



# education review // reseñas educativas

editors: david j. blacker / gustavo e. fischman / melissa cast-brede / gene v glass

a multi-lingual journal of book reviews

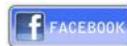
August 20, 2013

Volume 16 Number 8

ISSN 1094-5296

*Education Review/Reseñas Educativas* is a project of the  
College of Education and Human Services of the University of Delaware  
the National Education Policy Center, and the  
Mary Lou Fulton Teachers College, Arizona State University

Follow *Education Review* on Facebook



and on Twitter: #EducReview

## Remembering History:

An Essay Review of *Testing Wars in the Public Schools*

Robert L. Hampel

University of Delaware

Reese, William J. (2013) *Testing Wars in the Public Schools: A Forgotten History*. Cambridge, MA and London: Harvard University Press.

Pp. 298

ISBN 978-0-674-07304-3

Citation: Hampel, Robert L. (2013 August 20) Remembering History: An Essay Review of *Testing Wars in the Public Schools*. *Education Review*, 16(8). Retrieved [Date] from <http://www.edrev.info/essays/v16n8.pdf>

Students memorize but rarely understand their lessons. Several countries surpass us, so we should copy what they do. Test scores are used to evaluate teachers, who vary from good to awful. Local school board elections attract

candidates who vow to bring about sweeping change.

Sound familiar?

End of year oral exhibitions. Separate writing and reading schools within the same building. The Whig Party. A veteran teacher drinks a beer in class. Another uses a rattan to keep the boys in line.

Oh. So we're not talking about American schools today?

No, we're in the 1840s in Boston, Massachusetts, thanks to the masterful storytelling of one of the best historians of education. The unprecedented use of a one hour written exam to gauge student achievement is the central event in this book. Exploring why that change came about and tracing its repercussions makes Bill Reese's excellent book far broader than a case study of one test in one city.

In the 18<sup>th</sup> and early 19<sup>th</sup> century, American students demonstrated their accomplishments at an annual exhibition. Speeches ("orations" in English and Latin), debates, plays and music lasted several hours. A procession might begin and end the ceremony, as would the local ministers' prayers. Open to everyone, the events drew large and appreciative crowds. They judged the achievements of the students and their teachers on the basis of what could be seen and heard. That faith in an occasional display of talent might seem very odd today, but in that era, good deportment was a marker of virtue and merit. Trust and respect accrued from impressive public performances, not just technical prowess.<sup>1</sup>

Whenever an exhibition featured only the best students, it was unclear how well everyone else had done in school, and even the stars could be so thoroughly coached that their accomplishments might be superficial. Doubts intensified when observers felt sure that the teachers' favorite instructional methods were questionable. That was the case for Boston's school "masters" in charge of the "reading schools" (for spelling, history, geography and reading) and the "writing schools" (for mathematics, penmanship, and commercial subjects)—the two branches within each of the 19 "grammar schools" for students who today would attend elementary and middle school. The masters taught the oldest students, relying on memorization and corporal punishment far too often to please reformers like Horace Mann and Samuel Howe. As a group the masters were well-educated, politically connected, and self-confident, so the modest oversight from the School Committee and its "visiting committees" for each school never shook their faith in rote recall and occasional

