What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

May 2014 Revised*

WWC Review of the Report "The Effectiveness of Secondary Math Teachers from Teach for America and the Teaching Fellows Programs"[1,2,3]

The findings from this review do not reflect the full body of research evidence on *The New Teacher Project Teaching Fellows* program.

What is this study about?

The study examined whether students taught by teachers in the *Teach for America (TFA)* and *The New Teacher Project Teaching Fellows (Teaching Fellows)* programs had greater mathematics achievement than students taught by teachers who were not in either of these programs. This WWC report focuses on the *Teaching Fellows* intervention, and a separate single study review provides information about the *TFA* intervention.

The *Teaching Fellows* study included two cohorts of students in grades 6–12 (one from the 2009–10 school year and one from the 2010–11 school year). Students came from eight states, nine school districts, and 44 schools. Within each school, students were randomly assigned to a mathematics classroom taught by either a *Teaching Fellows* teacher or a comparison teacher who did not enter teaching through a highly selective alternative route. In total, the *Teaching Fellows* study randomly assigned over 7,000 students to either a *Teaching Fellows* teacher or a comparison teacher.

Mathematics achievement was assessed for students in grades 6–8 using scores from state-required assessments. For students in grades 9–12, the authors administered end-of-course math assessments developed by the Northwest Evaluation Association (NWEA). The study compared the student performance on these assessments for *Teaching Fellows* teachers and all comparison teachers. The study also compared student performance for subgroups of *Teaching Fellows* and comparison teachers in the same schools based on: (1) whether comparison teachers entered teaching through traditional routes or through less selective alternative certification routes; (2) whether or not *Teaching Fellows* teachers and comparison teachers were in their first 3 years of teaching (novice or experienced teachers); (3) whether *Teaching Fellows* teachers and comparison teachers had similar levels of experience (i.e., difference of no more than 2 years); and (4) whether teachers were in middle schools or high schools.

Features of *The New Teacher Project Teaching Fellows (Teaching Fellows)*

Teaching Fellows is a highly selective alternative certification program for new teachers. The program recruits, selects, trains, places, and provides support to teachers in low-income, high-need schools across the country.

Teaching Fellows recruits both new college graduates and professionals who would like to switch careers. Training is intended to provide over 68 hours of instruction and 64 hours of fieldwork during a summer institute, and most *Teaching Fellows* teachers also attend state-authorized alternative certification programs. *Teaching Fellows* teachers are asked to make an open-ended commitment to teaching.

What did the study find?

The study authors found, and the WWC confirmed, that *Teaching Fellows* teachers did not differ in effectiveness on improving student mathematics achievement scores from comparison teachers overall and from comparison teachers who entered teaching through traditional routes. The study authors also found, and the WWC confirmed, that there were no differences in effectiveness among the subgroups of experienced teachers, middle school teachers, and high school teachers. The study authors found, and the WWC confirmed, that *Teaching Fellows* teachers were more effective than comparison teachers from less selective alternative certification routes, novice *Teaching Fellows* teachers were more effective than novice comparison teachers, and novice *Teaching Fellows* teachers were less effective than experienced comparison teachers.

WWC Rating

The research described in this report meets WWC group design standards without reservations

The *Teaching Fellows* study is a randomized controlled trial with low levels of sample attrition.[4]

Appendix A: Study details

Clark, M. A., Chiang, H. S., Silva, T., McConnell, S., Sonnenfeld, K., Erbe, A., & Puma, M. (2013). *The effectiveness of secondary math teachers from Teach For America and the Teaching Fellows programs* (NCEE 2013-4015). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Setting

The New Teacher Project Teaching Fellows (Teaching Fellows) study was conducted in eight states, nine school districts, and 44 schools.

