

The Dissolution of Education Knowledge

Richard P. Phelps

A key component of our faith in progress is the corollary belief that our base of knowledge continually expands: that is, we know what we already know, and we are always learning more. This continual expansion of knowledge requires both that the historical accumulation of knowledge be preserved and that new knowledge be disseminated. A key component of our faith in democracy is the corollary belief that citizens can easily access a wide variety of policy-relevant information needed to make responsible political decisions. A key component of our faith in modern republican democracy is the corollary belief that elected representatives have unrestricted access to and are aware of the variety and scope of detailed policy-relevant research available to them.

For various reasons, these corollaries do not always function as well as they might. Essentially, the public and policymakers have lost access to much accumulated wisdom about education, and over time the squeeze seems to have become tighter.

Such is the case with the research literature regarding the effects of standardized testing on achievement. I write about testing for two reasons. First, it is one of the few topics I know a lot about. Second, if most available information about what is arguably the most important current topic in education can be successfully censored, suppressed, or otherwise removed from public consideration, the same fate can befall any topic. No research literature is safe from dissolution. It is difficult to imagine how any other issue in education could be considered more important.

Hiding in Plain Sight

Given the implementation of the No Child Left Behind (NCLB) Act, one might reasonably assume that the research literature on the effects of standardized testing would have been exposed, made widely familiar,

and meticulously analyzed in the early 2000s. But just the opposite happened. Assertions about a sparse or nonexistent research literature have been common (see, for example, Barth 2006; Education Commission of the States 2006; Olson 2002). Moreover, both opponents and supporters of high-stakes standards and standardized testing assert the claim (see, for example, Cizek 2001; Figlio and Lucas 2000; Hanushek and Raymond 2002a, 2003; Jacob 2001, 2002, 2003; Koretz 1996; Linn 1993; Loveless 2003; Mehrens 1998; Nave, Miech, and Mosteller 2000; and Roderick, Jacob, and Bryk 2002).

As belief in the research literature's nonexistence has spread, efforts to reference it have become less thorough or casually dismissed. After all, why bother to search a literature you believe does not exist? As ironic (and inexplicable) as it may be, the bulk of an available research literature that could help guide our society in implementing its primary and most controversial education policy remains largely unknown.

I. Does Your Research Matter? Only as Much as You Do.

Nobody listens any more. . . . I just want someone to hear what I have to say. And maybe if I talk long enough, it'll make sense.

—Ray Bradbury, *Fahrenheit 451*

Thousands of scholarly journals that publish education research exist, and tens of thousands of education researchers. Perhaps you are a researcher who has published in one or more of those journals. If so, you may recall how much effort you invested toward satisfying the editors and reviewers as your work was shepherded toward publication and, likewise, how much dedication was required to complete your graduate degree.

Was all that effort just for exercise? Or did you endure it because you thought that your work might actually be relevant and that you might contribute something to society? If you select the latter option—that you intended and thought your work to be relevant—you should be heartened to learn about a center created and operated by one of the most prominent national associations of professional educators: in its own words, “America’s one-stop shop for clear, concise, and trusted information about the nation’s elementary and secondary public schools. . . .”¹

This center is real, but identifying it by name would serve little purpose, for its characteristics are far from unique. It maintains a sophisticated Web site and provides guides, glossaries, “go-to backgrounders,” e-mail alerts, online chats, and packets of “rigorous education research that cut[s] through the noise of opinion and politics to get at what really works.” In addition, center personnel write books and articles for popular publications and make presentations—apparently many presentations—at

meetings and conferences throughout the country, under titles such as “High-Stakes Testing and Instruction: What the Research Says.”

A resource this complete would surely provide the public and policy-makers thorough coverage of the research literature on the effects of standardized testing. Here’s what the center had to say, as of January 2007:

We need to recognize a problem presented by much of the current literature—most is opinion that does not report empirical research. . . . Research on the relationship of assessment to teaching has grown since 2002, but it remains overbalanced by essays, anecdotal reports, testimonials. . . .²

A ten-page research review documents this statement with thirty-nine sources: fourteen articles from popular practitioner journals, magazines, or newspapers;³ ten articles or reports from both privately and federally funded research centers;⁴ six articles from a single online journal;⁵ four journal articles that were freely accessible on the Web; two documents from a well-known testing expert’s Web site;⁶ two journal articles from the ERIC database; and one primary data source.⁷ Not included: anything from a book, anything that required a trip to a library, or (with the possible exception of subscriptions to the practitioner magazines) anything that required a fee to obtain. In short, little of the information came from individuals or organizations lacking a dedicated public-dissemination function and public relations staff.

