

## Teacher Preparation: Evaluation and Consequences

Michael Feuer

The George Washington  
University

TeachingWorks working papers are unpublished manuscripts that focus on the professional training of teachers. They involve analysis of data or literature and reflect “good thinking” – clear, systematic interrogation of issues critical in the field of teacher training.

These working papers are circulated to promote discussion. As such they are freely available to a broad audience interested in the study and improvement of ideas and practices in teacher education.

TeachingWorks working papers are subject to a blind review process that focuses on the relevance of the proposed work to pressing problems in teacher education, the transparency and relevance of the methods to the questions asked, as well as the quality of the writing. All submissions should be original.

**The views expressed herein are those of the authors and do not necessarily reflect the views of the University of Michigan and/or TeachingWorks.**

## Teacher Preparation: Evaluation and Consequences

Michael Feuer

The George Washington University

**Michael Feuer** is Dean of the Graduate School of Education and Human Development and Professor of Education Policy at The George Washington University, and President of the National Academy of Education. Before coming to GW, for the previous 17 years Feuer held positions at the National Research Council of the National Academies, most recently as the executive director of the Division of Behavioral and Social Sciences and Education. Prior to joining the NRC he was a senior analyst and project director at the US Congress Office of Technology Assessment. Feuer received his BA (cum laude) in English literature from Queens College (CUNY), an MA in public management from the Wharton School, and a PhD in public policy analysis from the University of Pennsylvania. He has studied at the Hebrew University of Jerusalem and the Sorbonne, was on the faculty of the business school at Drexel University from 1981-1986, and has taught courses in education policy and research at Penn and Georgetown. His 2006 book, "Moderating the Debate: Rationality and the Promise of American Education" (Harvard), was praised for its applications of economics and psychology to education policy and research. Feuer consults regularly to educational institutions and government in the US, Israel, Europe, and the Middle East. He has published in education, economics, and policy journals and has had reviews, essays, and poems in newspapers and magazines in Washington, Philadelphia, Los Angeles, and New York. He is a fellow of the American Association for the Advancement of Science and of the American Educational Research Association. He lives in Washington with his wife, Dr. Regine B. Feuer. The Feuers have two grown children.

**Abstract:**

It is now commonly understood by researchers and educators that the quality of teaching matters to the lives and learning opportunities of children, and that the ways teachers are prepared for their profession therefore matters greatly. Efforts to evaluate teacher preparation programs encounter many problems: the definition of desired outcomes, the possibility of unintended effects on teacher motivation and morale of applying various measures, and the potential for perpetuation of disparities in the allocation of teachers to schools serving disadvantaged students, to name a few. Still, the education research and practice communities have made substantial progress in appreciating these challenges and responding with an array of measures and metrics intended to satisfy various goals, including public accountability, program improvement, and information to prospective teacher candidates. This paper draws from a recent report of the National Academy of Education and offers a framework for the analysis of existing evaluation systems and the possible invention of new and better ones.

## INTRODUCTION

In his charge to the Department of Education to take steps toward the improvement of America's teacher education programs, President Obama said last year that "... the vast majority of new teachers – almost two-thirds – report that their teacher preparation program left them unprepared for the realities of the classroom" (Obama, 2014). With this flourish the President was echoing an oft-heard (if somewhat logically challenged) lamentation about the sorry state of teacher preparation – and the lack of credible data about what actually goes on in the institutions that prepare most of our teachers: "... for decades, institutions that prepare teachers have had [sic] lacked the feedback needed to identify their strengths and weaknesses, and had little information on where program graduates go to teach, how long they stay, and how they perform in the classroom. Existing federal regulations on teacher preparation focus on information that is not sufficiently meaningful to preparation programs, potential teachers or potential employers" (Obama, 2014).

