

# JoLLE@UGA

JOURNAL OF LANGUAGE & LITERACY EDUCATION

Volume 10 Number 1

Spring 2014

Editor  
Stephanie Anne Shelton

<http://jolle.coe.uga.edu>

---

## Review of *The Mismeasure of Education*

Horn, J., and Wilburn, D. (2013). *The Mismeasure of Education*. Charlotte, NC: Information Age Publishing, Inc.

ISBN 978-1-62396-391-0

Pages: 271

William G. Wraga  
University of Georgia

---

Please cite this article as:

Wraga, W. (2014). Review of the mismeasure of education. *Journal of Language and Literacy Education* [Online], 10(1), 216-220. Retrieved from <http://jolle.coe.uga.edu/>

In *The Mismeasure of Education* Horn and Wilburn present an analysis of recent efforts to deploy value-added modeling as a policy tool in education reform and attempt to place these efforts in the context of the history of educational testing in the United States. On the former the book succeeds spectacularly; on the latter it falls far short.

Horn and Wilburn contrive a historical account of testing in American education that is incomplete and oversimplified. They depict the development of standardized testing as animated by a monolithic intent to sort and fit students to the socioeconomic status quo:

What resulted from that first generation of testing and sorting was a system that continues today to provide 'scientific' rationalization for the creation and maintenance of measures whereby children of the privileged display test results, on average, consistently higher than those children under the privileged on tests that were devised to show as much. By using measures stamped with the seal of science, then, high test scorers are guaranteed seemingly legitimized access to the a [sic] legacy of privilege that accompanies higher performance, thus reproducing social and economic dominance by descendants of the middle class elites who first established their dominion in the Colonies during the seventeenth and eighteenth centuries. (p. 14)

Granted, an element of truth resides in this representation; but, in an act of historical exaggeration, Horn and Wilburn confuse that element with the whole entity of testing and assessment. They achieve this by cherry-picking evidence that supports their perspective, by ignoring evidence and research that contradicts it, and by misrepresenting other historical developments.

For example, in a section titled "Zealots for the Elimination of the Unfit," they associate Franklin Bobbitt's (1918) *The Curriculum* with Tayloristic scientific management, although in that book Bobbitt considered Frederick Taylor's principles as a management system a "relative failure" (p. 84). They associate the Commission on the Reorganization of Secondary Education's (1918) *Cardinal Principles of Secondary Education* not only with social efficiency-social control doctrine, but also with standardized testing and even with eugenics, although the report advocated none of these practices. They also imply that the priorities of the 1893 Committee of Ten Report were somehow more humane and egalitarian than these latter initiatives, despite the fact that it condoned the exclusion of over ninety percent of the adolescent population from the secondary school. In reconstructing their history of testing, Horn and Wilburn rely too heavily on the social efficiency-social control thesis that dominated curriculum history scholarship for three decades, but which has been revealed as considerably overdrawn (Fallace & Fantozzi, 2013). They also overlook historical accounts of resistance to the introduction of group testing to the military and to public schools (Chapman, 1988), depicting the testing movement as a unilateral and efficient hostile takeover of schooling. And on the frequently noted limitations of standardized tests featured in the psychometric literature, and on the use of other sources of information about student learning that had emerged by the 1930s, they are silent.

Although it would be safe to suggest that a strain of what Horn and Wilburn attribute to testing practice certainly existed during the early twentieth century, the entire reality was much more complex than their narrative demonstrates. The reader would do well to beware that a fair-

mind history of testing is not to be found in this book, and should seek such from other sources (e.g., DuBois, 1970; Chapman, 1988; Madaus, Russell, & Higgins, 2009).

Horn and Wilburn's reductionist history of testing may have been provoked by their apparent zeal to expose the egregious limitations of value-added measurement as a form of educational assessment. The detailed account they offer of the legislative history and implementation of value-added modeling as the main component of the educational accountability scheme in Tennessee, institutionalized in the Tennessee Value-Added Assessment System (TVAAS), documents the sheer careerism and cynical political maneuvering of elected state officials in chilling fashion.

The crux of Horn and Wilburn's argument is that, "In using input-output analyses like TVAAS, the focus on teacher effect is substituted for education funding and resource allocation as the educational input and student test scores are substituted for student learning as educational outputs" (p. 149). As inequities in funding are left unaddressed, and as students' SES to a significant extent influences their standardized test scores, input-output analyses function not as "reforms" that improve student achievement and life chances, but rather as mechanisms to maintain educational, social, and economic inequality. While it may be difficult to establish that this is the intent of recent accountability reforms, Horn and Wilburn make a compelling case that it will be the pernicious effect.

After recounting the implementation of valued-added modeling in Tennessee, Horn and Wilburn offer an incisive analysis of the methodological flaws in such production function measurements. This analysis amounts to an indictment of the Tennessee initiative and of value-added models that William L. Sanders has marketed nationally. These flaws make the widespread use of value-added modeling in educational assessment not only educationally problematic, but also psychometrically indefensible.

Among other things, Horn and Wilburn report that the consensus among researchers who have systematically examined value-added modeling as an education assessment tool is that the practice should not be used for high stakes purposes; that in Tennessee principals have adjusted their ratings of classroom observations of teachers based upon TVAAM scores; that despite the promising rhetoric behind TVAAM, the State has effectively failed to meet its own explicit goals for school reform; and that TVAAM scores are inconsistent with and significantly more positive than NAEP scores for the state.

Horn and Wilburn's principal recommendation for using value-added modeling in education assessment is that its use for high stakes consequences cease and be replaced by strictly diagnostic uses. This would put value-added modeling in line with the principal defensible use of standardized test scores that emerged from the field of psychometrics--as a diagnostic tool. They also recommend that "policy makers and implementers must focus on learning, rather than on test scores for accountability purposes" and that resources be directed away from overreliance on high stakes standardized assessment to the "use of locally produced curriculum and assessments" (p. 215). They also express concern that the vast and entrenched testing and assessment industry in the US remains unmonitored and unregulated, despite its significant impact on the lives of children and families.

A brief word may be in order about the production of this book. Unfortunately, dense text, numerous typographical errors, and other formatting flaws distract from book's important findings. The authors also too frequently allow their anger with the political and educational debacle of value-added modeling in education, although justified, to get the better of their academic analysis as they resort to intemperate language and labeling. These imperfections, and the problematic historical contextualization that they present at the outset of the volume, detract from an otherwise trenchant and timely expose of misguided and miseducative corporate-style school "reform."

## References

- Bobbitt, F. (1918). *The Curriculum*. Boston: Houghton Mifflin.
- Chapman, P. D. (1988). *Schools as sorters: Louis M. Terman, applied psychology, and the intelligence testing movement, 1880-1930*. NY: New York University Press.
- DuBois, P. H. (1970). *A history of psychological testing*. Boston: Allyn and Bacon.
- Fallace, T., & Fantozzi, V. (2013). Was there really a social efficiency doctrine?: The uses and abuses of an idea in educational history. *Educational Researcher*, 42(3), 142-150.
- Madaus, G., Russell, M., & Higgins, J. (2009). *The paradoxes of high stakes testing: How they affect students, their parents, teachers, principals, schools, and society*. Charlotte, NC: Information Age Publishing.