



education review // reseñas educativas

editors: gene v glass gustavo e. fischman melissa cast-brede

a multi-lingual journal of book reviews

November 23, 2011

ISSN 1094-5296

Education Review/Reseñas Educativas is a project of the National Education Policy Center <http://nepc.colorado.edu>

Follow Education Review on Facebook



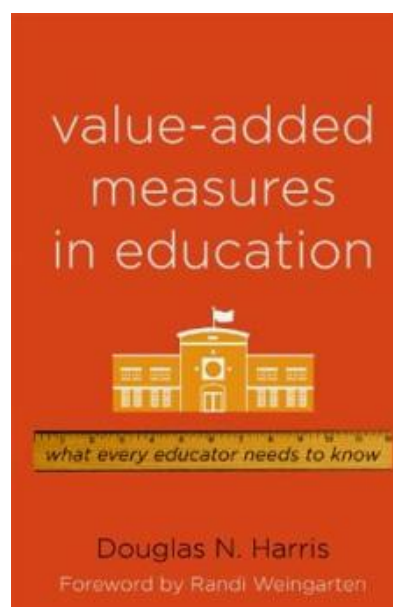
Harris, Douglas N. (2011) *Value-added measures in education: What every educator needs to know*. Cambridge, MA: Harvard Education Press.

Pp. ix + 276

ISBN 978-1-61250-000-3

Reviewed by Clarin Collins &
Audrey Amrein-Beardsley
Arizona State University

As pointed out in Randi Weingarten's foreword in *Value-Added Measures in Education*, Doug Harris sets out to shift the conversation about value-added from the technical language of econometricians and statisticians into a more common explanation especially for "those directly responsible for improving our schools," (p. vii). With the goal to alleviate some of the confusion and mistrust in value-added, Harris attempts to integrate information in a "clear, and accessible" way to both policy makers and educators (p. 3). Harris packs *Value-Added Measures in Education* full of information about value-added statistics and modeling, value-added accountability, uses and misuses of value-added data, its potential



Citation: Collins, Clarin & Audrey Amrein-Beardsley. (2011 November 23) Review of *Value-added measures in education* by Douglas N. Harris. *Education Review*, 14. Retrieved [Date] from <http://www.edrev.info/reviews/rev1126.pdf>

impact on teaching and learning, and value-added misconceptions and recommendations.

In the introduction, **Harris** explains his cardinal rule of accountability, which states that people should only be held accountable for what they can control (p. 4). He continues to explain how current test-based accountability measures violate this rule as teachers have relatively little control over students and various factors that contribute to student learning and test-taking abilities (p. 5).

Additionally, Harris explains that tests do not fully capture the contributions that teachers and schools make to student achievement (p. 6). However, value-added measures, according to Harris, offer a desirable deviation from such archaic accountability systemic issues.



From this comes his cardinal conclusion: value-added models, while imperfect, are good enough. Harris supports the adoption of value-added systems as the nation's new and improved educational accountability system because value-added measures provide more accurate information about individual student academic growth by taking students' academic starting points into consideration (p. 43). Focusing on student growth using value-added measures, allows student achievement to be more accurately and fairly captured, unlike traditional snapshot tests common under No Child Left Behind (NCLB). While the topic of value-added measures is certainly a controversial one, Harris strives to maintain a "middle of the road" position; playing it safe, all the while pointing out the pros and cons of value-added. With this, Harris offers the most comprehensive review of pros and cons to date. This is one of the strengths of this book. However, the logic Harris follows is at times confusing, even for readers who are familiar with value-added and the research. As such, this book's greatest strength comes with its greatest weaknesses.

One of the most controversial debates surrounding value-added is whether or not demographic information should be accounted for when calculating value-added. Harris believes that demographics are best left unaccounted, much like the opinion of the creator of the most popularly

used Education Value-Added Assessment System (EVAAS), William Sanders. Harris notes that accounting for demographic information could reduce expectations for particular students, resulting in a wider achievement gap between schools serving low-income students and schools serving their more advantaged peers (p. 75). Harris also explains if demographic information is associated with learning, the influence of demographic information should be relatively consistent, or “cancel out” during each year of learning (p. 76).

Harris further elaborates on how the inclusion of demographic information can result in incorrect causal assumptions about race/ethnicity, income, and poor academic success, when in reality race and income are proxies for learning (p. 74). However, Harris writes that “research evidence” shows that disadvantaged students grow at slower rates (p. 75), countering his logic that these factors should go away over time. He adds that summer learning loss is also highly associated to demographics. This too does not follow the pattern of constant influence, complicating and confusing Harris’s position further. Additionally, Harris notes that disadvantaged students grow at slower rates (p. 75) and as such are going to have a more difficult time achieving comparative growth. In the end, to account or not to account for demographics is still the question.

