



**Carnegie Foundation** for the Advancement of Teaching

# HOW DO VALUE-ADDED INDICATORS COMPARE TO OTHER MEASURES OF TEACHER EFFECTIVENESS?

**DOUGLAS N. HARRIS**

TULANE UNIVERSITY

CARNEGIE KNOWLEDGE NETWORK

**What We Know Series:**

Value-Added Methods and Applications

**ATIL**

ADVANCING • TEACHING • IMPROVING • LEARNING

# HOW DO VALUE-ADDED INDICATORS COMPARE TO OTHER MEASURES OF TEACHER EFFECTIVENESS?

## HIGHLIGHTS

- Value-added measures are positively related to almost all other commonly accepted measures of teacher performance such as principal evaluations and classroom observations.
- While policymakers should consider the validity and reliability of all their measures, we know more about value-added than others.
- The correlations appear fairly weak, but this is due primarily to lack of reliability in essentially all measures.
- The measures should yield different performance results because they are trying to measure different aspects of teaching, but they differ also because all have problems with validity and reliability.
- Using multiple measures can increase reliability; validity is also improved so long as the additional measures capture aspects of teaching we value.
- Once we have two or three performance measures, the costs of more measures for accountability may not be justified. But additional *formative* assessments of teachers may still be worthwhile to help these teachers improve.

## INTRODUCTION

In the recent drive to revamp teacher evaluation and accountability, measures of a teacher's value added have played the starring role. But the star of the show is not always the best actor, nor can the star succeed without a strong supporting cast. In assessing teacher performance, observations of classroom practice, portfolios of teachers' work, student learning objectives, and surveys of students are all possible additions to the mix.

All these measures vary in what aspect of teacher performance they measure. While teaching is broadly intended to help students live fulfilling lives, we must be more specific about the elements of performance that contribute to that goal – differentiating contributions to academic skills, for instance, from those that develop social skills. Once we have established what aspect of teaching we intend to capture, the measures differ in how valid and reliable they are in capturing that aspect.

Although there are big holes in what we know about how evaluation measures stack up on these two criteria, we can draw some important conclusions from the evidence collected so far. In this brief, we will show how existing research can help district and state leaders who are thinking about using multiple measures of teacher performance to guide them in hiring, development, and retention.

## WHAT DO WE KNOW ABOUT HOW ALTERNATIVE MEASURES OF TEACHER EFFECTIVENESS COMPARE?

The simplest way to judge how well various measures compare with each other is to calculate a “correlation,” a statistic that indicates the extent to which two numbers move in tandem. When two measures are unrelated to one another the correlation is 0.0. When two measures are perfectly correlated the correlation is 1.0.<sup>1</sup>

The Measures of Effective Teaching (MET) study, a much-cited study funded by the Gates Foundation,<sup>2</sup> finds correlations of 0.12 to 0.34 between value-added measures and classroom observation rubrics such as the Danielson Framework. The connection is nearly identical to the correlations that prior studies have found between value-added measures and confidential low-stakes evaluations of teachers by their principals.<sup>3</sup> MET found stronger relationships between value-added measures and student surveys of teacher practice.<sup>4</sup> Although students certainly are not expert judges of effective teaching, they are with teachers every day, and it is their performance on standardized tests that ultimately determines a teacher’s value added.

To put these numbers into perspective, I created two hypothetical performance measures and placed 100 teachers into one of four equally-sized performance categories, “A” being high and “D” being low. Correlations of 0.2, such as those above, are consistent with 32 teachers remaining in the same categories under both measures (e.g., performance level A in both cases), 44 teachers switching by one performance level, 18 switching by two levels (e.g., A to C), and 6 teachers switching from the top to the bottom or vice versa.

