Stop Punishing Teachers:
What We Should Be Doing Instead to Improve Education

Theodore W. Frick

Department of Instructional Systems Technology
School of Education, Indiana University Bloomington

Last revised: February 21, 2016

Introduction

Imagine that a strategy to improve the quality of medical practice was to punish doctors whose patients are more likely to die, by paying them less than other doctors or possibly firing them. Doctors who specialize in oncology, cardiology, and geriatric care would be punished because of higher patient mortality rates. Who would want to become an oncologist, a cardiologist, or a geriatrician? Under such an unfair evaluation system, these doctors would be considered ineffective, because more of their patients die from cancer, heart attacks, and old age. Even worse, this strategy would further result in a shortage of medical doctors with these specialties.

Yet, a similar strategy is being considered for improving the quality of education by effectively punishing teachers whose students achieve less than those of other teachers in their district. In the long run, such disincentives will discourage people from entering the teaching profession. Such a teacher accountability strategy is clearly irrational. It will do more harm than good. It is ill conceived and should be abandoned immediately.

Do Teachers Add Value?

In today's political climate there is considerable talk about making teachers more accountable as a means to improve the quality of our schools. Through an empirical statistical procedure called *value-added modeling*[i], teachers whose students achieve more would be rewarded with salary increases and keep their jobs, while those whose students achieve less would be receive no salary increases or possibly be fired. But the fact is that poor student performance on standardized tests is affected by factors that teachers cannot control—including poverty, living in high-crime neighborhoods, lack of early childhood education, limited learning opportunities outside of school, and lower parent education levels. In fact, "teacher differences account for about 10 % of variance in student test score gains in a single year" (Haertel, 2013, p. 5). Other factors beyond teacher control account for about 60 % of the variance in student learning achievement. Another 20 % to 30 % is attributed to random variation (unobserved in- and out-of-school factors). While teachers *do* make a difference, other factors make a much greater difference in student learning achievement—factors that teachers have no control over (Haertel, 2013).

Who would want to become a teacher under such accountability methods? Teacher salaries are already relatively low, when compared with other professions[ii], and now teachers will be further punished if their students do not perform well on standardized achievement tests.

While it is true that K-12 school systems are now instead using *multiple* indicators for annual teacher evaluations[iii], the use of standardized student achievement test scores for any significant portion of teacher evaluation is patently unfair to teachers. On the other hand, such tests can be used as indicators of student achievement, and potentially also for diagnostic purposes. That is what these tests are designed to measure, not *teacher* effectiveness.

A Different Metaphor

To use a different metaphor, this value-added modeling approach would be similar to a strategy where chefs in a restaurant are given no raises or fired because they do not *add value* in terms of profitability. To continue the metaphor, imagine that you sit down for a meal in a restaurant. A more holistic evaluation of the restaurant could consider 4 dimensions of quality:

1. *Content* of food to be served (e.g., menu items; nutritional value of ingredients; dietary balance; quality of ingredients),

2. *Context* of the restaurant in which meals will be eaten (e.g., comfort of seating; cleanliness of facilities; dining atmosphere; adequacy of food preparation area, food storage, and waste disposal),

3. *Process* of experiencing the dining event (preparation of meals by chefs; timeliness of courses served; customer consumption of food and drinks; service provided by waiters), and

4. *Outcomes* (e.g., satisfied dining customers with no subsequent food-born illness; tips for waiters and payment for meals by customers; sustainability—sufficient numbers of returning and new dining customers to stay in business; retention of competent restaurant staff).

All of these dimensions of quality are important. For example, it would be shortsighted to focus solely on the quality of chefs without consideration of the choices on the restaurant's menu and the quality of its dining environment. There are many reasons that customers might not come to the restaurant: they may not like the choices available on the menu due to high caloric content or to omission of vegetarian or seafood entrées; they might not like the dining atmosphere in the restaurant because it is dirty, noisy, smoky, too warm or cold, and uncomfortable seats; patrons could be dissatisfied with the service by waiters; and they could find that the meals are too expensive for their dining budget. Yet chefs could be doing a good job preparing meals.

