Carnegie Foundation for the Advancement of Teaching

HOW STABLE ARE VALUE-ADDED ESTIMATES ACROSS YEARS, SUBJECTS AND STUDENT GROUPS?

SUSANNA LOEB
AND CHRISTOPHER A. CANDELARIA
STANFORD UNIVERSITY

CARNEGIE KNOWLEDGE NETWORK
What We Know Series:
Value-Added Methods and Applications

ATIL ASSESSING • TEACHING • IMPROVING • LEARNING

HIGHLIGHTS

- A teacher's value-added score in one year is partially but not fully predictive of her performance in the next.
- Value-added is unstable because true teacher performance varies and because value-added measures are subject to error.
- Two years of data do a meaningfully better job at predicting value added than just one.
- A teacher's value added in one subject is only partially predictive of her value added in another, and a teacher's value added for one group of students is only partially predictive of her valued added for others.
- The variation of a teacher's value added across time, subject, and student population depends in part on the model with which it is measured and the source of the data that is used.
- Year-to-year instability suggests caution when using value-added measures to make decisions for which there are no mechanisms for re-evaluation and no other sources of information.

INTRODUCTION

Value-added models measure teacher performance by the test score gains of their students, adjusted for a variety of factors such as the performance of students when they enter the class. The measures are based on desired student outcomes such as math and reading scores, but they have a number of potential drawbacks. One of them is the inconsistency in estimates for the same teacher when value added is measured in a different year, or for different subjects, or for different groups of students.

Some of the differences in value added from year to year result from true differences in a teacher's performance. Differences can also arise from classroom peer effects; the students themselves contribute to the quality of classroom life, and this contribution changes from year to year. Other differences come from the tests on which the value-added measures are based; because test scores are not perfectly accurate measures of student knowledge, it follows that they are not perfectly accurate gauges of teacher performance.

In this brief, we describe how value-added measures for individual teachers vary across time, subject, and student populations. We discuss how additional research could help educators use these measures more effectively, and we pose new questions, the answers to which depend not on empirical investigation but on human judgment. Finally, we consider how the current body of knowledge, and the gaps in that knowledge, can guide decisions about how to use value-added measures in evaluations of teacher effectiveness.

WHAT IS THE CURRENT STATE OF KNOWLEDGE ON THIS ISSUE?

Stability across time

Is a teacher with a high value-added estimate in one year likely to have a high estimate the next? If some teachers are better than others at helping students improve their performance on standardized tests, we would expect the answer to be "yes." But we also know that teachers are better at their jobs in some years than in others, so we would not expect their value added from one year to the next to be exactly the same. That is, part of a teacher's influence on students persists over time, and part of it varies. It may vary for a range of reasons. Professional development might lead to improvement, or family responsibilities or school reforms might cause distractions in one year but not in another.

Instability is not necessarily a bad thing; if a school is engaging in substantial change, one might expect teacher performance to change as well. In addition, value-added measures are subject to error because of a number of factors, such as whether the students have good or bad days when taking the test. The smaller the number of test scores used to create value-added measures, the more important students' idiosyncratic performance is likely to be. So, we can think of value-added estimates as measuring three components: (1) true teaching effectiveness that persists across years; (2) true effectiveness that varies from year to year; and (3) measurement error. The relationship between a teacher's value added in one year and her value added in another year combines all three of these factors.

One statistic that describes the year-to-year persistence in value-added estimates is the correlation coefficient, a measure of the linear relationship between two variables. The correlation coefficient is 1.0 if a teacher's value-added measure in one year is perfectly predictive of her score in the following year and zero if her score one year tells us nothing about how she will fare the next. A study of elementary school teachers found correlations from one year to the next of 0.6 to 0.8 for math and 0.5 to 0.7 for reading.[1] Other studies have found lower correlations, ranging from 0.2 to 0.6, depending on location, school level, and the specifications of the value-added model.[2] These estimates are similar to the year-to-year performance correlations found for other occupations, including college professors, athletes, and workers in insurance sales.[3] Overall, a teacher's value-added estimate in one year provides information about his or her value added in the following year, but it is not perfectly predictive.