Study sample

In each participating school, researchers matched two or more math classrooms so that at least one classroom would be taught by a *Teaching Fellows* teacher, and at least one classroom would be taught by a teacher who did not participate in a highly selective alternative route to certification (such as *Teaching Fellows* or *TFA*). Students in grades 6–12 were randomly assigned to these classrooms, typically at the beginning of the school year. However, students who entered a participating school after the beginning of the school year were randomly assigned to classrooms through rolling random assignment. The *Teaching Fellows* study was conducted with two cohorts of students in the 2009–10 and 2010–11 school years. The *Teaching Fellows* study sample was comprised primarily of high school classrooms (69%).

The *Teaching Fellows* study was conducted in 44 schools located in eight states, with 118 classroom matches covering 69 intervention and 84 comparison teachers. For this study, 3,659 students were randomly assigned to a *Teaching Fellows* teacher, and 3,629 students were randomly assigned to a comparison teacher. The analytic sample included 2,127 students in the intervention group and 1,989 students in the comparison group.

Intervention group

The intervention group consisted of classrooms of *Teaching Fellows* teachers. *Teaching Fellows* recruits, selects, trains, places, and provides support to teachers in low-income, high-need schools across the country. The *Teaching Fellows* program recruits both new college graduates and professionals who would like to switch careers. The program is highly selective (*Teaching Fellows* accepts 13% of applicants), and recruits apply to programs located in the region where they intend to teach. The application process for the *Teaching Fellows* program involves a written application, an analysis and writing exercise, an in-person interview, a monitored group discussion, and a 5-minute sample teaching lesson. Once selected, *Teaching Fellows* teachers perform 25 hours of independent study, participate in a summer institute, and participate in a local certification program. The summer institute consists of 68 hours of instruction and 64 hours of fieldwork. *Teaching Fellows* teachers are asked to make an open-ended commitment to teaching.

Comparison group

Classrooms were eligible to be in the comparison group if their teachers were certified through a route other than *TFA*, *Teaching Fellows*, or a similar program. The comparison group includes teachers who entered teaching through either a traditional route or a less selective alternative certification route.

Outcomes and measurement

For students in grades 6–8, the authors measured mathematics achievement using scores obtained on state-required assessments. For students in grades 9–12, the authors administered end-of-course mathematics assessments developed by the Northwest Evaluation Association (NWEA). All test scores were converted to *z*-scores to facilitate comparisons across classrooms, schools, and districts. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Teaching Fellows recruits are required to (a) complete 25 hours of independent study and a 4-hour orientation meeting and (b) attend a 5–7 week summer institute that includes 68 hours of instruction and an average of 64 hours of fieldwork via fifteen 5-hour days in a summer school classroom. In addition, most teachers attend state-authorized alternative certification programs. During the school year, *Teaching Fellows* teachers are provided information, feedback, and mentoring through the *Teaching Fellows* program.

Reason for review

This study was identified for review by the WWC by receiving media attention.

Appendix B: Outcome measures for the mathematics achievement domain

Mathematics achievement	
Mathematics assessments	For students in grades 6–8, the authors measured mathematics achievement using scores from state-required assessments. These assessments were administered at the end of the school year in which the students were randomly assigned. For students in grades 9–12, the authors administered end-of-course mathematics assessments developed by the Northwest Evaluation Association (NWEA). These computer-adaptive assessments covered general high school math, Algebra I, Geometry, or Algebra II, depending upon the course content. All test scores were converted to *z*-scores. For state assessments administered to middle school students, the reference population was the full population of students in the same state, year, and grade who took the same assessment. For NWEA end-of-course assessments administered to high school students, the reference population was the NWEA's nationwide norming sample for that assessment.

Appendix C: Study findings for the mathematics achievement domain, *Teaching Fellows* study

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mathematics achievement								
Mathematics assessments	All teachers	44 schools/ 153 teachers/ 4,116 students	–0.39 (nr)	–0.39 (nr)	0.00	0.00	0	.96
Domain average for mathematics achievement						**0.00**	**0**	**Not statistically significant**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The statistical significance of the study's domain average was determined by the WWC. nr = not reported.