In short, we could have excellent chefs in this particular restaurant who can prepare excellent meals, but customers could be dissatisfied with other aspects of the restaurant. So would it make sense to punish the chefs if the restaurant is losing money? This is not rational as a strategy for improvement.

Now, instead of dining in a restaurant, let us turn to education.

Would it make sense to create disincentives for teachers because their students are not meeting academic standards? Yet, this is the conversation going on in many state

departments of education in the U.S. who are concerned about improving educational outcomes (Cody, 2012).

To be consistent under such a line of reasoning, if teachers' students do not meet state standards, should we not also shut down the state department of education who licensed those teachers and the local school boards who hired them? Surely, a state department of education whose schools are failing to advance student achievement should be a sufficient reason to punish or eliminate the state department. By using the very same logic for getting rid of poor teachers, state departments and local school boards should all resign if they are not adding value either.

Then the U.S. federal government could take over local school districts. Since the NCLB law guaranteed that our K-12 schools would fail to get 100 % of their students to meet academic standards by 2014[iv], the federal government should also relinquish control of the schools since it too does not add value. Perhaps we should then look to the governments of Singapore or Finland whose students appear to be faring much better on standardized tests than do those in the U.S. They must be doing something right.[v]

The absurdity of NCLB expectations is further mirrored by this analogy that Diane Ravitch (2013) provides:

> It was as though Congress had passed a law saying that every city in America should be crime-free. Who could disapprove of such a laudable goal? What city would not want to be crime-free? But imagine if the law set a deadline twelve years off and said that any city that did not meet the goal would be punished; its police stations would be closed and privatized; its police officers would lose their badges. The first to close would be the police stations in the poorest neighborhoods, where crime rates were highest. Eventually, the scythe would swing even in affluent neighborhoods, because no city is completely crime-free. Wishing that it might be so, or passing laws to require that it be so, does not make it so. (Locations 311-316)

Perhaps these analogies in the fields of medicine, restaurant management, and in criminal justice will have exposed the shortsightedness of current discussions about improving our schools. Let us now turn to a more holistic approach.

Four Dimensions of Educational Quality

The essence of education is intended, guided learning (Steiner, 1988). When students intend to learn and teachers attempt to guide student learning, then education is occurring.

The following dimensions of education can be assessed with respect to quality:

1. *Content*: goals of learning, design of teaching-learning activities, and resources to support those activities;

2. *Context*: environment for teaching and learning;

3. *Process*: what teachers and students do with the content in that context;

4. *Outcomes*: results of what students learn, sustainability, unanticipated side effects.

Some Examples

When evaluating the first dimension of educational quality, we can evaluate the worthwhileness of *goals* of education. One of the most basic questions ought to be: What should be the purpose of education?[vi] Very little current discussion appears to be on the goals, beyond state adoption of Common Core Standards[vii]. What if we are expecting students to achieve the wrong goals?

Similarly, we can consider the *learning tasks* that students are expected to do, how they are sequenced, and task authenticity (van Merriënboer & Kirschner, 2007). In other words, what is likely to motivate students and help them to achieve desired goals?

In the first dimension we can further evaluate curriculum resources that are made available to students and teachers. For example, the quality of textbooks used in schools should be evaluated. If the textbooks are poor, this would be analogous to a chef in a restaurant who has poor ingredients to work with when preparing meals, and a narrow range of ingredients to use, including lack of ingredients that are associated with a balanced healthy diet.

When considering the second dimension, we can evaluate the environment in which students and their teachers are working. For example, if students and their teachers feel unsafe, and the school's roof is leaking and windows are broken, these contextual factors would not support teaching and learning, and instead serve as obstacles.

When considering the third dimension, we could evaluate the presence or absence of principles of instruction known to promote learning. We could evaluate student engagement in learning tasks and their success in doing them (Merrill, 2012; Frick, Chadha, Watson, & Zlatkovska, 2010).