While one year of data does only a modest job of predicting teachers' future value added, averaging the value-added measures across more than one year makes predictions more accurate. Using the same data and model specification, one study found a correlation of 0.4 comparing one year of data to the next and a correlation of 0.6 using two years of data.[4] In other words, in this study, two years of data does a meaningfully better job at predicting value added than does just one. It is not as clear from the literature whether more than two years of data is helpful. The one available study found that the correlation of two years of value-added data to future value-added data was significantly higher than one year, but that additional years of data (up to six) added only a little more predictive power.[5]

Thought of in another way, teachers who have a high value-added estimate in one year are likely to have a high value-added estimate in the next, but some teachers with high value-added scores in one year have quite low scores in the next. Table 1 presents a transition matrix that displays the value-

added rankings of one sample of teachers over two years.[6] Teachers are ranked and grouped into quintiles based on their value-added score in each year. Half of teachers in the top fifth in the first year remained there, while 8 percent dropped all the way to the bottom fifth. We also see that 43 percent of teachers who had value-added estimates in the bottom quintile in one year remained there the next. If there were no relationship between years, we would expect only 20 percent of teachers to be in this group since they would be evenly divided between the five quintiles.[7] The table shows some stability in the value-added measures as most teachers either remain in the same quintile from one year to the next or move just one quintile up or down.

Table 1: Stability of Teacher Value-Added Across Time (Percent of Teachers by Row)

	Ranking in year 2 (Quintiles)				
Ranking in year 1 (Quintiles)	Q1 (Bottom)	Q2	Q3	Q4	Q5 (Top)
Q1 (Bottom)	43%	29%	14%	10%	4%
Q2	26%	21%	25%	18%	9%
Q3	12%	21%	28%	25%	15%
Q4	10%	19%	19%	28%	23%
Q5 (Top)	8%	11%	11%	19%	50%

Notes: Total teachers = 941. Source: Koedel & Betts (2007)

Stability across subjects

Teachers may teach multiple subjects over the course of a year, and they might be better at teaching one subject than another. For example, in elementary school, some teachers might be more effective teaching multiplication than reading. If a teacher who is very good at teaching one topic is also very good at teaching another—if her value-added measures are similar across topics—we might be comfortable using value-added measures based on a subset of student outcomes, for example, just math or just reading. However, we might be less comfortable using a single measure of student outcomes if the value-added estimates varied a lot across topics. If it did, we would want to measure teacher effectiveness in each subject that we care about.

To assess the consistency of teacher effectiveness across subjects, we can compare a teacher's value added in one subject to her value added in another. The consistency of these measures across subjects, similar to consistency over time, combines both true differences in effectiveness and measurement error. That is, one reason a teacher's value-added estimates are not perfectly correlated across subjects is that both the estimates are measured with error. Another reason is true differences in contributions to student learning.

Studies that have looked at teachers' consistency across reading and math have found correlation coefficients in the range of 0.2 to 0.6—similar to the correlations for the year-to-year comparisons.[8] To help make these findings clearer, we created a matrix based on the data from one of these studies.[9] The matrix in Table 2 shows rankings of teachers on both their reading and math value-added scores. We see that 46 percent of teachers who ranked in the bottom fifth in math also ranked in the bottom fifth for reading. Similarly, 46 percent of teachers who ranked in the top fifth for math also ranked in the top fifth for reading (compared with only four percent in the bottom fifth). The

matrix suggests that, on average, teachers who are effective in math are also effective in reading, although some teachers are better at one than the other.