One wonders about the validity and tone of the conclusion – that preparation programs don't prepare teachers – given the acknowledgement in the same statement that the data are so flimsy. If institutions are so lacking in relevant information about their programs, what is the basis for claiming they are doing such a lousy job? Indeed, the origins of the President's statement are a bit mysterious, especially because government surveys tell a different story. In the most recent compilation of results from the Department of Education's School and Staffing Survey, or SASS, more than 80% of public school teachers report they were well prepared (44%) or very well prepared (38%) to teach their subject matter; and more than half report being ready to handle classroom management.<sup>1</sup> Earlier reports provided similar findings: in 1998, well over 80% of respondents to a Department of Education survey (which preceded the development of SASS) reported feeling very well prepared or moderately well prepared to implement new methods of teaching and implement state or district curriculum and performance standards (Parsad, Lewis, & Farris, 2001).<sup>2</sup>

But one needs to take these survey results with a grain of salt. Or maybe a whole salt shaker. There is a rich and important literature, spanning cognitive psychology, assessment, and labor economics, on the topic of "expertise," and "proficiency," the details of which are beyond the scope of this paper. One relevant finding is that people become really good at their work through substantial practice and on-the-job learning; though prior formal education in certain relevant skills and knowledge is important, it is not generally considered sufficient (e.g., Lesgold, Feuer, & Black, 1997; Sternberg, 1998). Research on teaching, surely one of the most complex professions, yields similar results, i.e., that it takes time to become good, or competent, or "expert," in the classroom (e.g., Cuban, 2010; Shulman & Shulman, 2004). My sense, therefore, is that we need to be careful in defining what it means to be "prepared" for work, which teachers who responding to SASS and other surveys most likely interpret as something other than the kind of proficiency that the President and others are interested in. In any case, some clarity in the way terms are defined would be useful in such discussions.

But let's assume for argument's sake that the 40% number is real. I would argue then it would be worthy of celebration, not grieving. After all, how does that estimate compare with rookies in other professions? In a recent blog posting, for example, a freshly minted MBA from the Harvard Business School poignantly admitted that "getting an MBA didn't prepare me to DO actual skill work in my industry at all really. It taught me how to see a big problem and make snap decisions on what should be done. But actually DOING work? I'm learning as I go and it's a steep learning curve..." (Perryman, 2013). Even more critically, a prominent lawyer with substantial experience as partner in a major firm and director of private and public organizations (including the Manpower Development Research Corporation) has argued that "... law school graduates, having gone deeply into debt, find they actually don't know how to practice law and increasingly

---

<sup>1</sup> The Schools and Staffing Survey (SASS) is a system of related questionnaires that provide descriptive data on the context of elementary and secondary education and the condition of education in the United States. See <http://nces.ed.gov/datalab/sass/index.aspx>.

<sup>2</sup> I am grateful to Bob Floden for pointing me to the SASS database.

can't find work. They were taught interesting theory but typically weren't taught the skills and even the substantive material they need for their profession... after seven years of this advanced education, our law school graduates largely lack the knowledge and skills to be lawyers" (Roster, 2015). Turning to another profession, one with no shortage of publicly available performance statistics, the National League overall team batting average in 2014 was .249 (Sports Reference Sports Reference LLC, 2000-2015b); among rookies the average was .171 (Sports Reference LLC, 2000-2015a); and as baseball fans know, individual players who get even close to batting .400 are literally off-the-charts (For a delightful and rigorous explanation see: Gould, 2011). Some perspective would be welcome, then, regarding the seemingly disastrous finding that only 40% of new teachers report feeling "prepared."