The center thus relegated several hundred studies conducted prior to 2002 (for full citations, see Phelps 2005) to the category of nonexistence or, at best, mere “essays, anecdotal reports, [or] testimonials.” Among them were meta-analyses of several hundred controlled experiments;⁸ meta-analyses of hundreds of effective-schools studies;⁹ multiple case studies;¹⁰ research reviews of incentivized testing programs in higher education;¹¹ and data-driven program evaluations of the effects of national, state, and provincial testing programs.¹²

What did these several hundred orphaned studies have in common? Very few were conducted by organizations that devote substantial resources to marketing their work. They did not morph into media packets, press releases, press conferences, promotional brochures, or cross-marketing campaigns. To find some of these studies, the analysts at “THE place to find up-to-date information” would have needed to look beyond what was available on the tops of their desks. In short, what the center in question presents as “the research” on this topic is a summary of an infinitesimal and highly unrepresentative proportion of the available research.

One could have uncovered similar claims at the Web sites of other well-funded, high-profile organizations and publications. All these actors have asserted that little to no research on the effects of standardized

testing existed prior to a few years ago. They cite few scholars on the topic—typically only a few prominent scholars and pundits, what I call “celebrity researchers.” Yet these organizations’ research efforts are underwritten by millions of dollars in donations from a variety of corporate entities, including some testing firms.¹³

These feeble research summaries apparently assume some or all of the following:

- Any information of importance is easily available on the Web.
- The most-knowledgeable experts can be found at the research organizations with the most money and the highest public profiles.
- Celebrity researchers can be trusted to fairly represent all the existing information on a topic.
- A worthwhile research study will be accompanied by press releases and a marketing campaign. Other research is of doubtful value; ergo, little purpose remains to reading scholarly journals or visiting libraries.

II. Education-Information Solvents

I have my books and my poetry to protect me.

—Paul Simon, “I Am a Rock”

If even this topic—the effects of standardized testing—can be persistently exposed to the glare of journalists’ floodlights these past several years, yet censored and suppressed, then any topic can be. Is the increasing concentration of education-research dissemination in fewer and fewer hands likely to improve education? It may not matter how one answers the question, for the factors and forces working to dissolve and disintegrate the hard-won accumulation of education knowledge seem to be growing only stronger.

Solvents are substances that dissolve other substances, the way paint thinner dissolves paint. Several factors and forces currently serve, by intention or not, to dissolve education information.

Celebrity Education Research

As Daniel Boorstin (1961) put it, celebrities are people well-known for their “well-knownness.” I define celebrity research as mass-marketed research conducted to attract media attention. It is characteristic of only a tiny proportion of all research—that produced by high-profile organizations with public relations shops and access to press microphones as well as in-house, or otherwise dedicated, publication outlets—federal labs, think tanks, well-endowed advocacy organizations, and the like.

Celebrity researchers benefit from guaranteed publication through their organizational sponsors, while noncelebrity researchers must run the gauntlet of often-arbitrary (and sometimes biased) editorial review. Celebrity researchers can get their work published almost immediately upon completion; noncelebrities typically wait months or years.

Celebrity research is disseminated at Internet speed. The sponsoring organizations, operating with less restraint than traditional scholarly journals, can post Web reports in an instant and e-mail them to hundreds of journalists. Under this new order, a nationwide debate over a topic may transpire in the time that a traditional journal takes to circulate manuscripts among its reviewers. Indeed, we now see nationwide debates over a small subset of studies on a topic—studies undertaken by celebrity researchers—that bear little resemblance to the debate that would take place over a full set of studies on a topic. Celebrity debates are no more likely to represent all the available research on a topic than is celebrity research itself.