Teachers may differ in their value added not only *across* subject areas but *within* subject areas. Many standardized exams are composed of tests within tests. For example, the math section of the Stanford Achievement Test contains one subtest on procedures and one on problem-solving. Two studies have examined the consistency of value-added estimates based on these subtests, and one found correlations within the range of 0.5 to 0.7, while the other found correlations within the range of 0.0 to 0.4.[10] We have seen that the range in correlations across years and between subjects is due largely to differences in the methods for estimating value added. But in the second study, which looks *within* subjects, the lower correlations come from differences in data sources and differences across the subtests of the Stanford Achievement Test.[11]

Table 2: Stability of Teacher Value-Added Across Subjects (Percent of Teachers by Row)

	Ranking in reading (Quintiles)					
Ranking in math (Quintiles)	Q1 (Bottom)	Q2	Q3	Q4	Q5 (Top)	*Total*
Q1 (Bottom)	46.4%	27.3%	14.1%	8.7%	3.6%	704
Q2	23.2%	28.0%	22.3%	17.4%	9.2%	622
Q3	16.4%	24.5%	22.8%	20.9%	15.5%	580
Q4	8.3%	17.5%	24.7%	28.1%	21.4%	612
Q5 (Top)	4.1%	9.5%	16.1%	24.8%	45.5%	589
Total	641	671	616	608	571	3,107

Source: Loeb, Kalogrides, & Béteille (forthcoming).

Stability across student populations

The final type of consistency that we address is that which crosses student populations. Despite the great variation among student groups – demographic, academic, and otherwise – few studies have examined the extent to which teachers are more effective with one group of students than with another. Understanding these differences helps tell us whether a teacher who is effective with one group is likely to be effective with another. It also tells us how important it is to calculate value-added measures by student sub-group.

Two studies have compared teachers' value added with high-achieving students to the same teachers' value added with lower-achieving students. Both studies find that teachers who are better with one group are, in most cases, also better with the other group. The first study calculates two estimates: one based on the test scores of the teachers' high-performing students and the other based on the scores of their low-performing students.[12] While the samples used in the analyses are small, the authors find a correlation between the estimates of 0.4. As with the cases discussed above, the differences could come from variations in teachers' true value-added across student groups or from measurement error enhanced by the small sample size. The second study uses a different approach to ask a similar question. It finds that approximately 10 percent of the variance in teachers' value-added estimates can be attributed to the interaction between teachers and individual students, indicating substantial similarity in teacher effects between groups.[13]

Another study conducted a similar comparison of teachers' value added with English language learners and English-proficient students.[14] The authors found a correlation between teachers' value-added measures with the two groups of 0.4 to 0.6 for math and 0.3 to 0.4 for reading. For comparison, and to distinguish measurement error from true differences in teacher effectiveness, the authors ran similar correlations with randomly separated groups of students. They found correlations of 0.7 for math and 0.6 for reading -- evidence that teachers' value added for English learners is similar to their value added for students who already speak English, though again, they found that some teachers are somewhat better with one group than with the other.

Table 3 provides a matrix similar to those described above. It shows that 50 percent of teachers in the bottom fifth in value added for math with English-proficient students are also in the bottom fifth with English learners. That compares to less than four percent of teachers who are in the top fifth with English learners.[15] Similarly, 59 percent of the teachers in the top fifth in value added for math with English-proficient students are also in the top fifth in valued added with English learners; less than four percent are in the bottom fifth. Thus, the relationship between teachers' value-added measures for these student groups is at least as strong as the relationship between value-added measures across years and subject areas.

Table 3: Stability of Teacher Value-Added Between English Learners and Others (Math)

Ranking for English Learners (Quintiles)	Ranking for Non-English Learners (Quintiles)					
	Q1 (Bottom)	Q2	Q3	Q4	Q5 (Top)	*Total*
Q1 (Bottom)	49.90%	25.10%	14.80%	7.10%	3.70%	467
Q2	22.70%	32.40%	22.90%	14.10%	5.40%	480
Q3	15.10%	23.10%	27.50%	20.60%	11.80%	484
Q4	8.60%	15.10%	24.60%	30.20%	19.90%	480
Q5 (Top)	3.70%	4.40%	10.20%	28.00%	59.20%	475
Total	409	525	541	504	407	2,386

Source: Loeb, Soland, and Fox (2012)

WHAT MORE NEEDS TO BE KNOWN ON THIS ISSUE?