Another measure that is positively related to value-added measures is a teacher’s overall score from the National Board for Professional Teaching Standards (NBPTS).<sup>5</sup> Not surprisingly, the individual components of that score—videos of instruction, teacher portfolios, and standardized tests—are also positively related to value-added estimates. All of this evidence reinforces the general conclusion that almost all the other measures now being considered are positively correlated with value-added measures.<sup>6</sup>

How should we interpret these correlations? Are they strong enough to justify the use of value-added measures? At least one report has criticized teacher value-added measures because they do not line up closely enough with other measures.<sup>7</sup> Certainly, the level of agreement is not very high. Below, we consider why that is and what the implications are.

### Validity, reliability, and classification errors

When selecting performance measures, it is important to establish criteria for evaluating them and to apply the same criteria to all. When we talk about “accurate” measures of teacher performance, researchers mean measures that are valid and reliable. *Validity* refers to the degree to which something measures what it claims to measure. *Reliability* refers to the degree to which the measure is consistent when repeated. A measure could be valid on average, but inconsistent when repeated. Conversely, a measure could be highly reliable but invalid—that is, it could consistently provide the same invalid information.

## HOW DO VALUE-ADDED INDICATORS COMPARE TO OTHER MEASURES OF TEACHER EFFECTIVENESS?

Validity is closely related to the idea of *bias*. We tend to think of bias as arising in subjective decisions. Someone who works at the Ford Motor Company might say, “I love Ford cars, but I work there, so I’m biased,” which is to say that the person’s views might not correspond to the objective qualities of the car. Likewise, with teacher evaluation, if the observer knows the teacher personally, we might worry that she is hampered in her ability to objectively judge that teacher’s performance. Also, some school principals and other observers may simply not know what to look for when assessing classroom practice; this, too, can introduce bias.

Value-added measures can also be biased, but in a somewhat different way. A common criticism of value-added measures is that some teachers are at a disadvantage because they are assigned students who are more difficult to educate, even after the measures account for students’ prior test scores; this is what researchers call *selection bias*. No matter how many times we calculate value-added for these teachers, this form of bias means the results will still be invalid.

Serious consequences arise when measures are not valid and reliable: they increase *classification errors* – the placement of teachers into incorrect performance categories. In Florida, the District of Columbia, and a growing number of states and districts, performance measures, including value-added estimates, play a key role in placing teachers in performance categories. Being misclassified as unsatisfactory could mean losing a job.

### Evidence on validity and reliability

The evidence on the validity and reliability of value-added estimates is evolving and may be misunderstood.<sup>8</sup> Strictly speaking, establishing that a measure is valid requires comparing it to a “true” measure that we know to be correct, but this is essentially impossible with teaching. Instead, researchers try to establish validity indirectly. One widely publicized study, which created a statistical test of the validity of value-added measures,<sup>9</sup> found reason for concern. Several subsequent studies suggest that the measures are probably reasonably valid on average.<sup>10</sup> Another study randomly assigned students to teachers and found that selection bias was only a small problem, although there is debate about the interpretation of this result.<sup>11</sup>

It is important to note, however, that even if the conclusions from these studies are right, they provide evidence about whether value-added measures are valid *on the average* across large numbers of teachers.<sup>12</sup> They could still be—and apparently are—invalid for specific subgroups of teachers. For example, one study, which points out that almost all the evidence about validity is based on studies in elementary schools, provides evidence that typical value-added measures are biased in middle and high school.<sup>13</sup> Another study suggests that teachers whose students start off with very high achievement will receive lower performance ratings than they deserve because of the “test ceiling”.<sup>14</sup> Value-added measures are probably also highly sensitive to the context of teachers’ classrooms, including behavioral issues and the school culture.<sup>15</sup> The assumptions underlying value-added models have also been shown to be false,<sup>16</sup> which may influence different teachers in different ways. In general, a measure cannot be considered valid if it is heavily influenced by factors that are outside the control of teachers. So these findings raise important concerns.<sup>17</sup>