In discussing the systematic errors caused by students who have serious out-of-school problems, Harris dismisses this as all schools have students with similar home situations, and that these errors also “cancel out” (p. 101). Harris’ logic might work at the school-level, but for teachers evaluated by teacher-level value-added, whose scores will be negatively impacted by such variables beyond the control of the teacher, this logic becomes more perverse. This is also where things get most confusing as Harris changes his cardinal rule from holding people accountable for what they can control to holding people accountable “even for factors only partly within their control” (p. 82).

On the topic of uncontrollable school resources (school size, administration, student absences), for example, Harris explains how more sophisticated value-added measures should allot “bonus points” to schools with such uncontrollable variables. He provides a hypothetical

example that if research shows that smaller class sizes have higher achievement, then the value-added measures can be “adjusted” to “offset” the disadvantage for larger class sizes (p. 87). While value-added measures can be adjusted, real life situations do not “offset” or provide bonus points for students, teachers, or schools so influenced by uncontrollable resources. In addition, as research already suggests that teachers teaching small class sizes are more likely to regress to the mean, nowhere does Harris address the very real adjustments already needed. For now, such teachers are excused because the systemic and random errors are too large.

Explaining many other statistical errors inherent in value-added measures, Harris does, however, provide easy-to-follow analogies and charts to explain the various errors. Yet all the while he still holds firm to his prevailing mantra, that even given the errors and imperfections, value-added measures, are still good enough. In one of the best sections of the book, Harris discusses teacher-level value-added. This is most valuable given the current federal and local push to adopt teacher value-added for evaluation and accountability purposes. Nevertheless throughout the book Harris bypasses many opportunities to point out the issues and concerns presented in this chapter that violate the very core of value-added and its practicality.

Harris cites a study conducted by Jacob and Lefgren (2007) who evidenced how value-added measures are only able to differentiate between teachers with very low or very high value-added scores (p. 140). He points out the instabilities prevalent when analyzing value-added from year-to-year, which means that teachers are falling into different effectiveness levels from one year to the next although their instruction has not changed. But, again, Harris does not pay the appropriate amount of attention to this value-added blemish that it really deserves. The very real fact that 25%+ teachers of virtually the same effectiveness levels are consistently being labeled as differentially effective is most problematic (Baeder, 2010; Baker et al., 2010; Haertel, 2011; Koedel & Betts, 2007; Papay, 2010), and this, in itself, counters Harris’ recurring logic. Unless value-added measures can capture the teacher effectiveness construct more consistently and reliably, they cannot be used for much decision-making,

especially if high-stakes consequences are to be attached to such decisions. They are not, in fact, good enough.

The lack of reliability, stability, or consistency researchers are finding across value-added years and between value-added subjects, is quite possibly the biggest problem confronting the practicality and applicability of value-added. Yet Harris downplays these issues quite a bit, even noting at times that this matters less in the policy world as it does in the research field (p. 105, 107), and he holds tight to his recurring logic that even though this too is imperfect, it also may be good enough for policy purposes. To counter this, Harris emphasizes that a teacher's value-added scores should be used in tandem with and other, corresponding indicators of teacher effectiveness. As long as the other indicators measure the teacher effectiveness construct, as he suggests value-added does, using multiple measures to capture teacher quality makes sense. Most if not all value-added researchers agree with this, especially because using multiple measures to measure most anything is in line with the standards of educational measurement. Various measures of teacher effectiveness are needed to capture the "multifaceted" responsibilities necessary in the teaching profession (p. 154), and it is most critical to appreciate how "multifaceted" the teaching profession is. Yet after citing a series of moderate correlations between, for example, teachers' value-added and supervisor evaluation scores, Harris again makes the argument that these correlations, while less than we would like, are strong enough to complement value-added output. Again, value-added is "good enough."

Harris also repeatedly notes that value-added should not be used as a sole indicator of teacher effectiveness. More attention could have been paid to the very real potential and actuality that school districts across the country are already using value-added as the primary decision-making indicator. This includes one of Harris' featured case-study districts, the Houston Independent School District (HISD). In HISD, preliminary evidence suggests that although teachers are being evaluated using both EVAAS value-added scores and supervisor evaluations, district leaders are increasingly skewing observational data to align with and match EVAAS output (Capital Hill Briefing, 2011).