This would certainly not be the first time that the rhetoric of education policy and reform jumped ahead of the empirical evidence (See, for example: Cremin, 1990; Feuer, 2006; Tyack & Cuban, 1995). But discomfort with overstatement is not an excuse for understatement. So, let me reassure readers that I am not apologizing for what I believe is an untenable status quo: there is sufficient evidence that average performance of teachers could be higher, that the skewed distribution of well-qualified teachers away from disadvantaged and minority schools and children needs to be corrected, that useful knowledge about how to improve classroom learning is often ignored, and that the combined projections of increased student population and an aging teaching force create a somewhat urgent need to revisit the structure and content of our teacher preparation programs (For a good overview see, e.g.: Sass, Hannaway, Xu, Figlio, & Feng, 2012). Surrounding these legitimately worrisome trends in teaching, inequality, and teacher preparation is the shrill sound of alarm bells heralding imminent decline in American education compared to our global competitors, with ominous implications about our long-term economic stature (See, e.g.: Hanushek & Woessmann, 2010). I have argued elsewhere that this rhetoric is outsized with respect to the underlying evidence (Chatterji et al., 2014; Feuer, 2012, 2013a, 2013b). But even if one looks at the international data more dispassionately, the idea that our honored historical tradition of public school excellence is immune from technological, demographic, and global change is foolish and dangerous (For the best recent history see Goldin & Katz, 2008).

It is no wonder, then, that the idea of evaluating teacher preparation programs and holding them accountable for results has gained so much traction in the popular and professional discourse. On the need for better evaluations and accountability my position has been clear, starting with a response in 2011 to the National Council of Teaching Quality plan for a new rating system, continuing with the report of the National Academy of Education (NAEd) in 2013, and then more recently with a commentary in the *Chronicle of Higher Education* (Feuer, 2011, 2015; Feuer, Floden, Chudowsky, & Ahn, 2013). As I argued there, "with high rates of retirement by an aging teaching force and continuing growth in school enrollments, we as a nation need more than ever to focus on how, where, and how well we prepare our future educators... The federal government spends close to \$3-billion a year on a variety of programs aimed at improving the quality of elementary and secondary teachers... [and those] well-intentioned investments in teaching [create] a legitimate demand for evidence that taxpayers' dollars are being well spent ... [but] the metrics now in place have not produced very believable or reassuring answers..." (Feuer, 2015).

Which is why TeachingWorks is such an important endeavor. The initiative facilitates the development and monitored implementation of "high-leverage" classroom practices designed to improve student learning, it relies on a healthy blend of evidence from research and experience, it avoids the temptation of focusing solely on the flaws and imperfections in other teacher professional development systems, it does not fire silver bullets at fast-moving complex targets, and it opens the way to constructive design of teacher preparation strategies along with appropriate metrics for their evaluation.

As part of the partnering and engagement strategy that is central to the TeachingWorks mission, the 2014-15 seminar series focused on "the central challenge of teacher preparation: the demand to show results and impact of initial training on the quality of candidates' effectiveness as

teachers.” Here I will summarize my presentation at the April 2015 seminar,<sup>3</sup> which drew on themes and recommendations of the recent National Academy of Education report (Feuer et al., 2013). The upshot of my argument is that the evaluation of teacher preparation is too important to not be afforded the most rigorous tools, metrics, and methods. And I will advocate for what Greg Duncan and Richard Murnane have called, in a somewhat related context, “sensible accountability” (2014), which I respectfully paraphrase as acknowledging the legitimacy of accountability and striving for methods that provide relevant, timely, and reliable information with limited downside risk of unintended negative consequences.<sup>4</sup>

## THE VALUE OF EVALUATION

Jim March, one of the most enlightened and prolific scholars of education and its intersections with psychology, economics, and politics, once said that “too much accountability is often the symptom of pathology in social systems.”<sup>5</sup> By that criterion, our education (and political) system is beyond pathological. On the other hand, my congenital optimism leads me to a more charitable interpretation: to the extent that one of the most valued traditions of our education system has been to honor and promulgate core principles of democracy, the fact that we hold our teachers and schools accountable for so many aspects of their performance could be taken as evidence that the system actually is, in fundamental ways, working.<sup>6</sup>

Still, and such optimism notwithstanding, there is plenty of room for improvement in what seems to be an accountability system run amok. The so-called “opt-out” movement is borne of frustration with an overzealous reliance on test scores as indicators of student learning and teacher performance; though I fundamentally disagree with the carte-blanche demonization of testing<sup>7</sup>, the fact there is such a movement, especially in a society with a long and deep track record of respect for quantitative data to inform decision making, suggests why a more “sensible” approach is called for.