Reliability is also considered a significant issue with value-added measures. As with correlations, the highest possible reliability measure is 1.0, which means that the performance measure does not change at all over time.<sup>18</sup> At the other extreme, reliability of zero means that any given measure of

## HOW DO VALUE-ADDED INDICATORS COMPARE TO OTHER MEASURES OF TEACHER EFFECTIVENESS?

teacher performance tells us nothing about what the measure will be for the same teacher the next time.<sup>19</sup> When creating student tests, for example, designers usually set a standard of at least 0.9.<sup>20</sup> The MET study reports reliability for teacher value-added measures of about 0.3 to 0.5 when three years of data are used.<sup>21</sup> To put this in perspective, one study finds that only 28 to 50 percent of teachers who were ranked in the top fifth on value-added measures one year were still ranked in the top fifth in the subsequent year, and 4 to 15 percent of teachers switched from the top fifth to the bottom fifth.<sup>22</sup>

Critics of value-added measures might stop right here and argue that the limited reliability argues against using value added at all. But here again, we have to compare the alternatives. A single classroom observation has lower reliability than a value-added measure, but a combination of four classroom observations yields a higher reliability of about 0.65.<sup>23</sup> The validity of classroom observation measures is less clear. The MET study involved highly trained observers who had literally passed an exam demonstrating their skill; this level of training is unlikely in everyday school settings. Older evidence suggests that classroom observations can be influenced by factors unrelated to performance, such as age and race. Also, it seems likely that classroom context will affect observation measures just as it appears to affect value-added measures. For example, it may be difficult to make valid comparisons between the classroom management skills of a teacher who has emotionally impaired students, subject to frequent disruptions, to the skills of a teacher whose students are less disruptive.

On other evaluation measures, we have almost no evidence about validity and reliability. With the measure known as student learning objectives (SLOs), teachers work with their instructional leaders to identify student needs, create specific objectives, and establish metrics based on student work to establish progress toward those objectives. This process combines the outcomes orientation of student test scores with the more subjective elements of classroom observations. While these kinds of evaluations are somewhat similar to the teacher portfolios used in NBPTS, they are difficult to assess.

SLOs are potentially attractive because they can be used in all classrooms, allow local autonomy, and fit well with customized instruction. Some view the autonomy as a weakness because the measures are unlikely to be comparable across teachers and may be too easily manipulated to give the appearance of high performance. In any event, there is essentially no evidence about the validity or reliability of SLOs.

Most of the evidence cited above involves low-stakes measures, but the validity and reliability of all the measures is likely to be influenced by the stakes attached to them.<sup>24</sup> When a district uses value-added measures, some teachers might try to “game the system” and get assigned to students who are most likely to make achievement gains or to move to schools where the biases seem to give teachers an advantage. Alternatively, high stakes might cause classroom observers to take their task more seriously -- to be more careful in their work.

Measures that allow for more subjectivity and local control, such as classroom observations and SLOs, are subject to their own types of bias. If classroom observers have personal relationships with teachers they may, however unintentionally, help their friends. Likewise, if teachers can largely set their own objectives with SLOs, they may set the bar low to ensure that they appear successful.

## HOW DO VALUE-ADDED INDICATORS COMPARE TO OTHER MEASURES OF TEACHER EFFECTIVENESS?

### From validity and reliability to practicality

In addition to questions of validity and reliability, it is worth briefly considering the practicality and costs of the various evaluation measures. Value-added measures, in most districts, can only be used with about one-third of teachers who are in tested grades and subjects and who have at least two years of data. This necessitates some other approach with other teachers. On the other hand, value-added measures are fairly inexpensive once the testing regime is in place. Also, while some have criticized the complexity of value-added measures, at least one handbook for SLOs is nearly 60 pages long, and classroom observations can involve more than 100 sub-measures. A complete comparison of multiple measures requires that these practical considerations be accounted for as well.