From the standpoint of society's welfare, the chief drawback to celebrity research is its propensity to displace other research on a topic, even among researchers themselves. Because celebrity research can attain far more attention than the noncelebrity variety, its influence can be parasitic. Researchers new to a topic may cite easily available celebrity research and ignore other, sometimes superior, research that is more difficult to find. If these new researchers accept celebrity researchers' contentions about the research literature—including the extraordinarily common suggestion that no other research worth reading has been conducted on the topic—they may not even bother to search for the other work. The ultimate result is a society denied access to the complete research literature, in favor of a very small and restricted subset. Given this dynamic, only celebrity research will have any influence in the long run, irrespective of quality. The few gaining the best access to the widest audience become the only scholars who matter.

The economist and jurist Richard A. Posner (2001) is perhaps the most prominent critic of celebrity researchers, whom he labels "public intellectuals." Posner's main complaint is their tendency to operate outside their field of expertise, rendering them unaccountable for their claims. Within most academic disciplines, for example, one's work is subject to criticism and review by experts in the discipline or specialty. Public intellectuals, however, whether expert or not, may speak on all sorts of matters, and little of their work is subject to blind review. Few journalists, policymakers, or ordinary citizens can tell the difference.

The research literature on the effects of testing is not the only victim of this syndrome. Myron Lieberman (2007) has documented how well-known think tank analysts have suppressed the substantial research

literature on teacher collective bargaining. He found that these commentators claimed expertise they lack and dismissed as nonexistent research they made no apparent effort to find. Lieberman attributes the blind trust that many place in public intellectuals to a dependence on “credence goods.” Some consumers of research may trust a Brookings Institution analyst because of that organization’s reputation for quality research and, therefore, reliable information. These consumers may not know how to evaluate the same information from an unaffiliated party, for they likely lack in-depth understanding of the topic themselves.

Some of the trust in public intellectuals undoubtedly derives from ideological preference. An individual with generally conservative views may lack the requisite knowledge to evaluate the findings of an education-policy analyst with the Heritage Foundation. Similarly, those of generally liberal views may trust an education-policy analysis from the Economic Policy Institute, whether or not they fully understand it.

Pack Journalism

A friend of mine once wrote a research report, one of thousands of reports published every year on public policy issues. A reporter at the *New York Times* somehow learned about it, gave him a call, and then wrote a front-page article about it. My friend had been doing the same type of work for years without receiving a call from any journalist; in the following weeks, he received more than two dozen inquiries from other journalists throughout the country. He was stunned and sobered to experience the “pack journalism” method of finding a story.



I assume that most education journalists would claim adherence to standard journalistic principles: diversifying sources; not blindly accepting assertions of fact from sources; learning of and telling all credible sides of a story. In my observation, however, many education journalists simply call one or another think tank or public intellectual for their stories, and leave it at that. I wanted to understand why journalists would so thoroughly censor themselves, essentially by restricting their sources to a certain, quite obviously unrepresentative, type, so I asked some journalists. The most-frequent explanation was convenience. Journalists cite those most familiar to them and easiest to reach.

To paraphrase *A Nation at Risk* (National Commission on Excellence in Education 1983): if an authoritarian government had attempted to impose on U.S. education journalists the level of censorship they have voluntarily imposed upon themselves, it would have been viewed as a betrayal of our democracy, an unconstitutional act. A press conference hosted by a think tank or a federally funded research center illustrates how drastically public intellectuals can reduce the public information available on a topic. Journalists are invited and the organization's version of the knowledge base is presented. In the last two events of this type I attended, the meeting rooms were filled with dozens of people; judging from the sign-in sheets, probably half were journalists.

If a few dozen journalists devote half a day of their busy schedules to absorb one organization's take on a specific issue, how much attention will other researchers' perspectives receive? Probably none. After all, the journalists already have the story; they got it at the press conference. Imagine, by contrast, what might happen to the coverage of a critical public policy issue if there were no think tanks or federally funded research centers. Those same two or three dozen journalists might, instead, cover the issue each in his or her own way. Given a variety of approaches, a wide range of research and researchers could be uncovered and presented to the public. The public would then benefit from a wider pool of information and a greater range of perspectives. Society would be far better informed.