When considering the fourth dimension of educational quality, we can evaluate the integration of cognitive, conative and affective learning outcomes as kinds of student achievement. Cognitive outcomes pertain to what students think and know, including generalizable concepts, relations and criteria; practical know-how (follow protocols, adapt, and create); and qualitative knowing of uniques (recognition, acquaintance and appreciation). Conative outcomes refer to student intentions—not only in the short-term, but for life—e.g., wanting to learn, to seek truth, to excel; to become a nurse, a social worker, a teacher, a physician; to own and run a business; to be a leader; to do what is right; to be rational. Affective outcomes pertain to development of sensitivity—e.g., attentiveness, immediate awareness, compassion, kindness, caring for others.

Readers should note that the Common Core Standards largely address *cognitive* outcomes in mathematics and English language arts in grades K-12. Conative and affective goals appear to be missing from the Common Core Standards, nor are these important goals assessed by standardized tests. This is particularly salient in light of findings about prevalent student feelings about school. For example, the majority of U.S. high school students are bored every day in school. Yazzie-Mintz (2007) summarizes results from a survey of 81,499 students in 110 high schools across 26 U.S. states. Approximately 2 out of 3 students said that they were bored in class every day. When asked why they were bored, the top reasons were that learning materials were uninteresting, irrelevant and not

challenging enough. Yazzie-Mintz cited one student who stated, "Our school needs to be more challenging. Students fall asleep because the classes aren't really that interesting." Another said, "School is easy. But too boring. Harder work or more is not the answer though. More interesting work would be nice" (p. 10). Students who considered dropping out of school indicated that the main reasons are dislike of their school and teachers. Sixty percent further said, "I didn't see the value in the work I am asked to do" (Yazzie-Mintz, 2007, p. 5). For those who stay in school, the primary reason they do so is to get their high school diploma, so that they can go on to college.

The lack of integration of cognitive, conative and affective outcomes does not bode well in terms of student learning. Greenspan and Benderly (1997) have noted that since the ancient Greek philosophers, the rational or cognitive aspect of mind has often been viewed as developing separately from emotion. They argue that this view has blinded us to the role of emotion in how we organize what we have learned: "In fact, emotions, not cognitive stimulation, serve as the mind's primary architect" (p. 1). They identify the importance of emotion during human experience: "… each sensation … also gives rise to an affect or emotion…. It is this dual coding of *experience* that is the key to understanding how emotions organize intellectual capacities …" (p. 18, italics added). Greenspan and Shanker (2004) provide further evidence of how emotion is central to how we organize our thinking.

There is a biological basis for formation of mental structures (i.e., learning) as they are encoded through neural connections in the nervous system (Kandel, 1989; 2001; Squire & Kandel, 1999). Kandel (1989), a Nobel-prize winning neuroscientist, concludes from empirical evidence that:

> Learning produces changes in neuronal architecture (p. 103)…. Whereas short-term memory does not require the synthesis of new proteins … the consolidation of long-term memory … does require new protein synthesis (p. 109). … [T]he long-term process differs from the short-term process in two important ways: one, the long-term process requires translation and transcription, and two, the long-term process is associated with growth in synaptic connections. (p. 115) …. Our evidence suggests that learning produces enduring changes in the structure and function of synapses... (p. 121)

Kandel recommends further study on the "… the *power of experience* in modifying brain function by altering synaptic strength…" (p. 123, italics added).

If emotion is indeed the architect of mental structures, as mounting evidence appears to support (Greenspan & Shanker, 2004), then it follows that many students are likely to be developing ill-formed mental schema for the subject matter they are expected to learn in school—mental structures which are weakened or disconnected from existing mental structures due to feelings of meaninglessness, irrelevance, boredom and even disdain with respect to the content of their education (Frick, 2015). Ideally, students should instead be developing mental structures that are strengthened through *authentic life experience* and *positive emotion*. If so, then those positive feelings and the authenticity of purposeful learning activities will facilitate organization of mental structures that constitute long-term memory.