While the district is taking a multiple measures, multifaceted approach, EVAAS output are trumping the other measures that are to be used to capture what it means to be an effective teacher. In Harris's book, one written in particular for "those directly responsible for improving our schools," (p. vii), the seriousness of this potential was seriously understated. If, as Harris points out, the goal here is to improve the understandings of value-added measures (p. 10) for teachers, administrators, and state officials, here lies a missed opportunity.

Another missed opportunity is that Harris mentions only briefly how value-added measures are being used to hold educators to different standards by placing increased pressure on only those teachers evaluated by value-added. While it is "unfair to hold educators in a given district to different standards," (p. 164), Harris marginalizes the fact that teachers who do not teach core curriculum classes in grades three or four through eight are typically left out of the "systemic reform" associated with value-added. This in itself makes the value-added approach limited as it only really impacts approximately 30% of a school's teachers (p. 122). Yet Harris might argue that because traditional snapshots did the same, this issue still favors value-added. Harris brings the book to a close with nine value-added recommendations from his perspective as a researcher, educator, parent, and advisor to local, state, and national policy makers. He also introduces five different case studies highlighting how value-added measures are being used in the state of Tennessee, Dallas Independent School District, Milwaukee Public Schools, Houston Independent School District, and England.

And while Harris stresses throughout the book that relatively little is known about whether value-added measures actually improve teaching and learning, districts and states are rampantly moving forward with value-added measures in response to federal and state policies. Harris also strongly cautions others against using value-added output for high-stakes purposes, and with this most if not all value-added researchers would also agree. Though Harris does go back and forth on whether he would use value-added output to make merit pay decisions. Regardless, given Harris's implicit and explicit messages throughout the book, he supports the value-added movement for two main reasons: (1) value-added offers

much more in terms of educational accountability than the other dominant, wide-scaled educational accountability measures and policies the country has seen to date, and (2) as long as users proceed with caution, no real harm can come from the adoption and application of value-added systems and policies. Harris might pay a visit to the teachers in Houston to see for himself about that.

References

- Baeder, J. (2010, December 21). Gates' measures of effective teaching study: More value-added madness. *Education Week*. Retrieved from http://blogs.edweek.org/edweek/on_performance/2010/12/gates_measures_of_effective_teaching_study_more_value-added_madness.html
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J., F Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: Economic Policy Institute. Retrieved from <http://www.epi.org/publications/entry/bp278>
- Capitol Hill Briefing. (2011, September 14). Getting teacher evaluation right: A challenge for policy makers. A briefing by E. Haertel, J. Rothstein, A. Amrein-Beardsley, and L. Darling-Hammond. Washington DC: Dirksen Senate Office Building (research in brief). Retrieved from <http://www.aera.net/Default.aspx?id=12856>
- Haertel, E. (2011). *Using student test scores to distinguish good teachers from bad*. Paper presented at the Annual Conference of the American Educational Research Association (AERA), New Orleans, LA.
- Jacob, B. & Lefgren, L. (2007). Principals as agents: Subjective performance assessment in education. *Journal of Labor Economics*, 26(1), 101-136.
- Koedel, C., & Betts, J. R. (2007, April). *Re-examining the role of teacher quality in the educational production function*. Working Paper No. 2007-03. Nashville, TN: National Center on Performance Initiatives.

Papay, J. P. (2010). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal*, 48(1), 163-193. doi:10.3102/0002831210362589

About the Reviewers

Clarín Collins is a doctoral candidate in Educational Leadership and Policy Studies at Arizona State University. Her research interests include national and local policy implementation at the classroom level, teacher influences on policy making and implementation, and education evaluation and accountability systems. Her dissertation study analyzes teachers' understanding of and experiences with the Educational Value-Added Assessment System (EVAAS) in a large, urban, high-needs school district using the EVAAS to evaluate teachers with high-stakes consequences.



Audrey Amrein-Beardsley is currently an Associate Professor in the Mary Lou Fulton Teachers College at Arizona State University. Audrey's research interests include educational policy, research methods, and more specifically, high-stakes tests and value-added measurements and systems. In addition, she researches aspects of teacher quality and teacher education. She is also the creator and host of a show titled Inside the Academy during which she interviews some of the top educational researchers in the academy. For more information please see <http://insidetheacademy.asu.edu>.



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

Editors

Gene V Glass

glass@edrev.info

Gustavo Fischman

fischman@edrev.info

Melissa Cast-Brede

cast-brede@edrev.info