Too many journalists accept it on faith, perhaps because of convenience, that the most-familiar sources of information just coincidentally provide full coverage, or at least representative samples, of education issues and points of view. Too many journalists accept it on faith that those researchers who make the most effort to draw attention to themselves also deserve that attention.

Cable Television News

Suppose Professor Smartstuff at Mimosa State College completes a wonderfully well-executed study on a topic highly relevant to current

national decision-making on education policy, and he publishes the results in an academic journal. Will anyone involved in the process hear about it from C-SPAN, the cable television news network? Not likely, unless Professor Smartstuff also happens to be affiliated with one of those public-intellectual groups for which the network provides so much airtime. Observe C-SPAN programming yourself and keep a tally of where the network points its cameras. You will notice, first, that the coverage tends to concentrate on the Washington, D.C., area, where many think tanks, but few of you, are located. Second, you will observe a tendency to focus on “events,” and individual researchers typically do not host “events.”

Independent researchers provide information the old-fashioned way, by submitting a manuscript to a scholarly journal; waiting for reviews; resubmitting a revised manuscript; waiting for more reviews or confirmation; and then waiting for publication. The entire process can take months or even years, but in the end, there’s neither a public relations announcement nor a marketing campaign. The article simply appears in the journal, one of thousands of such journals.

An outlet such as C-SPAN effectively helps those who already have abundant resources to promote themselves even more.

The World Wide Web

In the past, you had to memorize knowledge because there was a cost to finding it. Now, what can’t you find in thirty seconds or less?

—Mark Cuban

If it’s not on the Web, it doesn’t exist at all!

—Sarah Stevens-Rayburn

Mark Cuban (2006) was serious in posing his question; Sarah Stevens-Rayburn (1998) was being facetious. Mark Cuban owns a professional sports team, but made his fortune in high-tech telecommunications. Sarah Stevens-Rayburn is a librarian. Perhaps many readers of *Time* magazine, in which Cuban was interviewed, concurred that almost any information one could need can be found on the Internet, and quickly. Just as likely, I would guess, those same readers would not get Stevens-Rayburn’s quip.

But consider this: by an overwhelming margin, most of humanity’s information accumulated before the World Wide Web even existed. Granted, some documents from the Dark Ages—that hazy historical era prior to the mid-1990s—have been scanned, often imperfectly, into electronic form. But that accounts for a minuscule proportion of the whole.

By far, most of the world's accumulated wisdom is too old to be found on the Web.¹⁴

Hence, the nagging problem of the scholarly journals that still exist only in paper form. Many others have converted their current issues to both paper and electronic form, but some journals' past issues date back deep into those Dark Ages. Moreover, much to the consternation of some consumers, most journals in electronic form charge a fee for downloading their articles.

There also exists a cornucopia of administrative records that may or may not be in electronic form but for the most part have not been posted on the Web. With ecological metaphors, Jeremy White, as quoted by Stevens-Rayburn (1998), puts the situation into perspective:

In the great sea of human knowledge, the Web is like a coastal swamp—shallow and chaotic, full of debris and rotting matter, and often stagnant and murky, if nonetheless always crowded and busy with life. Meanwhile the great sea of human knowledge stretches out undiscovered before it.

Donors to Public-Intellectual Organizations

Funding solicitations I have received from think tanks have asserted the following:

- Not enough (or not any) research is conducted on topics that I care about.
- They are conducting and will continue to conduct such research.
- Without my donations these topics will be ignored.
- The money I give them will be used efficiently.

Think tanks have little incentive to make potential donors aware of previous work conducted on a topic, or even to search for it: awareness of that other work might diminish the sense of need upon which the appeal is based. Typically, the evidence provided to buttress the efficiency claim is “media impact”—how often the group's research or researchers are mentioned in the media. Consequently, think tanks have no more incentive to suggest that a reporter interview an expert outside the think tank—even if that other expert might be more qualified to respond to the inquiry—than an automobile dealer has to suggest that a customer can get a better deal from a competitor. Thus, they may obscure the fact that a valid, or at least plausible, solution has already been proposed.