Students could attain the Common Core Standards while remaining unenthusiastic towards learning itself, and fail to be inspired and to persevere in discovering lifetime pursuits. That is, students could perform well on standardized achievement tests, but not be able to answer the important question: What should I do with my life? This is especially problematic, according to the Theory of Totally Integrated Education (TIE), which predicts that such disconnected learning achievement is highly vulnerable to forgetting under these conditions (Frick, 2015, Figure 6b).

In summary, my major contention is that there are many important elements that are part of educational quality, not just the quality of those whom we call teachers. For further reading about what constitutes worthwhile education, see Steiner (1981) and Frick (2015).

How Do We Justify Criteria for Determining Educational Quality?

If we are going to determine quality, we must have *justifiable criteria* for making such judgments. Reasoned argument is paramount for such justification. Rationality is required.

Reasoned argument for criteria should not be based on what *is*, but rather on what *ought to be*. Reasoned argument for justifying criteria should not rely solely on empirical evidence, for to do so would be to commit the *naturalistic fallacy*. For example, it does not make sense to argue that murder of human beings is worthwhile, based on the empirical fact that murders do occur.

The ultimate criteria for making such judgments must be based on initial principles that are justified by means other than empirical evidence. As an example, the Greek philosopher, Plato, put forth the fundamental principles of *truth, goodness,* and *beauty*.[viii]

Another well-known philosopher, Immanuel Kant, reasoned that *justice* should be determined by the *categorical imperative*: "Act as though the maxim of your action were to become, through your will, a universal law of nature" (1785, p. 24). In other words, it is right for one person to do this action, only if it also should become a universal law for everyone to do so. For example, one should treat others with respect, because everyone ought to do so. On the other hand, murder of human beings cannot be justified, when judged rationally by the *categorical imperative*.

The educational philosopher, Elizabeth Steiner, further argued for these criteria: "The justification of the principles of *universality* (impartiality), *autonomy* (liberty), and *humanity* (rational benevolence) resides in the intuition of *rationality* as the essential characteristic of humanness" (2009, Section 13.5, italics added). Simply put, to be truly free, we must become rational. Therefore the primary goal of education should be to guide students to become rational (Steiner, 1981).

Who should be determining these criteria? We need educologists who contribute to sound philosophical educology, that is, "quantitative educology that consists of universal signs of criterial knowing that about worthwhile education" (Frick, 2015, http://educology.indiana.edu/philosophicalQuantitativeEducology.html). Educology is "knowledge of education" (Steiner, 1988). In addition to scientific and praxiological educology, we need philosophical educology.

In summary, justification of criteria for determining educational quality must be through reasoned argument from initial principles—i.e., through rationality—not from empirical fact. Philosophical educology is needed for such justification.

The Practical Need for Educology

Educators who have been around several decades have seen widely touted changes come and go. Most changes that have occurred in U.S. K-12 schools appear to be trial and error. For example, there may be more use of computer tablets and Wi-Fi networks in schools, more standardized achievement testing, more teacher accountability for student learning achievement, less state funding for public schools, more tax dollars going to private charter schools, and increased regulation of schools by state and federal governments.

Have any of these changes significantly improved K-12 education? While apparently well-intentioned state legislators and state departments of education are mandating changes in K-12 education, there are no guarantees of improving matters. Worse, these changes may cause more harm than good. The stakes are very high. The consequences of mistakes can be devastating for our children and our future.

The following questions have not been adequately addressed:

- "Change what?"
- "Change how?" and
- "How do you know the change is likely to work?"

We must know what to change in order to know how. Without knowing what to change, the "how" is irrelevant (Duffy, 2009; Frick, Thompson & Koh, 2006). We must know whether the change is likely to accomplish the goal and that the change will not have negative, unintended effects.

For example, attempts to hold teachers accountable for student achievement not under their control may drive the best teachers to leave the profession due to frustration with such working conditions. It may also discourage potentially good teachers from entering the profession. Moreover, the best students might leave the public schools to attend private schools, if their parents can afford it. This would leave public schools in possibly worse straits, with the least capable teachers and lowest achieving students remaining, and less money from public tax dollars to support them. Then what?

