



School of Education, University of Colorado at Boulder
Boulder, CO 80309-0249
Telephone: 802-383-0058

NEPC@colorado.edu
<http://nepc.colorado.edu>

DUE DILIGENCE AND THE EVALUATION OF TEACHERS

A REVIEW OF THE VALUE-ADDED ANALYSIS UNDERLYING THE
EFFECTIVENESS RANKINGS OF LOS ANGELES UNIFIED SCHOOL DISTRICT
TEACHERS BY THE *LOS ANGELES TIMES*

Derek C. Briggs and Ben Domingue, University of Colorado at Boulder

The research on which the *Los Angeles Times* relied for its teacher effectiveness reporting was demonstrably inadequate to support the published rankings. Using the same L.A. Unified School District (LAUSD) data and the same methods as the *Times*, Derek Briggs and Ben Domingue of the University of Colorado at Boulder probed deeper and found the earlier research to have serious weaknesses making the effectiveness ratings invalid and unreliable.

Last August, the *Times* published the results of a “value added” analysis of student test data, offering ratings of elementary schools and teachers in the LAUSD. The analysis was conducted by Richard Buddin, a senior economist at the RAND Corporation, as a project independent of RAND itself.” He found significant variability in LAUSD teacher quality, as demonstrated by student performance on standardized tests in reading and math, and he concluded that differences between “high-performing” and “low-performing” teachers accounted for differences in student performance.

Yet, as Briggs and Domingue explain, simply finding that a value-added model yields different outcomes for different teachers does not tell us whether those outcomes are measuring what is important (teacher effectiveness) or something else, such as whether students benefit from other learning resources outside of school. Their research explored whether there was evidence of this kind of bias by conducting what researchers call a “sensitivity analysis” to test whether the results from the *L.A. Times* model were valid and reliable.

The results of this new critical review of the research adopted and published by the *L.A. Times* are troubling: with each test of sensitivity, Briggs and Domingue found that Buddin’s approach came up short. They first investigated whether, when using the *Times* model, a student’s future

teacher would appear to have an effect on a student's test performance in the past—something that is logically impossible and a sign that the model is flawed. They found strong evidence of these illogical results when using the *L.A. Times* model, especially for reading outcomes: “Because our sensitivity test did show this sort of backwards prediction, we can conclude that estimates of teacher effectiveness in LAUSD are a biased proxy for teacher quality.”

Next, they developed an alternative, arguably stronger value-added model and compared the results to the *L.A. Times* model. In addition to the variables used in the *Times*' approach, they controlled for (1) a longer history of a student's test performance, (2) peer influence, and (3) school-level factors. If the *L.A. Times* model were perfectly accurate, there would be no difference in results between the two models. But this was not the case. For reading outcomes, the findings included the following:

- *More than half (53.6%) of the teachers had a different effectiveness rating under the alternative model.*
- *Among those who changed effectiveness ratings, some moved only moderately, but 8.1% of those teachers identified as effective under our alternative model are identified as ineffective in the L.A. Times model, and 12.6% of those identified as ineffective under the alternative model are identified as effective by the L.A. Times model.*

The math outcomes weren't quite as troubling, but the findings included the following:

- *Only 60.8% of teachers would retain the same effectiveness rating under both models.*
- *Among those who did change effectiveness ratings, some moved only moderately, but 1.4% of those teachers identified as effective under the alternative model are identified as ineffective in the L.A. Times model, and 2.7% would go from a rating of ineffective under the alternative model to effective under the L.A. Times model.*

Accordingly, the effects estimated for LAUSD teachers can be quite sensitive to choices concerning the underlying statistical model. One reasonable model leads to very different conclusions about individual teachers than does a different reasonable model.

Briggs and Domingue then examined whether Buddin's approach for determining teacher effect estimates reliably distinguishes between teachers assigned different value-added ratings. Once the specific value-added estimate for each teacher is bounded by a 95% confidence interval, they find that between 43% and 52% of teachers cannot be distinguished from a teacher of “average” effectiveness. Because the *L.A. Times* did not use this more conservative approach to distinguish teachers when rating them as “effective” or “ineffective”, it is likely that there are a significant number of false positives (teachers rated as effective who are really average), and false negatives (teachers rated as ineffective who are really average) in the *Times*' rating system.

The new report also finds evidence that conflicts with Buddin's finding that traditional teacher qualifications have no association with student outcomes. In fact, they found significant and meaningful associations between their value-added estimates of teachers' effectiveness and their experience and educational background.

For the full research report, please visit: <http://nepc.colorado.edu/publication/due-diligence>.

*The mission of the **National Education Policy Center** is to produce and disseminate high-quality, peer-reviewed research to inform education policy discussions. We are guided by the belief that the democratic governance of public education is strengthened when policies are based on sound evidence. For more information on NEPC, please visit <http://nepc.colorado.edu/>. The policy brief was made possible in part by the support of the Great Lakes Center for Education Research and Practice (greatlakescenter.org).*