What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

July 2014

WWC Review of the Report "Incentives, Selection, and Teacher Performance: Evidence from IMPACT"[1,2]

The findings from this review do not reflect the full body of research evidence on the *IMPACT* evaluation system.

What is this study about?

The study examined the effects of *IMPACT*, the teacher evaluation system used in the District of Columbia Public Schools (DCPS), on teacher retention and performance. *IMPACT* assigns each teacher a single performance score based on classroom observations, student achievement, core professionalism, and their contributions to the school. Based on this score, teachers are assigned one of four ratings: Highly Effective, Effective, Minimally Effective, or Ineffective. Highly Effective teachers receive sizeable increases in compensation, Minimally Effective teachers are scheduled for dismissal if improvement does not occur in 1 year, and Ineffective teachers are immediately dismissed.

Researchers used a regression discontinuity design to examine the effects of receiving a rating of Minimally Effective or Highly Effective (instead of Effective) on teacher retention and *IMPACT* scores in the next school year. To measure the effect of receiving a Minimally Effective rating, the study compared about 4,000 teachers who had *IMPACT* scores just above and just below the cutpoint between Minimally Effective and Effective. Similarly, the study measured the effect of receiving a Highly Effective rating by comparing about 2,000 teachers who had *IMPACT* scores just above and just below the cutpoint between Effective and Highly Effective.

Features of the *IMPACT* Evaluation System

DCPS introduced the *IMPACT* evaluation system in the 2009–10 school year. *IMPACT* assigns each teacher a single performance score that is a weighted average of classroom observations (scored with the Teacher and Learning Framework), a measure of student achievement based on standardized tests for reading and math or other assessments, a measure of all students' achievement at the school, a measure of the teacher's commitment to the school community, and a measure of the teacher's core professionalism. The Teacher and Learning Framework used for classroom observations includes a scoring rubric based on dimensions that DCPS has identified as relating to effective instruction, such as the structure and clarity of lessons.

Based on the performance score, teachers are assigned one of four ratings: Highly Effective, Effective, Minimally Effective, or Ineffective. During the first 3 years of *IMPACT*, teachers rated as Highly Effective received an immediate bonus as large as $25,000 and were eligible for permanent increases in base salary. If they received a Highly Effective rating in 2 consecutive years, their annual base salary could increase by up to $27,000. Teachers rated as Effective received a typical base pay salary increase. Teachers rated as Minimally Effective did not receive a typical base salary increase and were subject to dismissal if they received a Minimally Effective rating for a second year. Teachers rated as Ineffective were immediately dismissed.

What did the study find?

None of the analyses presented in this study meet WWC standards, and therefore, the study findings are not presented in this WWC report.

WWC Rating

The research described in this report does not meet WWC regression discontinuity design standards

The information needed to determine attrition and baseline equivalence was not provided in the study. Consequently, the WWC could not assess whether changes in teacher retention and performance were related to differences in the teacher characteristics at baseline.

Endnotes

[1] Dee, T., & Wyckoff, J. (2013). *Incentives, selection, and teacher performance: Evidence from IMPACT* (NBER Working Paper 19529). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org

[2] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the authors) to assess whether the study design meets WWC regression discontinuity design standards. The review reports the WWC's assessment of whether the study meets WWC regression discontinuity design standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the single study review protocol, version 2.0. A quick review of this study was released on December 18, 2013, and this report is the follow-up review that replaces that initial assessment.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2014, July). *WWC review of the report: Incentives, selection, and teacher performance: Evidence from IMPACT.* Retrieved from http://whatworks.ed.gov

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Improvement index Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.

Single-case design (SCD) A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample are spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 3.0) for additional details.