Some scholars argue that an entire paradigm change is needed in education. For example, Reigeluth & Karnopp (2013) have promoted a vision and strategies to help schools transform. These include significant curriculum expansion, individualized learner-centered instruction, and attainment-based evaluation of learning—that contrasts with existing time- and age-based structures for moving student groups through lock-step grade levels.

But do we know how well such a new paradigm will work? This does not mean that a new education system that is learner-centered and attainment-based is not worthwhile. Nor does it mean that changes to expand and revamp curriculum in school are not needed. It

just means that we lack knowledge to predict outcomes of new designs of education systems.

As an analogy, consider an old bridge that is failing—it is structurally weak and is impeding the flow of traffic. If the bridge is not fixed, it will collapse and vehicles will plunge into the river. When engineers design a new bridge, they utilize adequate scientific theories. No one in modern times would consider designing a new bridge by trial and error. Nor would they let politicians try to do it.

Yet, in education we are essentially proceeding by trial and error in attempts to improve education—whether tinkering around the edges or by creating new paradigms. We lack sound knowledge to make reasonable predictions whether or not the proposed remedies will fix the problems in education we face.

The Need for Precise Terminology in Educational Research

In disciplines where knowledge has significantly advanced, there has been careful development of terminology so that researchers know what each other is actually talking about. For example, in physics the concepts of *atoms* and *molecules* are clearly defined. Each *atom* has a particular combination of subatomic elements called *electrons*, *protons*, and zero or more *neutrons*. For example, a molecule of water is comprised of two hydrogen atoms and one oxygen atom. A hydrogen atom consists of one electron and one proton. A stable oxygen atom contains eight each of electrons, protons and neutrons.

As another example, not that long ago the field of medicine was not a discipline. There was no medical science, as there now is. At one time, physicians would prescribe *bloodletting* to treat all kinds of disease, which turned out to be an ineffective practice and has been largely abandoned ("Bloodletting," n.d.). Many people were harmed by such ignorance.

Medicine advanced, in part, because researchers in the field became more disciplined in their inquiry. Terms are now precisely defined in medicine. *Osteoarthritis* does not mean whatever people want it to mean (i.e., construct their own meanings). *Osteoarthritis* is the medical term for a particular disease. Researchers and practitioners in the field of medicine have agreed on what this term means. So when treatments of this particular disease are investigated, competent medical professionals know what they are talking about.

On the other hand, in the field of education, such precise terminology does not exist. Steiner has long-argued that such terminology is sorely needed for our field to advance (1986, 1988).

Basic terms of educology include *learning*, *knowing*, *signs*, *education system*, *teaching-studenting processes*, *teaching-studenting structures* and many others (Frick, 2015). See the basic glossary at http://educology.indiana.edu/glossary.html .

This vocabulary is proposed as a necessary step towards advancing educology as a discipline. Such advances in educology will be helpful in improving education, particularly, *worthwhile education for everyone*.

Summary

While having good teachers is clearly important for quality education, they are only part of what should be considered. Educational quality will not be improved unless we focus on multiple dimensions of education: content, context, process and outcomes. Criteria for judging these dimensions should be based not on what is, rather on what should be. *What should be* ought not be justified by empirical fact, but instead by reasoned argument. Criteria for judging educational quality must be consistent with initial principles that are justified rationally. Finally, to improve the quality of education, we need educology, that is, sound knowledge of education to inform decision-making and formation of strategies for change that are likely to result in worthwhile education for everyone: http://educology.indiana.edu/we2.html .