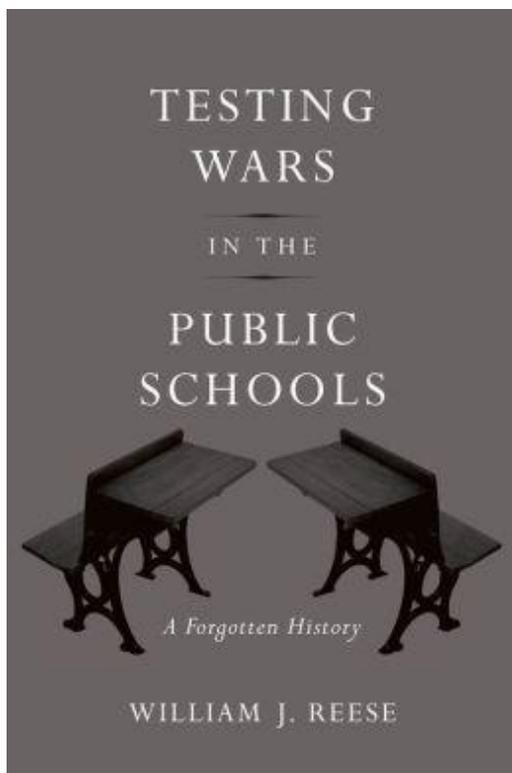
---

*History of a Founding Ideal from the American Revolution to the 21<sup>st</sup> Century* (Ithaca, N.Y.: Cornell University Press, 2013), Ch. 3. Also see Kett's insightful comments on Horace Mann's misgivings about the unbridled individualism fostered by exhibitions — youthful competition and "unnatural display" for the sake of public approval were morally dangerous, not just pedagogically unwise. Kett, Merit, pp. 99-111.

---

<sup>1</sup> For a superb recent account of (college) exhibitions, see Joseph Kett, *Merit: The*

floggings. “What is the square root of 5/9 of 4/5 of 4/7 of 7/9?” That was one question on the written exam created in 1845. Mann and his allies convinced the School Committee to test each school’s oldest students—the masters’ students—with a surprise one hour exam in place of the customary informal observation. The results were dismal. The average score was 30% correct, with especially low marks on the items calling for explanations and applications—for instance, 38% of the students knew the date of Thomas Jefferson’s embargo but only 28% could define the word embargo.



A report with dozens of tables revealed how each grammar school did on each test question. Four of the masters lost their jobs thanks in part to their students’ low marks. The quarrels between the masters

and the reformers receive close attention from Reese. He traces the back-and-forth in more detail than some readers will want (the same test received only a few pages in Stanley Schultz’s thorough history of ante-bellum Boston schools, and Jonathan Messerli did not dwell on it in his biography of Mann).<sup>2</sup> I think the painstaking account is worthwhile because tests were inextricably linked to other issues, as Reese indicated when I asked him how he could write a history of testing without also taking on the history of grades, promotion and honors.

Reese responds:

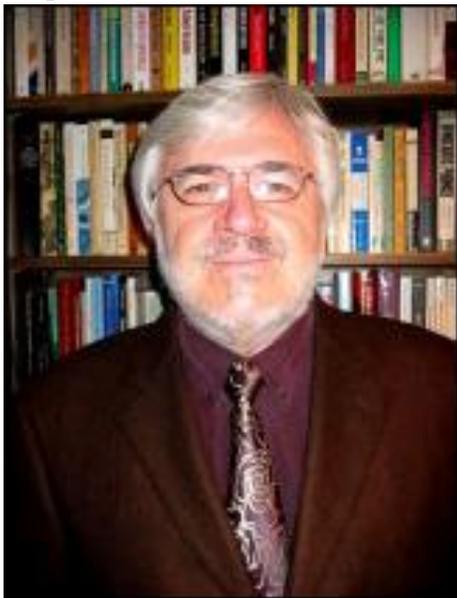
*Testing was intertwined with several interlocking, often mutually-reinforcing reforms. Mann and Howe didn’t just popularize competitive testing; they also promoted age-graded classrooms, which meant changes in school architecture, more specialization in teacher training, the hiring of more women as primary school teachers, and so forth. Changes outside of schools also accelerated the movement toward more written tests. Cheaper printing costs led to the adoption of more common textbooks; more uniform reading materials made testing easier. Textbooks often grew in size, filled with questions and answers. More testing, in*

---

<sup>2</sup> Stanley K. Schultz, *The Culture Factory: Boston Public Schools, 1789-1860* (New York: Oxford University Press, 1973), Ch. 6; Jonathan Messerli, *Horace Mann: A Biography* (New York: Knopf, 1972).

*turn, led to more written assignments in classrooms, weakening the time-honored tradition of favoring oral communication over the written word in assessing academic achievement.*

*And keeping better records, easier said than done, preoccupied many school administrators in the 19<sup>th</sup> century. The growth of competitive written tests and increased reliance on statistics undoubtedly affected the basis for calculating grades, assigning prizes, and determining promotion in different times and places.*



I found myself yearning for even more information on the political context of the educational controversies. As **Reese** notes, the Whig party dominated Boston elections in the mid-1840s, but he does not explore the factions within the party. The expansion of slavery split the party into “cotton” and “conscience” Whigs; did

those divisions on national policy also shape consideration of local issues?<sup>3</sup>

Moreover, the new Liberty Party siphoned Whig votes, as did nativist renegades, whose mayoral candidate in December 1844 deprived the Whigs of their customary majority.<sup>4</sup> I asked him why he did not analyze several types of voting—the turnout and results for School Committee elections, for instance, or the yeas and nays within the School Committee.