A donor can boost an organization's impact by subsidizing the production of its research documents. Given that think tanks and federally funded research centers, thanks to their donors, offer reports on education topics free and readily available for downloading, why should anyone

put up with the hassle and expense of scholarly journals? Likewise, donors can subsidize a research organization's book publishing and thus help undercut the sale of books published by those not affiliated with public-intellectual groups. For a student or scholar who will purchase just one book on a topic, the less-expensive, subsidized book will be the likely choice. Compilations of scholarly research published on the open market as books typically cost more than twice as much as those produced by think tanks and other subsidized groups.¹⁵

This narrowing of information sourcing is likely to worsen. "America has been fabulously successful at providing . . . lavishly-funded think-tanks. A growing number of Europeans are seeking inspiration in American-style think-tanks (which are spreading across the continent) . . ." (Lexington 2006).

Trends in Foundation Grant Funding

Some foundations, such as the Broad and Wallace foundations, not only discourage unsolicited proposals but claim to fund them only rarely. Others, such as the Bill and Melinda Gates Foundation, refuse even to consider them.

How, then, do such foundations decide where to place their investments? They may claim to scour the earth doggedly for the most-deserving hidden gems, but peruse their grant lists and you will see the most highly advertised research organizations well represented. Under this new foundation funding policy, it seems that those who already have get more.

To further concentrate resources into the hands of the few, some foundations, such as the Gates and the Rockefeller foundations, plan to maximize their impact by narrowing their focus, funding fewer projects, and giving the smaller number of grantees more money (Face Value 2006). There is logic to that. Some of humanity's problems are big, and solving them requires a big response. Any foundation, even that of Bill and Melinda Gates, has limited resources. If it disperses those resources widely among many little projects, the big problems may never get fixed.

But there are at least two problems with the new approach. One, big grants to solve big problems are most likely to go to big organizations, further concentrating resources. Second, what if the proposed big solution does not work or, even worse, makes the problem even bigger still? It is difficult to see how one could argue with a large investment to find a cure for River Blindness on these grounds. But given the state of affairs in education research and its dissemination, resource concentration may fund popular, yet ineffective, programs just as easily as effective ones. Even worse, given the reduced variety of programs funded, we will end up knowing less about what might work.

III. Where Do We Go from Here?

The lost art of the literature review

A thorough, sophisticated literature review is the foundation and inspiration for substantial, useful research. The complex nature of education research demands [it]. Such scholarship is a prerequisite for increased methodological sophistication and for improving the usefulness of education research.

—David N. Boote and Penny Beile

Few researchers would openly disagree with Boote and Beile (2005). Nonetheless, many manuscript authors apparently make little effort to search the literature, and likewise many journal reviewers put virtually no effort into verifying authors' literature search claims. It happens routinely, despite the fact that any new study is just one study among many, whereas a literature review is supposed to summarize the entirety of the research on a topic. Which is more important to do correctly?

One would think that the wonderful improvements in databases and computer search engines over the past quarter-century would have improved the literature search dramatically. Instead, they may have made it worse (see also Herring 2001; Wong 2004). Too often nowadays, researchers content themselves with a computer search on the most-obvious keywords. Even worse, some merely cite the conclusions of one of these casual, simplistic reviews, making no effort to familiarize themselves with the research base directly.

Relying on keyword searches, however, is inadequate to the task for several reasons. Just one is the matter of which keywords to choose. Different folks can attribute different keywords to identify the same concept. Sometimes the differences in wording are subtle; sometimes they are dramatic. Moreover, different research disciplines can employ entirely different vocabularies on the same topic. It is telling, moreover, that research articles based on extraordinarily superficial literature searches are published regularly in some of the same scholarly venues that minutely scrutinize analytic methodologies. Analytic methodology seems to matter quite a lot; an even minimal effort to survey or understand the research base almost not at all.

There is little glory in performing a literature search. It is mundane, tedious, "lower-order thinking" work. Research glory is found in the end game—the final analysis. A great literature search is unlikely to win one tenure, but a great analysis, even if combined with a superficial search and inaccurate claims about the existing research literature, may well do the job.

