagogy is likely to be a good strategy, given the diversity of our students.

"WHAT WORKS" ASSUMES UNIFORMITY IN INSTRUCTOR EXPERIENCE AND SKILL

> Bureaucratic solutions to problems of practice will always fail because effective teaching is not routine, students are not passive, and questions of practice are not simple, predictable, or standardized. Consequently, instructional decisions cannot be formulated on high then packaged and handed down to teachers."—Linda Darling-Hammond (1997a)

Inherent in the phrase "what works" is the notion that it should "work" in the hands of all instructors, if not equally well, then at least to some degree. In the same way that assumptions of student homogeneity are unfounded, assumptions of instructor equivalency are also problematic. Many assertions about the efficacy of a given instructional technique appear to be lovely and interesting examples of a successful teaching experience by a single instructor, often after a great deal of thought and years of practice. But perhaps the teaching success was the result of something other than the pedagogical technique under study. In fact, for some studies that examine serial approaches from year to year in an instructor's classroom, it would seem a plausible conclusion that the key variable in teaching success was a general increase in the instructor's skillfulness over time, as opposed to the particular technique that the author is assigning as the causative reason. Indeed, the often far-reaching nature of the "what works" conclusions and the lack of attention to individual instructors—and their professional development over time—is cause for concern.

It is well documented in the K–12 education research literature that student achievement is strongly linked to teacher quality (Darling-Hammond, 1997a,b). In addition, it is widely acknowledged that the development of highly skilled K–12 teachers takes years, if not decades, and that such development happens most often by ongoing improvement through participation in a professional learning community of other educators (Lave and Wenger, 1991; Darling-Hammond, 1997a). Why would we expect the development of a highly effective biology instructor at the college and university level to be different? In fact, the challenge of developing instructor skill in undergraduate science education is even more difficult. Few undergraduate biology instructors have any formal training in how to teach, in stark contrast to the extensive credentialing system required to enter K–12 science teaching. Those faculty that do have some training most often experienced limited teaching assistant workshops while graduate students; at best they have 5–12 d of faculty development workshops over the course of 1–3 yr (Pfund *et al*., 2009; Ebert-May *et al*., 2011). It is rare to encounter undergraduate science instructors who have the opportunity to regularly engage in professional development and conversations with colleagues about effective science teaching, that is, who have access to a professional learning community (Lave and Wenger, 1991). The scientific community would balk at the idea of relying on such limited training to develop skilled research scientists, so why would this appear sufficient to develop skilled and effective undergraduate biology instructors?

Yet many reports of "what works," as well as the conclusions of many biology education research publications, are silent on the role of instructor training and experience. Variability in this parameter will certainly impact investigations of teaching techniques. Lack of consideration of instructor variability and skill development can have the detrimental effect of encouraging instructors to try complex pedagogies that may take multiple years to master. But they have been led to believe that they need only put these strategies in place in their classroom and effective teaching and learning will happen in response. This disconnect is the likely source of a statement I encounter often as a source of faculty resistance to new teaching approaches: "Well I tried that teaching technique—[insert any teaching technique of your choice from a simple pair–share to highly structured problem-based learning]—and it was a disaster." As opposed to questioning his or her own skill, the speaker immediately rejects a new teaching technique as ineffective, assuming that the key variable is the teaching technique itself, as opposed to instructor skill. If we collectively acknowledge instructor skill, professional development, and training as essential components of biology education reform, then perhaps the response would be, "Well I tried that teaching technique, but I think that I need to change a few things about the way I was using it and try it again."

"WHAT WORKS" REQUIRES DEFINING WHAT IS MEANT BY "WORKS"

Assumptions of Alignment between Motivation and Conceptual Learning

As our efforts in biology education research and reform expand, we will no doubt develop a myriad of measures with which we evaluate effective teaching and learning. Currently, however, what is meant by "works" is often implied and left to the reader or listener to define and infer from the measures being presented. However, "what works" toward one goal in the context of undergraduate biology education may or may not be similarly effective toward a different goal.

There is overwhelming evidence that some students lose interest and motivation to study science to the extent that they simply leave the field (Tobias, 1990; Seymour and Hewitt, 1997). However, few conversations of "what works" appear to address the motivational and identity issues that are the core reasons why students leave the undergraduate sciences. The majority of students who leave our disciplines do so not because they are incapable of the conceptual learning, but because they are not adequately engaged (Tobias, 1990; Seymour and Hewitt, 1997). Yet, there appears to be a general assumption that all biology education reform efforts should "work" in the same direction on all of the challenges that we face. But what "works" in terms of increasing student test performance or grades may, for example, have inverse effects on student interest in the subject matter and career goals. In contrast, there may be pedagogical methods that are effective in increasing student engagement and motivation, but that do not translate to conceptual change or long-term retention of biological ideas. Examining more closely what is meant by "works" and what measures we would use to detect changes in things, such as attitude, motivation, and conceptual learning, will no doubt lead us to a more nuanced view of the complexities of what is going on in our classrooms. Questions about the multiple goals for student growth that may constitute "works" are ripe for codification, inquiry, and simultaneous measurement in biology education research efforts.