Formal program evaluation plays a central role in the drama of accountability because intuitive and anecdotal judgments of the impacts of policies and practices on educational outcomes, although at times interesting and informative, are simply inadequate. The ways in which formal systems of data collection, analysis, interpretation, and reporting have evolved over the past century (in particular in the last 50 years) reflect a noble tradition of preferring science over alchemy, of aspiring toward rigorous measurement over casual empiricism. Not coincidentally, many systems of program evaluation require attention to validity and reliability, attributes of inferences most commonly associated with standardized tests used to assess individual learning and teaching. It follows, then, unfortunately, that many of the ills of over reliance on standardized testing threaten to infect formal evaluations of programs that prepare future teachers.

A good way to unravel the complexities of program evaluation generally, and specifically in the context of teacher preparation, is to begin with the question about purpose: what is the goal of the evaluation?<sup>8</sup> The answer surely includes “accountability,” i.e., providing the general public

---

<sup>3</sup> See <http://www.teachingworks.org/training/seminar-series/event/detail/april-2015-seminar-state-context-for-issues-of-teacher-preparation-accounta>

<sup>4</sup> For more on the risks associated with measurement, and what to do about them, see, e.g., Michael Feuer, “Externalities of Testing: Lessons from the Blizzard of 2010,” (2010).

<sup>5</sup> Cited (without date) by Lee Cronbach and colleagues (Cronbach, Atkinson, Bradburn, & Horvitz, 1995).

<sup>6</sup> For a rumination on this topic see Michael Feuer, “The real meaning of ‘accountability’ and ‘trust’ in education,” speech at the Van Leer Institute, Jerusalem, June 2012, published later in the Washington Post (Feuer, June 21, 2012).

<sup>7</sup> See, e.g., the incessant anti-testing advocacy of FairTest: <http://www.fairtest.org/get-involved/opting-out>

<sup>8</sup> I am grateful to my longtime colleague, Patty Morison, for showing that the “why” question can be such a powerful organizing principle. See *Testing in American Schools: Asking the Right Questions* (U.S. Congress, 1992).

with evidence of quality; but there are other equally (or more) compelling purposes, such as institutional self-improvement and consumer information and protection. Thinking about purpose lays the groundwork for understanding risks and benefits of various evaluation systems, how data intended for certain purposes may not be appropriate for other purposes, and why using evaluation data can provoke controversy.

Our premise in the NAEd report was that as a first approximation there are three main purposes for evaluation, as suggested in the graphic below:



What are examples of these uses of evaluation? As we stated in the NAEd report, “The federal government, primarily through Title II of the Higher Education Act, seeks to hold TPPs [teacher preparation programs] accountable for performance by requiring them to report large amounts of data, and by requiring states to compile this information into publicly available “report cards” and to identify low-performing TPPs” (Feuer et al., 2013, p. 3). Similarly, “state governments evaluate TPPs as part of their responsibility to approve programs—an important designation because graduates of approved programs can automatically be recommended for state teacher certification... [and] some teacher preparation programs ... engage in self-evaluation to spur and inform program self-improvement. This can be done by a single institution or through a voluntary network of TPPs” (Feuer et al., 2013, pp. 3-4).

The key point here is that program evaluation has a range of different and legitimate purposes, which leads to the realization that those purposes are not necessarily satisfied by the same kinds of data. Again, from the NAEd report, “the policy challenge is to select the system or approach that is best suited for a defined purpose. For example, although an evaluation alone may not provide all the information needed to hold a TPP accountable for producing well trained and effective educators, it can provide relevant facts to the general public and to education policy makers. Evaluations with more of a consumer information purpose can give prospective teachers data to help them choose from among the broad array of preparation programs and can provide future employers of TPP graduates with information to assist them in hiring decisions. Evaluations for program self-improvement can yield information about an existing program’s strengths and weaknesses, which program faculty and leaders can use to guide innovation and positive change” (Feuer et al., 2013, p. 4).