*The Boston School Committee minutes often identified how members voted. Like any source, the minutes were helpful but lacked the details found in newspaper accounts of meetings, which I read in abundance. Why particular committeemen voted as they did is often impossible to determine, since they were never quoted anywhere or recorded their views. Whigs dominated Boston in the 1840s but disagreed about many policies. Their disagreements on the nature and uses of tests, or the reappointment of teachers or*

---

<sup>3</sup> Kinley J. Brauer, *Cotton versus Conscience: Massachusetts Whig Politics and Southwestern Expansion* (Lexington: University Press of Kentucky, 1967), 48-76.

<sup>4</sup> Michael F. Holt, *The Rise and Fall of the American Whig Party: Jacksonian Politics and the Onset of the Civil War* (New York: Oxford University Press, 1999), 228.

*administrators, produced considerable commentary in the daily press.*

*Committeemen ran for office on political tickets, and newspapers in the 19<sup>th</sup> century were intensely partisan. So, like every source, newspapers have to be read critically. Reporters and editors nevertheless routinely published valuable information on local schools. They identified the party labels of candidates during election season and later named the winners and losers. They made the flesh and blood conflicts over testing palpable.*

*Happily for the historian, Howe and Mann made lots of enemies, who were unafraid to voice their opposition to them in public. That Mann, Howe and their allies also corresponded extensively about the exams, the grammar masters, and the school committee provided me with a clearer sense of their motives and political strategies, which their enemies mostly had to speculate about. They offered a view from behind-the-scenes, another window to understand what was happening in Boston.*

What Reese demonstrates in *Testing Wars* is the enduring value of old-fashioned political history, culling manuscript collections, newspaper columns and pamphlets to analyze tactics, understand motives, and trace the unfolding of important events. A political historian well versed in quantitative methods could now supplement Reese's narrative by analyzing the voting behavior of the rank-and-file and the policymakers.

The final two chapters move beyond Boston to sketch the rapid adoption of written exams across the country. The pace of change varied from place to place, but exams were ubiquitous long before the rise of I.Q. testing in the early 20<sup>th</sup> century. From newspapers, magazines, professional journals, and national meetings, school administrators learned what their counterparts were doing; the popularity of written exams had nothing to do with state or federal mandates or guidelines. Other hard-to-trace but obviously important influences accelerated the spread of testing: cheaper and better pens, pencils, slate, chalkboards and paper so student could write during class; more and more written exams in colleges; and a growing faith that merit need not take the form of dazzling public speeches<sup>5</sup>. One revealing sign of the extent of change: by the 1870s many publishers sold books to coach students to do well on the exams (aspiring teachers also faced written exams, and several publishers wrote test prep books for them). Another useful indicator is the wave of complaints in education journals and the popular press in the 1870s and 1880s deploring excessive

---

<sup>5</sup> The classroom recitation survived even as end-of-year exhibitions waned. And high school grades by the 1880s still counted attendance and deportment as heavily as academic achievement, which included recitation marks, not just scores on written exams. See Reese, *The Origins of the American High School* (New Haven: Yale University Press, 1995), Ch. 10.

testing. Reese traces the ensuing reaction that reduced the frequency of “high stakes” testing for high school entrance, graduation, and promotion from grade to grade at all levels (and the backlash wasn’t because sophisticated statistical analyses revealed their shortcomings—I am struck by how so few people ever proposed the equivalent of multivariate analysis, but then again, statistics as a field was just taking shape, and it first took root in other fields)<sup>6</sup>. Tests with smaller consequences, however, continued to proliferate. What did not shift was the nature of the test questions. I was surprised to see the same types of question decade after decade, so I asked Bill if he found any innovative tests, defining innovative loosely to include True/False, multiple choice, and matching.