Study Notes: The WWC did not need to make corrections for clustering, multiple comparisons, or to adjust for baseline differences. The *p*-value presented here was reported in the original study. The effect size is the impact estimate from the study because the outcome was scaled to be in standard deviation units. This study is characterized as having an indeterminate effect because the effect for this intervention was neither statistically significant or substantively important. For more information, please refer to the WWC Standards and Procedures Handbook, version 3.0, pages 26–27.

Appendix D: Supplemental findings for the mathematics achievement domain, *Teaching Fellows* study

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mathematics achievement								
Mathematics assessments	*Teaching Fellows* teachers and comparison teachers from traditional routes	33 schools/ 113 teachers/ 3,268 students	–0.36 (nr)	–0.32 (nr)	–0.03	–0.03	–1	.25
Mathematics assessments	*Teaching Fellows* teachers and comparison teachers from less selective alternative certification routes	19 schools/ 46 teachers/ 902 students	–0.50 (nr)	–0.63 (nr)	0.13	0.13	5	.01
Mathematics assessments	Novice teachers	7 schools/ 17 teachers/ 354 students	–0.40 (nr)	–0.53 (nr)	0.13	0.13	5	< .01
Mathematics assessments	Novice *Teaching Fellows* teachers and experienced comparison teachers	19 schools/ 53 teachers/ 1,153 students	–0.63 (nr)	–0.53 (nr)	–0.10	–0.10	–4	< .01
Mathematics assessments	Experienced teachers	26 schools/ 80 teachers/ 2,408 students	–0.27 (nr)	–0.30 (nr)	0.03	0.03	1	.45
Mathematics assessments	Teachers with similar levels of experience	17 schools/ 46 teachers/ 1,283 students	–0.17 (nr)	–0.20 (nr)	0.03	0.03	1	.40
Mathematics assessments	Middle school teachers	17 schools/ 53 teachers/ 1,610 students	–0.35 (nr)	–0.39 (nr)	0.04	0.04	2	.38
Mathematics assessments	High school teachers	29 schools/ 101 teachers/ 2,506 students	–0.47 (nr)	–0.39 (nr)	–0.02	–0.02	–1	.47

Table Notes: The supplemental findings presented in this table are additional findings that do not factor into the determination of the evidence rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on individual outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. nr = not reported.

Study Notes: The study defines novice teachers as those in their first 3 years of teaching. Experienced teachers are those with more than 3 years of teaching experience. Teachers were determined to have similar levels of experience if their difference in experience was no more than 2 years. A correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The *p*-values presented here were reported in the original study.

Endnotes

* On September 1, 2015, the WWC modified this report to include additional information on subgroups of teachers that was made available during a review of this study for a grant competition. Based on this information, the WWC found that these subgroup analyses met WWC group design standards without reservations. The WWC has changed the description of the study and its findings, Appendix A, and added Appendix D to reflect the subgroup analyses. The WWC has also changed Appendix C to report the adjusted means for the intervention and comparison groups. The WWC has not changed Appendix B since the May 2014 report.

[1] 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 evidence standards. The review reports the WWC's assessment of whether the study meets WWC evidence 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 November 7, 2013, and this report is the follow-up review that replaces that initial assessment.

[2] Absence of conflict of interest: This study was conducted by staff from Mathematica Policy Research. Therefore, Mathematica reviewers were not involved in the WWC review of this study.

[3] The study also examined the impact of *Teach for America* relative to a separate comparison group that was formed by random assignment. The findings from that analysis are reported in a separate single study review by the WWC because the impacts of the two programs were measured separately. The authors also noted that the study was not designed to compare the effectiveness of *Teach For America* relative to *Teaching Fellows*.

[4] The WWC verified low study attrition through an author query.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2014, May). *WWC review of the report: The effectiveness of secondary math teachers from Teach for America and the Teaching Fellows programs*. 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% ($p < .05$).

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.