Assumptions of Alignment between Short-Term and Long-Term Impact

I have no doubt that undergraduate biology educators collectively aspire to teach students such that what they gain in our classrooms will continue to serve them 1, 2, 5, 10, and even 20 yr later. I have yet to hear a colleague express satisfaction with the idea that their teaching impact would last only until the student leaves their classroom at the end of a term. However, evidence for "what works" is often based on short-term and relatively ill-defined measures. Grades, in particular, as an assessment of "what works" would seem simultaneously both limited to short-term impact and vague in what is actually being measured, beyond success in the playing the game of school. Similarly, short-term assessments or attitudinal surveys are useful immediate measures, yet say little about "what works" in terms of what most value: long-term impact and retention of conceptual learning.. "What works" for short-term performance in a course, measured by immediate assessments, may or may not be the same as "what works" for deep conceptual change and long-term retention, yet we have little to no evidence beyond a single semester time frame.

The complex intersection between goals for learning and time can be clearly seen in an example from K–12 educational research (Leonhardt, 2010; Chetty *et al.*, 2011). In their study of 12,000 Tennessee kindergartners for more than two decades, economic researchers demonstrated the immediate positive impact of an experienced kindergarten teacher on student achievement as measured by test scores. As had been previously demonstrated, this impact "faded-out" and was not detectable in analyzing test scores for these same students in middle school. With only this evidence, some would argue against the critical importance of kindergarten teacher quality. However, the researchers went further and analyzed data on these same students as adults, including analysis of college attendance and job earnings. Those students who had been randomly assigned to kindergarten teachers with the most experience and who had learned more from these teachers than peers in other classrooms showed stark differences on a variety of life measures. Most importantly, they were more likely to go to college and were earning more as adults than peers taught by other less-experienced kindergarten teachers (Leonhardt, 2010; Chetty *et al.*, 2011). This study emphasizes that evaluation of teaching effectiveness may produce different conclusions using different measures at different points in time. In addition, it returns our focus to the ever-present question of what really are our goals in reforming undergraduate biology education, and what measures can we use reliably to assess effectiveness of teachers and teaching in reaching those particular goals?

Most of the assessments that are currently being used to gauge "what works" are short-term measures using data gathered by biology educators in their own classrooms. While these are accessible and convenient data to collect (and no doubt of high value to those particular instructors in shaping their teaching practice), self-studies are inherently limited in their relevance. In addition to the instructor bias (knowing the research questions and desired outcomes of the study), these self-studies are the extreme example of deriving "what works" in a very particular context, in the hands of one instructor, among one particular group of students. One goal of the emerging field of discipline-based biology education research is to foster research studies that move beyond individual efforts to broader studies involving multiple faculty at multiple institutions and research in both classroom contexts and in more controlled, nonclassroom settings.

BUILDING A COMMON LANGUAGE ABOUT THE SUBSTANCE OF THE "WHAT" IN "WHAT WORKS" IS NOT TRIVIAL

> "Fidelity of Implementation" refers to the "extent to which the user's current practice matched the developer's 'ideal.'" —Carol O'Donnell (2008)

In our current efforts to promote innovation in undergraduate biology education, we are still in the midst of building a common lexicon with which to describe the teaching strategies being used. I have frequently listened to conference presentations that describe the use of some teaching strategy, only to realize in conversations with the presenter afterward that what I had envisioned was happening in their classroom had little to no resemblance to what had actually occurred.

There have been extensive studies in K–12 education about the disconnects between the ideal of a curriculum or teaching technique and what is actually implemented in a classroom. How to deal with these disconnections between actual and intended teacher practice is a well-known challenge in the K–12 science education literature, and is commonly referred to as the problem of "fidelity of implementation" (O'Donnell, 2008). The range of teaching practices that may actually be happening, even when teachers report using the same curriculum and the same teaching strategies, have long been suspected to be a source of variation in student outcomes when using the same curricular approach. Fidelity of implementation is often a key assumption underlying assertions of "what works." An example of its importance can be seen in a recent investigation of the impact of faculty development programs on university biology faculties' actual teaching practice (Ebert-May *et al.*, 2011). The study asked a critical, yet simple, question: To what extent did faculty self-reports about teaching differ from data collected by independent observers in their classrooms? Analyzing videotapes of classrooms of college and university biology faculty who had participated in one of two faculty professional development programs, the researchers reported that while 89% of the faculty reported that they had moved toward "active, learner-centered instruction," 75% were observed to use "lecture-based, teacher-centered pedagogy" (Ebert-May *et al.*, 2011). While there are many issues at play in this study, it is presented here to highlight the importance of increasing the frequency and detail of our collective conversations about what we are actually doing in our classrooms. What "think–pair–share" or "clicker questions" or "problem-based learning" means to one instructor is sometimes a far cry from what it means to another. Increased collective understanding of and deep discussion about "what" is happening in classrooms would appear to be a first step toward increasing fidelity of implementation of innovative teaching techniques. In addition, such discussion would promote formation of professional development communities for biology faculty. Miscommunications about "what works," and differences in vision of what these teaching techniques look like when implemented in a classroom, point to the need for more dialog among biology faculty at large, as well as among biology education researchers.