If journal editors and reviewers, along with government and foundation research funders, treated literature reviews more seriously, however, they could do much to reverse or at least slow down the dissolution of education information. For starters, they could stipulate that:

- No manuscripts will be accepted for publication without sufficient acknowledgment of the existing research literature.
- No funding will be provided for new research projects without such documentation.
- Full representation of the research literature must be provided in literature reviews; censorship and cronyism should not be permitted.
- For absence- or paucity-of-research claims, manuscript authors and research grant applicants should provide a list of where they have looked; if they haven't looked, the publication or grant application should be denied.
- When it is discovered that manuscript authors have misrepresented the research literature or exaggerated the originality of their own work, a journal should publish a correction and provide accurate information.
- When it is discovered that grant applicants have misrepresented the research literature or exaggerated the originality of their own work, their funding should be withdrawn and they should be barred from applying for more.
- Funders should thoroughly investigate all absence- or paucity-of-research claims made by grant applicants. If they lack the resources to investigate, they should stop funding research until they do.

"Celebrities are different from you and me."

"Yes, they get more attention" might have been Ernest Hemingway's response if E. Scott Fitzgerald had asserted the above. To paraphrase Paddy Chayefsky's script for the film *Network* (1976): given all the possible dangers, drawbacks, and ethical affronts, why should we listen to what celebrity researchers have to say? Answer: because they're on television, dummy!

I propose that there now exist two different, separate education-research worlds whose borders only occasionally cross. Most of us live and work in one world, and what we do bears no influence on policy discussions. Celebrity researchers live and work in another world that matters quite a lot. It is the latter, smaller, celebrity-research world that education journalists cover and to which politicians pay attention.

Traditionally, researchers have carried out research and paid little attention to its marketing. But the success of organizations such as the

center I examined earlier demonstrate how pointless this traditional behavior can be. The center seems to put far more effort into marketing information than into gathering it. Moreover, the only researchers who seem able to get the center's attention are those sponsored by organizations that use the same approach.

To my observation, the marketing of research has become more important than the research itself. If, for a researcher's work to matter at all, a substantial proportion of effort and resources must be devoted to marketing, an equal proportion will be withdrawn from the research effort itself. It's a zero-sum game.

This situation begs the question: are 99 percent of education researchers wasting their time (and the resources of those who pay their salaries)? This question is not facetious.

Education Research: Why Bother?

There must be something in books, things we can't imagine, to make a woman stay in a burning house; there must be something there. You don't stay for nothing.

—Ray Bradbury, *Fahrenheit 451*

Perhaps some participate freely in censoring and suppressing information they dislike. Add the incentives of celebrity to those of self-interest and ideology, and most trustworthy knowledge in education may simply dissolve.

The chief problem is not that some people get attention and others do not. That is a side effect. More important, accurate and useful ideas and information in education—indeed, probably *most* accurate and useful ideas and information—are suppressed and ignored, not even considered, in policy discussions.

Most disturbing to me is that some people who are aware and disapproving of this situation assert that anyone given the same opportunity would behave the same way. Any of us, given the chance, they suggest, would censor or suppress information we do not like or that is inconvenient to retrieve. Presumably, if those now excluded from the dialogue had the power, they, too, would attempt to constrict the flow of education research and information to the most-conveniently retrieved subset they favor. They, too, would try to monopolize the attention of press and policymakers and misinform the public. They do not disapprove of censorship and suppression in general, just the current regime that others, not they, control.

For education researchers, these are very cynical times.