### WHAT MORE NEEDS TO BE KNOWN ON THIS ISSUE?

We know much more about value-added measures than we do about other evaluation methods, so clearly we need more research on the latter, as well as additional research on value-added measures to determine how valid they are for particular groups of teachers. The limited evidence is a big problem; it means that even what we think we know about value added does not get us very far in deciding what to do with it. Without conducting similar analyses on the other measures, we can't compare alternatives and choose the best options.

More research on other measures would also help us understand why the correlations among them are so modest—why they differ as much as they do. The first reason they differ is simply that each measure really captures a different notion of teacher performance; the measures *should* yield different results. For example, we have every reason to believe that principals care about students' academic achievement more than anything else. In one study, principals' assessments of overall teacher performance and their assessment of teacher contributions to student achievement are correlated at about 0.7, very high.<sup>25</sup> However, principals rank a "caring" disposition as one of the most important teacher traits.<sup>26</sup> Clearly, a principal who cares mainly about academic achievement thinks about teacher performance differently than one who prefers a caring personality.

If each measure were valid and reliable, the correlations would no doubt be much higher. But even then, the correlations would still be less than 1.0 for two reasons. First, one measure might be less valid than another, even when the intended notion of teacher performance is the same. Second, the maximum correlation is roughly equal to the reliability of the two measures, which is generally much less than 1.0.<sup>27</sup> For example, two measures with reliabilities of 0.5 (which seems realistic given the above measures) have a maximum correlation of 0.5.

These examples are largely hypothetical because we lack evidence on the validity and reliability of measures other than value-added. But the evidence does suggest that the main reason the measures differ is that each measure is unreliable. This is an important lesson because there are steps that can be taken to increase reliability, such as increasing the number of classroom observations and years of data used in value-added calculations.

## HOW DO VALUE-ADDED INDICATORS COMPARE TO OTHER MEASURES OF TEACHER EFFECTIVENESS?

### WHAT CAN'T BE RESOLVED BY EMPIRICAL EVIDENCE ON THIS ISSUE?

While choices about the mix of measures should be made partly based on evidence, they also require value judgments. We have to decide first what aspects of teaching we value. Are we more concerned about students obtaining academic skills or social skills or creativity? Choosing the right mix of measures therefore depends on what we think school should be trying to achieve. A valid measure of teacher performance is one designed to capture how well teachers contribute to the student outcomes we value most. On this, there are legitimate differences of opinion.

### PRACTICAL IMPLICATIONS

#### Does this issue impact district decision making?

There is wide support for using more measures in addition to value added to make high-stakes decisions. If a multiple-measures approach helps create a composite that is more representative of what stakeholders value, then validity improves. Using multiple measures also improves reliability, up to about 0.65 in the MET study.<sup>28</sup> This figure is still below conventional levels in educational assessment, but it is better than the alternative of a single measure.

But we can go further in thinking about how many, and which, measures should be used. Basic economic theory provides a useful perspective on multiple measures. First, economics recognizes that quality teacher evaluation is expensive and time-consuming. To observe teachers in class, for example, principals must take time away from other duties, and some of the best teachers must be pulled from the classroom to evaluate others. What matters is not just how many measures are used, but how much information is collected with each.

Second, basic economics suggests that when two measures are highly correlated, there is not much point in using both of them. This issue might seem moot since none of the measures *are* highly correlated, but the same principle applies. It's just that we have to interpret "highly correlated" based on the maximum correlation possible.

Yet many states and districts are considering using three or more measures. So the question then becomes: how much additional information would a third or fourth measure bring? The answer depends on the reliability of the additional measure, as well as how the random error in the additional measure correlates with the random errors in the other measures. In general, additional measures will increase both validity and reliability, but at some point the additional gain is not worth the cost. Austin, Texas schools use 13 measures to evaluate teachers – a costly strategy that may also confuse teachers about what they are supposed to be aiming for. States and districts can test the worth of adding more measures by calculating the correlations between simpler and more complex composite measures. If the correlations are very high, it might indicate that the additional measures are not worthwhile.