A very important consideration is the effect of any kind of external evaluation system on the behavior of individuals or organizations being evaluated. Years of research and experience with the unintended consequences of high-stakes testing, for example, provides the empirical foundation for concern over incentives associated with program evaluation and with the possibility that institutions will find ways to look better on the evaluations without necessarily improving the quality of their programs. Such “gaming” effects can be detrimental to the validity of the data and to the prospects for genuinely positive change in teacher preparation, and therefore warrant careful attention.

## MAPPING PURPOSE TO METHOD

Assuming, then, that evaluation has many potential uses, the next perhaps obvious point is to apply the methodology best suited to the determined purpose. By analogy, it would not make sense to use gross domestic product, a very useful summary for purposes of understanding aspects of the macro economy, to an assessment of individual worker productivity in a particular firm. Yet, in the world of educational assessment, testing, evaluation, and accountability, the benefits of good evaluation are often undermined by just this sort of “drift.”

In the NAEd report we offered a simplified tabular representation of the notion of “mapping” of purpose to the information contained in various evaluation methods and to intended and unintended effects. As shown here, the suggestion is to pay close attention to the design of evaluations and the data they require with respect to specific goals; implicit is the notion that some methods may work especially well for certain purposes and less well (if at all) for others.

	Evidence	Inferences	Incentives	Consequences
Accountability				
Information				
Improvement				

To illustrate this framework with more detail, consider the table below, which focuses on the provision of useful information to “consumers,” i.e., prospective teacher candidates deciding where to seek their professional pre-service preparation. Other goals of evaluation are treated similarly (Feuer et al., 2013, see Table 3-2 p. 76).

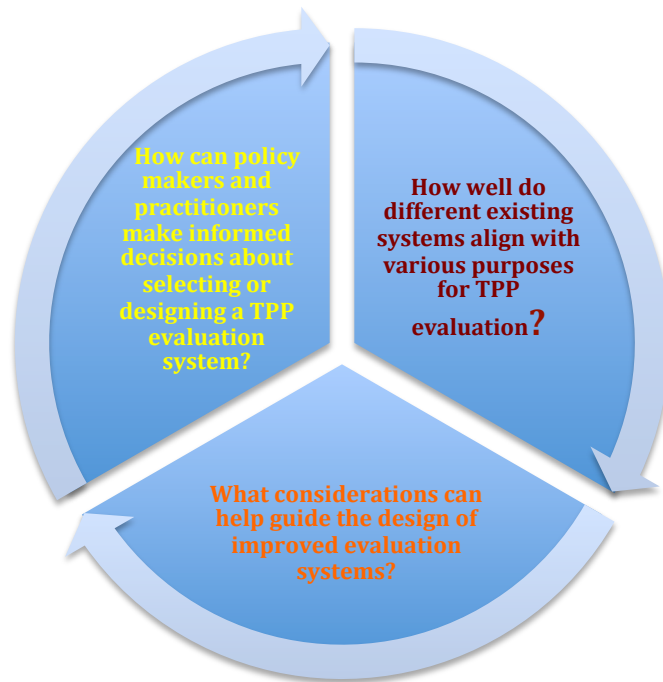
Advantages and Disadvantages of Existing Systems					
<i>If the goal is:</i>	Federal HEA-II	National Accreditation <sup>a</sup>	State Program Approval	Media and Independent Ratings <sup>b</sup>	TPP Evaluations for Program Improvement <sup>c</sup>