We have summarized the research on how well value-added measures hold up across years, subject areas, and student populations, but the evidence is based on a relatively small number of studies. Future research could better separate measurement error from true differences; more systematically compare estimates across model specifications; identify clear dimensions of time, topic, and student populations; and provide evidence on the sources of instability.

Again, we see changes in teachers' value-added estimates partly because the measurements are imprecise. While numerous papers have highlighted this imprecision, most studies of instability have not systematically considered the role of measurement error in estimates aside from the type that is caused by sampling error. Measurement error can come from many sources including errors in the

initial tests; the fact that many value-added estimates are based on small samples of students; errors in the data systems such as incorrect student teacher matches; effects of differential summer learning; and multiple support teachers working with same students

We also recommend further research on the differences that exist when value added is measured across datasets and model specifications.[16] With further study, researchers could clarify the role played by the model specifications and datasets in instability, and they could compare stability across a set of specifications. Then they could repeat the analysis across a range of datasets.

Our understanding of value-added consistency could also benefit from the examination of more subject areas and more student groups. We compared math to reading, but we could also consider outcomes in writing, science, social studies, and other topics. Across student groups, we could consider achievement levels and behavioral tendencies, as well as background characteristics such as gender, race/ethnicity, and family income.

Finally, further research could help uncover the causes of the inconsistency in value-added measures. To what degree are year-to-year changes due to differences in teacher effectiveness that schools can actually control? To what extent do factors outside the classroom affect teacher performance? To what extent are teachers handicapped simply because they aren't aware of the resources available to help students with different needs? When schools understand the causes of inconsistency in value-added measures, they not only learn about the appropriate use of these measures, they also learn how they can best help their teachers improve.

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

It is clear from the research so far that value-added measures will never be completely stable across time, topics, or student groups, nor would we necessarily want them to be because true teacher effectiveness likely varies across these dimensions. Given these general findings, there are a number of questions that do not depend on empirical evidence. They include choices about the number of years of data to combine, the subject areas or topics to measure, and the student groups among which to distinguish.

As described above, there are substantial yearly differences in teacher value-added scores. Districts can learn more about future value added by looking at more years of initial value added.[17] However, while using additional years leads to more stability, it also restricts the sample of students for whom these measures are available because teachers have to teach for multiple years to have multiple years of data. Deciding what level of instability is acceptable is a decision for practitioners, not researchers.

While instability across subject areas is perhaps not as great as instability across years, it is still substantial. It matters a lot which tests form the basis of the value-added measures. Do they measure math or reading? Do they measure factual knowledge or reasoning skills? Researchers can help develop more reliable tests, but the choice of subject areas and the tradeoffs across types of measures are matters of judgment. Similarly, considering groups of students separately increases instability, but if there are groups in which a district has a particular interest, it might be worth the cost in stability for the benefit of targeting this group.

More generally, while research can evaluate the effectiveness of policies that use value-added measures, research alone can never determine the optimal approach for a given district or school.

PRACTICAL IMPLICATIONS

How does this issue impact district decision making?

Year-to-year instability calls for caution when using value-added measures to make decisions for which there are no mechanisms for re-evaluation and no other sources of information. Similarly, the gain in stability from using multiple years of data points to benefits of using multiple years of data (at least two). However, even with multiple years of data, the instability can be consequential, a problem that points to supplementing value-added measures with other information

The instability across subjects also has important implications. First, it matters what is on the test. It should cover the skills and knowledge that are valued. Topic inconsistency can also have implications for teacher assignment and development. Also, knowing how teachers perform on different outcome measures, educational leaders can more carefully target professional development.