References

Bloodletting (n.d.). In *Wikipedia*. Retrieved February 1, 2015 from http://en.wikipedia.org/wiki/Bloodletting

Braun, H. (2005). *Using student progress to evaluate teachers: A primer on value-added models.* Princeton, N.J.: Educational Testing Service. Retrieved from April 22, 2015 from http://www.ets.org/Media/Research/pdf/PICVAM.pdf

Cody, A. (2012). How the war on teachers is changing the profession. Washington, DC: The Washington Post. Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/how-the-war-on-teachers-is-changing-the-profession/2012/01/27/gIQAbKyLWQ_blog.html

Duffy, F. (2009, Ed.). The Scarecrow Education Series on Leading School Improvement. Retrieved February 1, 2015 from: http://www.thefmduffygroup.com/publications/series.html

Frick, T. (2015). The theory of totally integrated education (TIE). Bloomington, IN: Monograph in Five Parts. Retrieved from http://educology.indiana.edu/Frick/TIEtheory.pdf

Frick, T. (2015). Educology: Knowledge of education. Retrieved from http://educology.indiana.edu/index.html

Frick, T., Chadha, R., Watson and Zlatkovska, E. (2010). Improving course evaluations to improve instruction and complex learning in higher education. *Educational Technology Research and Development*, *58*(2), 115-136.

Frick, T., Thompson, K., & Koh, J. (2006). Predicting education system outcomes: A scientific approach. In M. Simonson (Ed.), Proceedings of the Association for Educational Communication and Technology, Dallas, TX. Retrieved from http://www.aect.org/pdf/proccedings/2006I/06_17.pdf

Goldhaber, D., Brewer, D., & Anderson, D. (1999). A three-way error components analysis of educational productivity. *Education Economics*, *7*(3), 199–208.

Greenspan, S. & Benderly, B.(1997). *The growth of the mind and the endangered origins of intelligence.* Reading, MA: Addison-Wesley.

Greenspan, S. & Shanker, S. (2004). *The first idea: How symbols, language, and intelligence evolved from our primate ancestors to modern humans.* Cambridge, MA: Da Capo Press (Kindle edition).

Haertel, E. (2013). Reliability and validity of inferences about teachers based on student test scores. The 14th William H. Angoff Memorial Lecture, Education Testing Service Research and Development, Center for Research on Human Capital and Education, Princeton, NJ, March 21. Retrieved from https://www.ets.org/Media/Research/pdf/PICANG14.pdf

Kandel, E. (1989). Genes, nerve cells, and the remembrance of things past. *Journal of Neuropsychiatry, 1*(2), 103-125.

Kandel, E. (2001). The molecular biology of memory storage: A dialogue between genes and synapses. *Science, 294,* 1030-1038.

Kane, T., McCaffrey, D., Miller, T., & Staiger, D. (2013). Have we identified effective teachers? Validating measures of effective teaching using random assignment. Bill and Melinda Gates Foundation, *Measures of Effective Teaching* (MET) Project. Retrieved April 23, 2015 from http://www.metproject.org/downloads/MET_Validating_Using_Random_Assignment_Research_Paper.pdf

Kant, I. (1785). *Groundwork for the metaphysic of morals* (translated by J. Bennett, 2010). Retrieved from http://www.earlymoderntexts.com/assets/pdfs/kant1785.pdf

Merrill, M. (2012). *First principles of instruction: identifying and designing effective, efficient, and engaging instruction*. San Francisco, CA: Pfeiffer.

Mourshed, M., Chijioke, C. & Barber, M. (2010). How the world's most improved school systems keep getting better. London: McKinsey & Co. Retrieved from https://mckinseyonsociety.com/downloads/reports/Education/How-the-Worlds-Most-Improved-School-Systems-Keep-Getting-Better_Download-version_Final.pdf

National Education Association (n.d.). Myths and Facts about Teacher Pay. Retrieved April 22, 2015 from http://www.nea.org/home/12661.htm

Plato (~ 350 BC). *The republic of Plato* (Translated by F. Cornford, 1945). New York, NY: Oxford University Press.

Ravitch, D. (2013). *Reign of error: The hoax of the privatization movement and the danger to America's public schools*. New York, NY: Alfred A. Knopf.

Reigeluth, C. & Karnopp, J. (2013). *Reinventing schools: It's time to break the mold*. Lanham, MD: Roman and Littlefield Education.