IN CONCLUSION: SO WHAT IF WE NEVER FIND "WHAT WORKS"—A GOAL OF SCIENTIFIC TEACHING FOR ALL FACULTY

The use of the phrase "what works" appears to reflect a common desire to have principles of biology teaching that will generalize to all students, all instructors, and in all settings and courses. However, this desire for generalization may not be realistic. And, so what! The key insight in K–12 science education has been the absolutely critical role of a high-quality, experienced teacher engaged in reflective practice and dialogue (Darling-Hammond, 1997a). Perhaps the end goal, then, really is not to define the "best way to teach" and "what works," but rather to develop all biology instructors as reflective instructors who are analytical about their practice and who make iterative instructional decisions based on evidence from the students sitting right in front of them. If this were our goal, then "what works" becomes irrelevant. It would only matter whether or not instructors were monitoring how their current teaching approaches were or were not helping move their current students toward their goals for them. At some level "what works" arises from a desire to give scientists a shortcut to effective teaching, but there may not be any shortcuts.

Perhaps what is most frustrating about the language of "what works" is its implication that we are done—or that we ever will be—in our efforts to understand the wonderfully complex situation of teaching and learning about the biological world. That would take away all the wonder, excitement, and discovery that brought many of us to the scientific disciplines—and then to the desire to teach them well—in the first place.

REFERENCES

American Association for the Advancement of Science (AAAS) (1989). Science for All Americans, New York: Oxford University Press. www.project2061.org/tools/sfaaol/sfaatoc.htm (accessed 14 October 2011).

AAAS (2011).Vision and Change in Undergraduate Biology Education: A Call to Action, Washington, DC: American Association for the Advancement of Science.

Brown B (2004). Discursive identity: assimilation into the culture of science and its implications for minority students. J Res Sci Teach *41*, 810–834.

Chetty R, Friedman JN, Hilger N, Saez E, Schanzenbach DW, Yagan D (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. Q J Econ. (*in press*). http://obs.rc.fas.harvard.edu/chetty/chettypaperschrono.html (accessed September 4 2011).

Darling-Hammond L (1997a). Doing What Matters Most: Investing in Quality Teaching, Kutztown, PA: National Commission on Teaching and America's Future.

Darling-Hammond L (1997b). The Right to Learn: A Blueprint for Creating Schools That Work, Hoboken, NJ: Jossey-Bass.

Delpit L (1985). Other People's Children: Cultural Conflict in the Classroom, New York: New Press.

Ebert-May D, Derting TL, Hodder J, Momsen JL, Long TM, Jardeleza SE (2011). What we say is not what we do: effective evaluation of faculty professional development programs. BioScience *61*, 550–558.

Eshel N (2007). The Science Inside Learning, Washington DC: American Association for the Advancement of Science.

Handelsman J, *et al.* (2004). Scientific teaching. Science *304*, 521–522.

Johnson A (2007). Unintended consequences: how science professors discourage women of color. Sci Educ *91*, 805–821.

Kuhn TS (1962). The Structure of Scientific Revolutions, Chicago, IL: University of Chicago Press.

Lave J, Wenger E (1991). Situated Learning: Legitimate Peripheral Participation, New York: Cambridge University Press.

Lederman NG (2007). Nature of science: past, present, and future. In: Handbook of Research on Science Education, ed. SK Abell and NG Lederman, Mahwah, NJ: Lawrence Erlbaum, 831–879.

Leonhardt D (2010). The case for the $320,000 kindergarten teacher. New York Times. July 27, page A1. www.nytimes.com/2010/07/28/business/economy/28leonhardt.html (accessed 14 October 2011).

Loucks-Horsley S, Love N, Stiles KE, Mundry S, Hewson PW (2003). Designing Professional Development for Teachers of Math and Science, Thousand Oaks, CA: Corwin Press.

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

O'Donnell C (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. Rev Educ Res *78*, 33–84.

Pfund C, *et al.* (2009). Summer institute to improve university science teaching. Science *324*, 470–471.

Prusiner S (1991). Molecular biology of prion diseases. Science *252*, 1515–1522.

Schnepps M (producer) (1997). "Lessons from Thin Air," *Minds of Our Own*, segment 2, Cambridge, MA: Science Media Group, Harvard/Smithsonian Center for Astrophysics. DVD.

Seymour E, Hewitt NM (1997). Talking About Leaving: Why Undergraduates Leave the Sciences, Boulder, CO: Westview Press.

Steele CM, Aronson J (1995). Stereotype threat and the intellectual test performance of African Americans. J Pers Soc Psychol *69*, 797–811.

Tobias S (1990). They're not dumb. They're different. A new tier of talent for science. Change *22*, 11–30.

Wolfe-Simon F, *et al.* (2011). A bacterium that can grow by using arsenic instead of phosphorus. Science *332*, 1163–1167.