The economics-based approach, however, focuses on so-called *summative* performance measures that evaluators use to make high-stakes decisions about teachers' salaries and careers. Organizations also need *formative* information to help teachers improve; they need indicators of a teacher's specific skills in classroom management, for instance, or her ability to provide meaningful feedback to

## HOW DO VALUE-ADDED INDICATORS COMPARE TO OTHER MEASURES OF TEACHER EFFECTIVENESS?

students. Both types of measures are important.<sup>29</sup> So, even if an additional measure gives evaluators little in the way of new summative information, it may be quite valuable for the formative information it provides.

Performance measures are the lynchpin of teacher evaluation systems. The choice of measures is therefore the crucial first decision for administrators developing any system of teacher improvement and accountability. We have learned a great deal about the strengths and weaknesses of one of those measures – value added – but we need to know much more about the others. After all, we can't decide how best to use value-added measures without determining how the other measures compare. So far, the modest correlations we see imply that different evaluation measures will yield different results for the same teacher. We can reduce these classification errors by using multiple measures to improve validity and reliability, and by creating additional checks and balances when making high-stakes decisions. We can never eliminate classification errors, but we can reduce them.

### ENDNOTES

---

<sup>1</sup> Measures can also move in opposite direction, yielding correlations as low as -1.0.

<sup>2</sup> Bill and Melinda Gates Foundation, *Gathering Feedback for Teaching* (2012).

<sup>3</sup> Douglas N. Harris and Tim R. Sass, "What makes for a good teacher and who can tell?" (Working Paper no. 30, National Center for the Analysis of Longitudinal Data in Education Research (CALDER), Urban Institute, Washington, DC, 2009.), Brian Jacob and Lars Lefgren, "Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education," *Journal of Labor Economics*, 26, (2008): 101.

<sup>4</sup> The MET study does not report correlations between student surveys and the classroom observations. Gates Foundation, 2012, *ibid*.

<sup>5</sup> There is an important distinction between whether a teacher is NBPTS-certified and the overall score. NBPTS certification is at best weakly related to teacher value-added. The overall score used to make that determination is, however, related to teacher value-added. These apparently contradictory findings are possible because the overall score presents a nuanced picture of teacher performance where certification is a crude yes-no determination.

<sup>6</sup> Anytime I report that there is a relationship between two measures I mean that the results are also "statistically significant" by the usual statistical standards.

<sup>7</sup> Heather Hill, *Journal of Policy Analysis and Management*, 28, (2009): 700.

<sup>8</sup> There are many summaries of value-added research: Steven Glazerman et al., "Evaluating Teachers: The Important Role of Value-Added," (Brookings Institution, Washington, DC, 2010.), Douglas N. Harris, *Value-Added Measures in Education: What Every Educator Needs to Know* (Harvard Education Press, Cambridge, MA, 2011a); Douglas N. Harris, "Value-added measures and the future of educational accountability," *Science*, 333 (2011b): 826-827. Jonah E. Rockoff and Douglas O. Staiger, "Searching for Effective Teachers with Imperfect Information," *Journal of Economic Perspectives*, 24, no.3, (2010): 97–118.

<sup>9</sup> Jesse Rothstein, "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement," *Quarterly Journal of Economics* 125, no.1, (2010): 175-214.

<sup>10</sup> Dan Goldhaber and Duncan Chaplin, "Assessing the "Rothstein Falsification Test." Does it Really Show Teacher Value-added Models are Biased?" (CEDR Working Paper, University of Washington, Seattle, WA, 2012-13); Raj Chetty, John N. Friedman, and Jonah E. Rockoff, "The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood," (Working Paper 17699, National Bureau of Economic Research. <http://www.nber.org/papers/w17699>); Cassandra M. Guarino, Mark D. Reckase, and Jeffrey M. Wooldridge, (Paper presented at the annual meeting of the Association for Education Finance and Policy, Seattle, WA, 24- 26 March 2011); Cory Koedel and Julian Betts, "Does Student Sorting Invalidate Value-Added Models of Teacher Effectiveness? An Extended Analysis of the Rothstein Critique," Unpublished manuscript, 2009.