References

- Barth, P. 2006. "High Stakes Testing and Instruction." National School Board Association, Center for Public Education (April 9). Retrieved January 9, 2007, from <http://www.centerforpubliceducation.org/atf/cf/%7B13A13846-1CA6-4F8A-B52E-2A88576B84EF%7D/HIGH-STAKES_TESTING_04092006.pdf>.
- Boorstin, D. 1961. *The Image*. New York: Vintage.
- Boote, D. N., and P. Beile. 2005. "Scholars before Researchers: On the Centrality of the Dissertation Literature Review in Research Preparation." *Educational Researcher* 34 (6): 3-15.
- Bradbury, R. 1953. *Fabrenheit 451*. New York: Simon and Schuster.
- Chayefsky, P. 1976. *Network* [screenplay]. Metro-Goldwyn-Mayer and United Artists.
- Cizek, G. J. 2001. "More Unintended Consequences of High-stakes Testing?" *Educational Measurement: Issues and Practice* (September): 1-3, 7.
- Cuban, Mark. 2006. "Around the Corner." *Time* (March 20): What's Next Forum, 84-90.
- Education Commission of the States. 2006. "Assessment / High stakes / Competency / Selected Research and Readings." Available at <<http://www.ecs.org/ecsmain.asp?page=/html/issuesK12.asp>>.
- Face Value. 2006. "Rockefeller Revolutionary: Judith Rodin Is Shaking Up One of the World's Most Venerable Charitable Foundations." *The Economist* 68 (December 16).
- Figlio, D. N., and M. E. Lucas. 2000. "Do High Grading Standards Affect Student Performance?" NBER Working Paper, No. W7985, National Bureau of Economic Research.
- Hanushek, E. A., and M. E. Raymond. 2002a. Lessons about the design of state accountability systems. Paper prepared for the conference Taking Account of Accountability: Assessing Policy and Politics, Harvard University, June 9-11.
- . 2002b. Improving educational quality: How best to evaluate our schools? Paper prepared for the conference Education in the 21st Century: Meeting the Challenges of a Changing World, Federal Reserve Bank of Boston, June.
- . 2003. "Lessons about the Design of State Accountability Systems." In *No Child Left Behind? The Politics and Practice of Accountability*, ed. P. E. Peterson and M. R. West. Washington, D.C.: Brookings Institution, 126-151.
- Herring, M. Y. 2001. "Ten Reasons Why the Internet Is No Substitute for a Library." *American Libraries* (April).
- Jacob, B. A. 2001. "Getting Tough?" *Educational Evaluation and Policy Analysis* (Fall): 334.
- . 2002. "Accountability, Incentives and Behavior: The Impact of High-stakes Testing in the Chicago Public Schools." NBER Working Paper No. W8968, National Bureau of Economic Research: 2.
- . 2003. "High Stakes in Chicago." *Education Next* 1 (66).
- Koretz, D. 1996. "Using Student Assessments for Educational Accountability." In *Improving America's Schools: The Role of Incentives*, ed. E. A. Hanushek and D. W. Jorgeson. Washington, D.C.: National Academy Press.
- Lexington. 2006. "The Battle of Ideas: Public Intellectuals Are Thriving in the United States." *The Economist* (March 25): 38.
- Lieberman, M. 2007. "Credence Goods and the Accountability Gap." In press.
- Linn, R. L. 1993. "Raising the Stakes of Test Administration: The Impact on Student Performance on NAEP." In *CSE Technical Report 360*. National Center for

- Research on Evaluation, Standards, and Student Testing (March): 3.
- Loveless, T. 2003. Quoted in "New Report Confirms..." U.S. Congress: Committee on Education and the Workforce press release, February 11.
- Mehrens, W. 1998. "Consequences of Assessment: What Is the Evidence?" *Education Policy Analysis Archives* 6 (13) (July): 14.
- National Commission on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform*. Washington, D.C.: U.S. Government Printing Office.
- Nave, B., E. Miech, and F. Mosteller. 2000. "A Lapse in Standards: Linking Standards-based Reform with Student Achievement." *Pbi Delta Kappan* (October).
- Olson, L. 2002. "Accountability Studies Find Mixed Impact on Achievement." *Education Week* 13 (June 19).
- Phelps, R. P. 2005. *Defending Standardized Testing*. Mahwah, N.J.: Lawrence Erlbaum.
- Posner, R. A. 2001. *Public Intellectuals*. Cambridge, Mass.: Harvard University Press.
- Roderick, M., B. Jacob, and A. Bryk. 2002. "The Impact of High-stakes Testing in Chicago on Student Achievement in the Promotional Gate Grades." *Educational Evaluation and Policy Analysis* 24 (4): 333-57.
- Simon, P. 1966. "I Am a Rock." *Sounds of Silence*. New York, N.Y.: Columbia Records.
- Stevens-Rayburn, S. 1998. "If It's Not on the Web, It Doesn't Exist At All: Electronic Information Resources—Myth and Reality." *Library and Information Services in Astronomy III, ASP Conference Series*: 153.
- Wong, M. 2004. "Encyclopedias Gather Dust in Internet Age." *Associated Press* (March 9).