Consumer information	<ul style="list-style-type: none"> <li>+ Provides a "one-stop" shopping venue for an overview of data (on limited set of attributes) covering all states</li> <li>- Provides inadequate program-specific information to be of use to prospective students or other "consumers"</li> </ul>	<ul style="list-style-type: none"> <li>+ Provides equivalent of "stamp of approval" from credible body of professionals with deep content and pedagogical knowledge</li> <li>- Simple pass/fail designation does not provide readily accessible information to prospective students, who will not be able to sift through large quantities of dense program-specific reports to make informed consumer decisions</li> </ul>	<ul style="list-style-type: none"> <li>+ Indicates that TPPs have met state standards for teacher education</li> <li>+ Some states provide public reports that include more detailed information about each TPP than simply approved/not approved</li> </ul>	<ul style="list-style-type: none"> <li>+ Enables comparisons across programs on multiple criteria</li> <li>- Is limited to a subset of relevant indicators of quality that may bias overall comparative impressions</li> </ul>	<ul style="list-style-type: none"> <li>+ Faculty and staff are in the best position to provide detailed information (e.g., via website) about program content, etc., that can be useful to prospective students or other "consumers"</li> </ul>
----------------------	---	---	--	--	---

---

## FROM THEORY TO PRACTICE

Can this framework for decisions about whether and how to use evaluation be useful to policy makers and practitioners concerned with the quality of teacher preparation programs? The NAEd report is oriented toward use, and rests on a basic foundation of appreciation of the complexities faced by decision makers, as captured in the following graphic of what we called "the evaluator's dilemma:"



In hopes of providing a bridge between these principles and the real world of program evaluation and its consequences, the NAEd report includes a set of what we call “HFAQ” or “hopefully frequently asked questions,” intended as a kind of checklist for decision makers facing the evaluation challenge (Feuer et al., 2013, p. 94). As suggested below, the questions offer an opportunity to bring greater coherence and, most importantly, an appreciation of the imperfections along with benefits and risks of any particular evaluation program. In our report, each of these questions is followed by elaboration of its meaning and potential value. (The presentation that followed mine, by Jeanne Burns, suggests that such a framework can indeed be constructive.)

**Question 1:** What is the primary purpose of the TPP evaluation system?

**Question 2:** Which aspects of teacher preparation matter the most?

**Question 3:** What sources of evidence will provide the most accurate and useful information about the aspects of teacher preparation that are of primary interest?

**Question 4:** How will the measures be analyzed and combined to make a judgment about program quality?

**Question 5:** What are the intended and potentially unintended consequences of the evaluation system for TPPs and education more broadly?

**Question 6:** How will transparency be achieved? What steps will be taken to help users understand how to interpret the results and use them appropriately?

**Question 7:** How will the evaluation system be monitored?

## A CONCLUDING REMARK

As the dean of a school of education I may risk sounding self-serving, but for me there is abundant empirical evidence that investments in education are the most important a society can make (e.g., Goldin & Katz, 2008). And as a parent accustomed to hearing “duh” from my (now grown) children when they encounter something fairly obvious, it is perhaps a bit risky to note here that we now have strong empirical evidence to corroborate the intuition that teachers matter in the lives and learning of children (Cuban, 2010; Sass et al., 2012). How we prepare teachers, then, and how we know that we are preparing them well, are questions that need to be at the top of the research agenda if we are serious about the continuous improvement of education. Furthermore, in a society steeped in norms of accountability, diffused authority, and checks-and-balances, it is entirely likely that the public will continue to demand evidence that the money we allocate for teacher education is being spent wisely and effectively. The bottom line, then, is that how we evaluate programs of teacher preparation will remain a high priority for researchers, policy makers, and practitioners. It is my hope that the report of the National Academy of Education upon which this paper and presentation were based can be a useful tool and a catalyst for positive change.