Squire, L. & Kandel, E. (1999). *Memory: From mind to molecules.* New York, NY: Henry Holt and Co.

Steiner, E. (2009). Ethical theory (Lecture notes). Retrieved from https://www.indiana.edu/~tedfrick/steiner/Ethical%20Theory.pdf

Steiner, E. (1981). *Educology of the free.* New York, NY: Philosophical Library. Retrieved from https://www.indiana.edu/~tedfrick/steiner/Educology%20of%20the%20Free.pdf

Steiner, E. (1988). *Methodology of theory building.* Sydney, Australia: Educology Research Associates. Retrieved from https://www.indiana.edu/~tedfrick/steiner/Methodology%20of%20Theory%20Building.pdf

U. S. Department of Education (2015). Laws & Guidance / Elementary & Secondary Education / EASA Flexibility. Retrieved April 22, 2015 from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html

van Merriënboer, J. & Kirschner, P. (2007). *Ten steps to complex learning: A systematic approach to four-component instructional design*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Yazzie-Mintz, E. (2007). Voices of students on engagement: A report on the 2006 high school survey of student engagement. Retrieved from http://www.eric.ed.gov/PDFS/ED495758.pdf

End Notes

[i] Value-added modeling (VAM) is a complex statistical procedure with many technical problems, even when statistically controlling for factors beyond a teacher's control. For example, Henry Braun, a test and measurement expert at the Educational Testing Service, warns:

> VAM results should *not serve as the sole or principal basis for making consequential decisions about teachers*. There are many pitfalls to making causal attributions of teacher effectiveness on the basis of the kinds of data available from typical school districts. We still lack sufficient understanding of how seriously the different technical problems threaten the validity of such interpretations (2005, p. 15, italics added).

Edward Haertel, professor emeritus at Stanford University, and widely recognized as an expert in educational testing and assessment, issues an even more dire warning about the inherit unfairness of VAMs for teacher evaluation:

> Teacher VAM scores should emphatically *not* be included as a substantial factor with a fixed weight in consequential teacher personnel decisions. The information they provide is simply not good enough to use in that way. It is not just that the information is noisy. Much more serious is the fact that the scores may be systematically biased *for* some teachers and *against* others, and major potential sources of bias stem from the way our school system is organized. No statistical manipulation can assure fair comparisons of teachers working in very different schools, with very different students, under very different conditions. (2013, pp. 23-24.)

[ii] National Education Association, Myths and Facts about Educator Pay, http://www.nea.org/home/12661.htm .

[iii] When measuring teacher effectiveness by randomly assigning students to teachers and when controlling for prior student learning achievement, Kane, MacCaffrey, Miller and Staiger (2013) were able to predict between 1 % and 24 % of student learning gains in mathematics and English language arts in elementary grades, and between 3 % and 32 % of student learning gains in middle school grades (Table 1, p. 10). While such teacher effects were statistically significant (as measured by student surveys, multiple classroom observations and standardized tests), this still means that roughly two-thirds to three-quarters of the variance in student learning gains is unexplained.

[iv] Provisions of the *No Child Left Behind* law require that any school will be considered as failing where academic standards are not met by 100 percent of its students by 2014, and will be taken over. See http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf . Note that, as of early 2015, the U. S. Department of Education has since granted flexibility waivers to 43 of 45 states who have proposed to develop "rigorous and comprehensive plans … designed to improve educational outcomes for all students, close achievement gaps, increase equity, and improve the quality of instruction." (http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html).

[v] See Mourshed, M., Chijioke, C., & Barber, M. (2010), How the world's most improved school systems keep getting better.

[vi] See Steiner, E. (1981), *Educology of the free.*

[vii] According to their promotional website, the Common Core Standards have been adopted by 43 U.S. states, as of April, 2015: http://www.corestandards.org/standards-in-your-state/ .

[viii] Plato (~360 BC), *The Republic of Plato* (Translated by Francis M. Cornford, 1945).