*The test items were usually rote, short-answer, and drawn from the textbooks, which framed the core of the curriculum. Multiple-choice and other types of “objective” tests became popular in the early 20<sup>th</sup> century, and they received a huge boost with the invention of machine readable answer sheets [in the 1930s—RH]. What I found astonishing was how Mann and Howe devised some test questions to discern whether teachers emphasized rote teaching or understanding.*

What was not surprising was the wealth of evidence Reese used in each chapter. All

of his books display prodigious research. He always makes remarkably good use of education journals, annual school reports, and autobiographies. In this book, more than his earlier work, he also mines local newspapers. In chapter six alone, I counted 25 different papers, including small ones like the Idaho Avalanche and the Weekly Arizona Miner. And so I asked Bill why there is no description of his research strategy (unlike the long “essay on sources” in his previous book, *America’s Public Schools*). How did he decide which school reports, pamphlets, memories, and newspapers to read, apart from the obvious necessity to focus on Boston materials for the early to mid 1840s:

*A method to my madness? Originally, I planned to write a book on promotion practices in urban schools in the 20<sup>th</sup> century, with an opening chapter on the 19<sup>th</sup> century. After a year or two of research, I became fascinated with the origins and spread of written examinations and thus shifted my focus.*

*In 1995 I published a book on the birth of the American high school, so I knew that entrance tests to public secondary schools had once been common and were often controversial. Among other things, entrance tests tried to lift standards in the lower grades, but I knew little about how tests entered or were used in pre-high school grades.*

*So I tried to read broadly in national, state, and local sources. I read intensively in a breadth of primary sources on Boston,*

---

<sup>6</sup> Theodore Porter, *The Rise of Statistical Thinking, 1820-1900* (Princeton: Princeton University Press, 1986), Chs. 1-2, 6.

*once I realized its importance to the history of testing and recognized its utility in framing the arguments in my book. I read the annual reports of many urban school systems, from Maine to California, to uncover national trends.*

*Professionally-oriented periodicals such as the Common School Journal and the New England Journal of Education, as well as local and state-level magazines, helped track developments. So did articles that appeared in prominent national magazines such as Harper's or local newspapers such as the Cincinnati Daily Gazette. Once patterns began to emerge about the origins, politics, and spread of testing, I sketched a rough outline of chapters and began to write, and revise, and revise.*

Some readers might wish Reese had revised his Epilogue to comment on contemporary testing wars. Could the Lessons of History shed light on the current battles? Reese offers a few predictions—testing and politics will continue to intertwine, for instance—but I am glad that the bulk of the Epilogue isn't about the present or the future. Reese reminds us that many aspects of the current debates also marked the 19<sup>th</sup> century: the fretful overseas comparisons, the hyperbolic rhetoric of exam friends and foes, the use of tests to castigate teachers, and the faith that the best evidence of student achievement is numerical. Too many policymakers still act as if today's problems and solutions are brand new. On the other hand, Reese reminds us of the simple but powerful

point that testing bears the marks of the era when it took hold, and the 19<sup>th</sup> century in many respects was a very different world from the 21<sup>st</sup> century. Boston's writing school master Frederic Emerson could not drink a beer in class today or whip a student who was whispering in class, and the opposition to standardized testing in our generation will take different forms than Emerson used in the 1840s.

### About the Reviewer

Robert L. Hampel is a Professor and former Director, School of Education, University of Delaware. Bob Hampel is an historian of education (Cornell PhD). He served as Secretary-Treasurer of the History of Education Society from 2002 to 2011, and as Interim Director of the School of Education at the University of Delaware in 1998-2000 and 2010-2012. He wrote *The Last Little Citadel: American High Schools since 1940* (Houghton Mifflin, 1986) and several dozen articles and chapters on educational reform past and present.



Copyright is retained by the first or sole author, who grants right of first publication to the  
*Education Review*.

*Education Review/Reseñas Educativas* is a project of the  
College of Education and Human Services of the University of Delaware  
The National Education Policy Center <http://nepc.colorado.edu>, and the  
Mary Lou Fulton Teachers College, Arizona State University



Editors

David J. Blacker

[blacker@edrev.info](mailto:blacker@edrev.info)

Gustavo Fischman

[fischman@edrev.info](mailto:fischman@edrev.info)

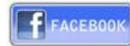
Melissa Cast-Brede

[cast-brede@edrev.info](mailto:cast-brede@edrev.info)

Gene V Glass

[glass@edrev.info](mailto:glass@edrev.info)

Follow *Education Review* on



Facebook and on Twitter:



#EducReview