## References

- Chatterji, M., Pizmony-Levy, O., Harvey, J., H. Schmidt, W., Noonan, R., Engel, L., . . . C. Rotberg, I. (2014). On the merits of, and myths about, international assessments. *Quality Assurance in Education*, 22(4), 319-338.
- Cremin, L. A. (1990). *Popular education and its discontents*. New York: Norton.
- Cronbach, L. J., Atkinson, R. C., Bradburn, N. M., & Horvitz, D. G. (1995). *A valedictory: Reflections on 60 years in educational testing*. Washington, D.C.: National Academy Press, Board on Testing and Assessment.
- Cuban, L. (2010). How long does it take to become a good teacher? . Retrieved from <https://larrycuban.wordpress.com/2010/04/20/how-long-does-it-take-to-become-a-good-teacher/>
- Duncan, G. J., & Murnane, R. J. (2014). *Restoring opportunity: The crisis of inequality and the challenge for American education*: Harvard Education Press.
- Feuer, M. (2006). Moderating the debate: Rationality and the promise of American education. *Harvard Education Press*.
- Feuer, M. (2010). Externalities of testing: Lessons from the blizzard of 2010. *Measurement*, 8(2-3), 59-69.
- Feuer, M. (2011). Rating teacher education: A fork in the road. *Voices in Education*(March 15, 2011).
- Feuer, M. (2012). No country left behind: Rhetoric and reality of international large-scale assessment. *Educational Testing Service: William Angoff Memorial Lecture*. Available online at [www.ets.org/Media/Research/pdf/PICANG13.pdf](http://www.ets.org/Media/Research/pdf/PICANG13.pdf).
- Feuer, M. (2013a). Spring Issue of The Bridge on STEM Education: Progress and prospects. 43(1).
- Feuer, M. (2013b). Validity of international large-scale assessments: Truth and consequences. In M. Chatterji (Ed.), *Validity and Test Use: an international dialogue on educational assessment, accountability and equity*: Emerald.
- Feuer, M. (2015, 2015). Improving teacher preparation: Right destination, hazardous route, *Chronicle of Higher Education* Retrieved from <http://chronicle.com/article/Improving-Teacher-Preparation-/190239>
- Feuer, M. (June 21, 2012). The real meaning of "accountability" and "trust" in education. *The Washington Post*.

- Feuer, M., Floden, R., Chudowsky, N., & Ahn, J. (2013). *Evaluation of teacher preparation programs: Purposes, methods, and policy options*. Washington, DC: National Academy of Education
- Goldin, C. D., & Katz, L. F. (2008). *The race between education and technology*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Gould, S. J. (2011). *Full house*: Harvard University Press.
- Hanushek, E. A., & Woessmann, L. (2010). The economics of international differences in educational achievement *Working paper 15949*. Cambridge, MA: National Bureau of Economic Research
- Lesgold, A., Feuer, M. J., & Black, A. M. (1997). *Transitions in work and learning: implications for assessment*. National Academies Press.
- Obama, B. (2014). FACT SHEET: Taking Action to Improve Teacher Preparation. *White House Briefing Room*.
- Parsad, B., Lewis, L., & Farris, E. (2001). *Teacher Preparation and Professional Development: 2000*. Washington, DC: US Department of Education, NCES 2001–088.
- Perryman, A. (2013). 10 things I wish I'd known about getting an MBA (before I took out my school loans). Retrieved from <http://postmbalife.com/10-things-i-wish-id-known-about-getting-and-mba-before-i-took-out-my-school-loans/>
- Roster, M. (2015). The New Normal: Facing up to the challenge: It's time to prepare law students for their profession. *ABA Journal*.
- Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., & Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics*, 72(2), 104-122.
- Shulman, L. S., & Shulman, J. H. (2004). How and what teachers learn: A shifting perspective. *Journal of Curriculum Studies*, 36(2), 257-271.
- Sports Reference LLC. (2000-2015a). 2014 Major League Baseball Rookie Players. from <http://www.baseball-reference.com/leagues/MLB/2014-rookies.shtml>
- Sports Reference LLC. (2000-2015b). 2014 National League Team Statistics and Standings. from <http://www.baseball-reference.com/leagues/NL/2014.shtml>
- Sternberg, R. J. (1998). Abilities are forms of developing expertise. *Educational Researcher*, 27(3), 11-20.
- Tyack, D. B., & Cuban, L. (1995). *Tinkering toward utopia: a century of public school reform*. Cambridge, Mass.: Harvard University Press.
- U.S. Congress, Office of Technology Assessment (1992). Testing in American schools: Asking the right questions *OTA-SET-519*. Washington, D.C.: U.S. Government Printing Office.