Finally, the consistency across student groups can affect practical decisions. By knowing how teachers perform on different outcome measures, education leaders can better target professional development to teachers' needs and can better target teachers to classes or subject areas for which they are effective. If teacher effectiveness varies a lot across subjects, leaders might make more careful distinctions among teaching jobs so teachers are teaching the subjects at which they are best. Moreover, if researchers can identify, in particular, why some teachers are more effective with underserved populations, school reform efforts might greatly benefit.

In summary, value-added measures vary for the same teacher from year to year, subject to subject, and student group to student group. Although the similarity across measures suggests they are useful tools for decision-making, the variation suggests the benefit of being able to update decisions when new information becomes available; if a teacher appears relatively ineffective in one year, it is worth seeing her measure in another year as well. Instability in value-added measures argues for using them in relatively low-stakes decisions that carry opportunities for reconsideration.

ENDNOTES

[1] The correlations are adjusted for sampling error, which is one part of the measurement error, see: Dan Goldhaber and Michael Hansen, "Is it just a bad class? Assessing the stability of measured teacher performance," (Paper for Center for Education Data and Research, 2010).

[2] For examples of studies with correlations in this range see: Daniel F. McCaffrey, Tim R. Sass, J. R. Lockwood, and Kata Mihaly, "The intertemporal variability of teacher effect estimates," *Education Finance and Policy*, 4 (2009): 572-606. doi: 10.1162/edfp.2009.4.4.572. Measurement of Effective Teaching Project, *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains*. Bill and Melinda Gates Foundation, Seattle (2012). We report ranges because different specifications of value-added models yield different

correlations. Models of student achievement in a given year as a function of prior achievement and other controls tend to give higher correlations than other models, see:
Daniel F. McCaffrey, Tim R. Sass, J. R. Lockwood, and Kata Mihaly, "The intertemporal variability of teacher effect estimates," *Education Finance and Policy*, 4 (2009): 572-606. doi: 10.1162/edfp.2009.4.4.572.

[3] McCaffrey et al. (2009).

[4] McCaffrey et al. (2009), show that correlations tend to be higher when using multiple years of data to estimate value-added models. The correlations generally range from 0.3 to 0.8. Using multiple years of data leads to a reduction in sampling error, see: Cory Koedel and Julian Betts, "Does Student Sorting Invalidate Value-Added Models of Teacher Effectiveness? An Extended Analysis of the Rothstein Critique," *Education Finance and Policy*, 6, no.1, (2011): 18-42. doi: 10.1162/EDFP_a_00027.

[5] Goldhaber and Hansen, 2010.

[6] The transition matrix comes from Table 8 of Koedel and Betts (2007): Cory Koedel and Julian Betts, "Re-examining the role of teacher quality in the educational production function," (Working Paper, University of Missouri, 2007). Retrieved from http://economics.missouri.edu/working-papers/2007/wp0708_koedel.pdf

[7] These results are from a value-added model that does not include school or student fixed effects. Those that do include these additional controls show lower persistence both in the transition tables and in the correlation coefficients The degree of instability shown in Table 1 is not unique to the Koedel & Betts (2007) study; Goldhaber and Hansen 2010, and McCaffrey et al. 2009, also construct transition matrices that have qualitatively similar patterns. For similar results from data collected on 9th grade high school teachers in Chicago, see: Daniel Aaronson, Lisa Barrow and William Sander, "Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25, no.1, (2007): 95-135.

[8] Koedel and Betts 2007, estimate that correlation coefficients between teacher value-added estimates derived from math and reading are in the range of 0.35 to 0.64. Similarly, Loeb, Kalogrides, and Béteille (forthcoming) find that the correlation between math and reading is in the range of 0.21 to 0.58 depending on the model specification: Tara Beteille, Demetra Kalogrides, and Susanna Loeb, "Effective schools: Teacher hiring, assignment, development, and retention," *Education Finance and Policy*. (Forthcoming).
Goldhaber, Cowen and Walch (2012) find correlations of 0.47-0.54.
Dan Goldhaber, James Cowan, and Joe Walch, "Is a Good Elementary Teacher Always Good? Assessing Teacher Performance Estimates Across Contexts and Comparison Groups," (Working Paper, 2012).
All adjust for measurement error.