## HOW DO VALUE-ADDED INDICATORS COMPARE TO OTHER MEASURES OF TEACHER EFFECTIVENESS?

---

- <sup>11</sup> Thomas Kane and Douglas Staiger, "Estimating teacher impacts on student achievement: An experimental evaluation," (National Bureau of Economic Research, Boston, MA, Working Paper, 2008). The study results are a subject of contention partly because it is not clear whether the schools were already essentially randomly assigned students to teachers. If they were then the study is not informative about whether value-added models solve selection bias.
- <sup>12</sup> Douglas N. Harris, *ibid.*, 2011a.
- <sup>13</sup> Douglas N. Harris and Andrew Anderson, Bias of Public Sector Worker Performance Monitoring: Theory and Empirical Evidence from Middle School Teachers, (Paper presented at the Association for Policy Analysis and Management 2012.)
- <sup>14</sup> Cory Koedel and Julian Betts, "Value-added to what? How a ceiling in the testing instrument influences value-added estimation," (NBER Working Paper 14778, National Bureau of Economic Research, 2009). Available at: <http://www.nber.org/papers/w14778>.
- <sup>15</sup> Sean F. Reardon, Stephen W. Raudenbush, Assumptions of value-added models for estimating school effects, *Education Finance and Policy*, 4, no.4 (2009): 492-519.
- <sup>16</sup> Douglas N. Harris, Tim R. Sass, and Anastasia Semykina, (Paper presented at the annual meeting of the Association for Education Finance and Policy, Seattle, WA, 24-26 March, 2011.)
- <sup>17</sup> Harris, 2011a, *ibid.*
- <sup>18</sup> Somewhat more formally, the reliability coefficient is the variation in that portion of teacher performance that is consistent over time (the portion not subject to random error) divided by the overall variation in the performance measure, which includes both the consistent and inconsistent (error) portions.
- <sup>19</sup> Unlike correlations, reliability measures cannot be negative.
- <sup>20</sup> It is important to note that this captures only "test-retest" reliability, referring to the consistency over time in the scores of individual students who take the same test twice.
- <sup>21</sup> Gates Foundation, 2012, *ibid.*
- <sup>22</sup> Daniel F. McCaffrey, Tim R. Sass, J. R. Lockwood, Kata Mihaly, "The Intertemporal Variability of Teacher Effect Estimate." *Education Finance and Policy*, 4, (2009):572.
- <sup>23</sup> Gates Foundation, 2012.
- <sup>24</sup> This is sometimes referred to as Campbell's Law.
- <sup>25</sup> Douglas N. Harris and Tim R. Sass, 2007, *ibid.*
- <sup>26</sup> Douglas N. Harris, Stacey Rutledge, William Ingle, and Cynthia Thompson, "Mix and match: What principals really look for when hiring teachers," *Education Finance and Policy*, 5, no.2 (2010): 228-246.
- <sup>27</sup> Specifically, the maximum correlation equals the square root of the product of the two reliabilities. So, if the two reliability coefficients are the same then this amounts to squaring and then taking the square root, which cancel out, to yield the reliability coefficient.
- <sup>28</sup> This is likely to be a true maximum in the sense that these are all low-stakes measures and the raters in the study were carefully trained. When stakes are attached the reliabilities are likely to diminish.
- <sup>29</sup> There is also some debate about whether more formative measures should be counted in the high-stakes composite measure; some believe formative assessments should always be low stakes. The case for including them in high-stakes decisions is that they can help offset the errors of value-added measures with regard to the quality of instructional practice.