Notes

1. The center's motto: "Practical information and analysis about public education." Here's more from its site:

The Center is a national resource for accurate, timely, and credible information about public education and its importance to the well-being of our nation. The Center provides up-to-date research, data, and analysis on current education issues and explores ways to improve student achievement. . . .

The Center serves as America's one-stop shop for clear, concise, and trusted information about the nation's elementary and secondary public schools, leading to more understanding about public education, more community-wide involvement in public schools, and better decision-making by leaders.

The Center aims to refocus the public debate about public schools. . . . Featuring a website, publications, communication tools, and more, the Center will be THE place to find up-to-date information on public education.

2. The text continues:

There is no doubt—and on this experts agree, whether they oppose or support large-scale high-stakes testing—that accountability systems and the tests on which they depend are in their infancy. . . .

Arguments on both sides are often passionate; they can even be described as polemical. But they beg the question: Where's the evidence?

Little research exists to show how testing is affecting classroom instruction.

Further in the text, the results of polls and surveys are classified as not “empirical” while large-scale databases wholly derived from surveys are.

3. E.g., *Pbi Delta Kappan*, *Educational Leadership*, *American School Board Journal*, *Education Week*.

4. E.g., *Education Next*, Manhattan Institute, Center on Education Policy, the Center for Research on Education Standards and Student Testing (CRESST).

5. The *Education Policy Analysis Archives*.

6. That expert being James Popham.

7. A Public Agenda opinion poll.

8. On extrinsic rewards (e.g., Cameron and Pierce); mastery testing (e.g., Anderson, Kulik, and Kulik); testing frequency (e.g., Bangert-Drowns, Kulik, and Kulik); human productivity (e.g., Csikszentmihalyi); the motivational effects of standards (e.g., Natriello and Dornbusch, Chaney and Burgdorf); the relationship between motivation and goals (e.g., Locke, Shaw, Saari, and Latham); the effect of testing in remediation programs (e.g., Roueche and Kirk; Roueche and Wheeler; Boylen); the motivational effect of literacy standards (Resnick and Robinson); classroom reinforcement programs (Kasdin and Bootzin); and many other, similar topics.

9. E.g., Levine and Lezotte; Clark, Lotto, and Astuto; Taylor, Valentine, and Jones; Rutter; Purkey and Smith.

10. E.g., Southern Regional Education Board; Bamberg and Medina; Eckstein and Noah; Heyneman and Ransom.

11. E.g., Cross; Banta.

12. E.g., in Alberta, British Columbia, Indiana, New York, Ohio, South Carolina, and Texas.

In addition, there remain hundreds of other program evaluations; summaries of administrative data; benefit-cost analyses; studies employing experimental, quasi-experimental, or interrupted time-series designs; and other multivariate analyses, all conducted prior to 2002.

13. The center receives financial help from more than a dozen corporate sponsors, which contributed more than \$1.6 million. Twenty-one state affiliates and eighty-three individuals added more than \$0.9 million to that sum.

14. Granted, much more has been added to the Web, of both good and poor quality, since Stevens-Rayburn and White wrote in 1998. Nonetheless, I would argue that their points remain valid. Much, if not most, of the old journal content that has been added to the Web since 1998, for example, is still obtainable only for a fee. Meanwhile, most public intellectual output is available for free.

15. Full disclosure: I have participated in think tank activities myself, writing a 1999 report for one later converted into a chapter for a subsidized book published by another in 2003. Also, I have been responsible for producing four unsubsidized books for the open market that have met limited competition from subsidized products.

Richard P. Phelps is the author of Standardized Testing Primer (Peter Lang, 2007) and Kill the Messenger: The War on Standardized Testing (Transaction Publishers, 2003). He is also the editor of The Anti-Testing Fallacies (APA Books 2007) and Defending Standardized Testing (Lawrence Erlbaum, 2005).