[9] Specifically, the data come from Loeb, Kalogrides, and Béteille (forthcoming).

[10] John P. Papay, "Different tests, different answers: The stability of teacher value-added estimates across outcome measures," *American Educational Research Journal, 48,* no.1, (2011): 163-193. doi: 10.3102/0002831210362589.
J. R. Lockwood, Daniel F. McCaffrey, Laura S. Hamilton, Brian M. Stecher, Vi-Nhuan Le and Jose Felipe Martinez, "The sensitivity of value-added teacher effect estimates to different mathematics achievement measures," *Journal of Educational Measurement, 44,* no.1, (2007): 47-67.

[11] While we do not know of any study that has tested differences across different domains of English language arts, Papay (2011) compares teacher value added estimates from three complete reading tests: a state test, the SAT, and the Scholastic Reading Inventory (SRI), all purportedly testing the same content domains. Correlations among the value-added estimates derived from the three reading tests range from 0.2 to 0.6. McCaffrey et. al. (2009) also looks at stability across test instruments.

[12] See: Daniel Aaronson , Lisa Barrow, William Sander (2007). High-ability students are those that score in the top half of the distribution on the Test of Achievement and Proficiency (TAP), administered by the Chicago Public Schools. Low-ability students are those that score in the bottom half of the distribution. In order to be included in this analysis, the teacher had to have at least 15 students who scored in the top half of the distribution and at least 15 students who scored in the bottom half.

[13] J.R. Lockwood and Daniel F. McCaffrey, "Exploring student-teacher interactions in longitudinal achievement data," *Education Finance and Policy, 4,* no.4 (2009): 439-467.

[14] Susanna Loeb, Jim Soland and Lindsay Fox, "Is a Good Teacher a Good Teacher for All? Comparing Value-Added of Teachers with Their English Learners and Non-English Learners," Working Paper (in progress).

[15] Table 4 is based on a value-added specification that includes student, classroom and school controls but not student or school fixed effects.

[16] See: Goldhaber, Dan, and Roddy Theobald. Carnegie Knowledge Network, "Do Different Value-Added Models Tell Us the Same Things?" Last modified October 2012. URL = <http://carnegieknowledgenetwork.org/briefs/value-added/different-growth-models/>, and McCaffrey et. al. 2009.

[17] See: Koedel and Betts 2011.

AUTHORS

Susanna Loeb is the Barnett Family Professor of Education at Stanford University, Faculty Director of the Center for Education Policy Analysis, and a Co-Director of Policy Analysis for California Education (PACE). She specializes in the economics of education and the relationship between schools and federal, state and local policies. Her research addresses teacher policy, looking specifically at how teachers' preferences affect the distribution of teaching quality across schools, how pre-service coursework requirements affect the quality of teacher candidates, and how reforms affect teachers' career decisions. She also studies school leadership and school finance, for example looking at how the structure of state finance systems affects the level and distribution of resources across schools. Susanna is a senior fellow at the Stanford Institute for Economic Policy Research, a faculty research fellow at the National Bureau of Economic Research, a member of the Policy Council of the Association for Policy Analysis and Management, and Co-Editor of Educational Evaluation and Policy Analysis.

Christopher Candelaria is a doctoral student in the Economics of Education program at Stanford University. He graduated from Stanford University in 2006 with a B.A. in Economics. Prior to his doctoral studies, he worked for four years as a research associate at the Federal Reserve Bank of San Francisco. His research interests include teacher labor markets, education finance, and quantitative methods in education research. At Stanford, Christopher is also pursuing a Master's degree in Economics.

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**.

CITATION

Loeb, Susanna, and Christopher Candelaria. Carnegie Knowledge Network, "How Stable are Value-Added Estimates across Years, Subjects, and Student Groups?" Last modified October 2012.
URL = <http://carnegieknowledgenetwork.org/briefs/value-added/value-added-stability/>

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

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 3
October 2012
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.