# HOW DO VALUE-ADDED INDICATORS COMPARE TO OTHER MEASURES OF TEACHER EFFECTIVENESS?

## AUTHORS



**Douglas N. Harris** is Associate Professor of Economics and University Endowed Chair in Public Education at Tulane University. His research explores how students' educational outcomes are influenced by school choice, standards, teacher evaluation, test-based accountability, college financial aid, and college access programs. A former school board member, his research marries theory and rigorous research with the practical realities of schooling with publications ranging from the general interest journal *Science* to the *Journal of Public Economics*, *Journal of Policy Analysis and Management*, and others. His recent book, *Value-Added Measures in Education* (Harvard Education Press, 2011) was nominated for the national Grawemeyer Award. *Washington Monthly* magazine has used his research on college performance measures in its college ratings and David Brooks has cited related work in his *New York Times* column. He has advised eight state departments of education, elected officials at all levels of government, and groups such as the National Academy of Sciences, National Council of State Legislatures, National Governors Association, and National School Boards Association. His work is frequently cited in the national media, including CNN, *Education Week*, *The New York Times*, and *The Washington Post*.

## ABOUT THE CARNEGIE KNOWLEDGE NETWORK

The Carnegie Foundation for the Advancement of Teaching has launched the Carnegie Knowledge Network, a resource that will provide impartial, authoritative, relevant, digestible, and current syntheses of the technical literature on value-added for K-12 teacher evaluation system designers. The Carnegie Knowledge Network integrates both technical knowledge and operational knowledge of teacher evaluation systems. The Foundation has brought together a distinguished group of researchers to form the *Carnegie Panel on Assessing Teaching to Improve Learning* to identify what is and is not known on the critical technical issues involved in measuring teaching effectiveness. Daniel Goldhaber, Douglas Harris, Susanna Loeb, Daniel McCaffrey, and Stephen Raudenbush have been selected to join the Carnegie Panel based on their demonstrated technical expertise in this area, their thoughtful stance toward the use of value-added methodologies, and their impartiality toward particular modeling strategies. The Carnegie Panel engaged a User Panel composed of K-12 field leaders directly involved in developing and implementing teacher evaluation systems, to assure relevance to their needs and accessibility for their use. This is the first set of knowledge briefs in a series of Carnegie Knowledge Network releases. Learn more at [carnegieknowledgenetwork.org](http://carnegieknowledgenetwork.org).

## CITATION

Harris, Douglas N. Carnegie Knowledge Network, "How Do Value-Added Indicators Compare to Other Measures of Teacher Effectiveness?" Last modified May 2013.

URL = <<http://carnegieknowledgenetwork.org/briefs/value-added/value-added-other-measures/>>



## **Carnegie Foundation** for the Advancement of Teaching

Carnegie Foundation for the Advancement of Teaching  
51 Vista Lane  
Stanford, California 94305  
650-566-5100

Carnegie Foundation for the Advancement of Teaching seeks to vitalize more productive research and development in education. We bring scholars, practitioners, innovators, designers, and developers together to solve practical problems of schooling that diminish our nation's ability to educate all students well. We are committed to developing networks of ideas, expertise, and action aimed at improving teaching and learning and strengthening the institutions in which this occurs. Our core belief is that much more can be accomplished together than even the best of us can accomplish alone.

**[www.carnegiefoundation.org](http://www.carnegiefoundation.org)**

We invite you to explore our website, where you will find resources relevant to our programs and publications as well as current information about our Board of Directors, funders, and staff.



This work is licensed under the Creative Commons Attribution-Noncommercial 3.0 Unported license.

To view the details of this license, go to:  
<http://creativecommons.org/licenses/by-nc/3.0/>.

**Knowledge Brief 5**  
**October 2012, Revised May 2013**  
**[carnegieknowledgenetwork.org](http://carnegieknowledgenetwork.org)**

*Funded through a cooperative agreement with the Institute for Education Science. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.*