What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

October 2013

WWC Review of the Report "Charter School Performance in New Jersey"[1]

The findings from this review do not reflect the full body of research evidence on charter schools.

What is this study about?

The study examined the effect of 60 New Jersey charter schools on annual student achievement growth in reading and math. The study was based on data from a large sample of students in grades 3–8 during 5 school years (2006–07 through 2010–11).

The study authors matched charter school students to students attending traditional public schools based on test scores and demographic characteristics. This matching process was conducted separately for each of the 5 study years. At least one match for 77% of charter school students in the sample (a total of about 16,500 students) could be identified.

Study authors assessed the effectiveness of the charter schools by comparing average 1-year achievement gains of charter and comparison students on standardized reading and math tests. This calculation was based on the pooled sample of students in grades 3–8 across the 5 study years.

The authors also examined charter school effects separately for Newark (a major city in New Jersey) and a collection of other major cities (Camden, Jersey City, Paterson, and Trenton).

Features of New Jersey Charter Schools

Charter schools are public schools that are established on the basis of a contract, or charter, that a private board holds.

They are released from many state and district regulations that govern traditional public schools, including those involving staffing, curriculum, and budget decisions.

New Jersey charter schools aim to improve student learning and achievement by encouraging the use of different and innovative learning methods and establishing accountability systems.

What did the study find?

The study authors reported that 1-year reading and math achievement gains of students in the charter schools were significantly greater than those of comparison students in traditional public schools.

Charter school effects in Newark were positive, statistically significant, and larger than for New Jersey overall, while charter school effects for the other major cities in New Jersey were negative and statistically significant in reading and zero in math.

WWC Rating

The research described in this report meets WWC evidence standards with reservations

Strengths: This is a large, multi-year analysis in which the authors matched charter school students with students from traditional public schools (that charter students would likely otherwise attend) using demographic and academic characteristics.

Cautions: Although the study matched charter school students with traditional public school students based on observed demographic characteristics and test scores, unobserved differences between the two groups may exist. For example, charter school students may be more motivated to do well in school or have other unobserved characteristics that influence student achievement. This means the study's results do not necessarily isolate the effect of charter schools.

In addition, the study's results do not have a straightforward interpretation because they blend the 1-year gains students experienced during their first year of charter school attendance and 1-year gains during subsequent years.

Finally, the effect sizes reported in this study (which are based on an analysis of achievement gains) are not directly comparable to effect sizes reported by other studies that analyzed achievement levels.

Appendix A: Study details

Center for Research on Education Outcomes. (2012). *Charter school performance in New Jersey.* Palo Alto, CA: Stanford University.

Setting

The study was based on data collected from students who attended 60 charter schools and students who attended traditional public schools in New Jersey.

Study sample

The study was based on an analysis of standardized reading and math test data for students in grades 3–8 from school year 2006–07 to school year 2010–11. During this 5-year period, the number of charter schools in New Jersey grew from 58 to 83. During each study year, all charter school students in grades 3–8 were included in the analysis, as long as they could be matched to a student from a traditional public school, for a total of 16,500 charter school students. For each year of the study, the authors identified feeder schools—defined as traditional public schools that charter school students either attended in the past or (for charter students who never attended a traditional school) would have attended had the charter schools not existed. Each charter student was matched to one or more students from the relevant feeder schools on exact values for a collection of measures, including race/ethnicity, gender, English proficiency, subsidized lunch status, special education status, and grade level. Additional matching was done using achievement test scores from the prior school year. To be considered a valid match, the prior year test scores of charter and comparison students had to be within 0.10 standard deviations of each other. Using this process, it was possible for a charter school student to be matched to more than one traditional public school student. For charter students that had multiple matches, one comparison record was created by averaging the outcomes of up to seven traditional public school students with whom the charter school student was matched. Valid matches were identified for 77% of the tested charter school students. The 23% of charter school students who did not have matches were excluded from the analysis; many of these were grade repeaters.

Intervention group

The intervention group included students who were enrolled in 60 New Jersey charter schools in grades 3–8 during 5 school years (2006–07 through 2010–11). Charter schools are public schools that are established on the basis of a contract, or charter, that a private board holds. They are released from many state and district regulations that govern traditional public schools, including those involving staffing, curriculum, and budget decisions. New Jersey charter schools aim to improve student learning and achievement by encouraging the use of different and innovative learning methods and establishing accountability systems.

Comparison group

The comparison group included students from traditional public schools that charter school students in the intervention group either attended in the past or, for charter students who never attended a traditional school, would have attended had the charter schools not existed.

Outcomes and measurement

The evaluation included two outcome measures: scores from standardized reading and math tests that were administered to students in grades 3–8 in the spring of each school year from 2006–07 through 2010–11. The data were obtained from the New Jersey Department of Education. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Not applicable.

Reason for review

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

Appendix B: Outcome measures for each domain

General reading achievement	
New Jersey Assessment of Skills and Knowledge (NJ ASK): Language Arts scale	According to the New Jersey Department of Education's website, this standardized test was designed around New Jersey's Core Curriculum Content Standards. It is intended to provide information about a student's achievement in language arts. The test includes reading passages, multiple-choice items, constructed-response items, and writing tasks.
General math achievement	
NJ ASK: Mathematics scale	According to the New Jersey Department of Education's website, this standardized test was designed around New Jersey's Core Curriculum Content Standards. It is intended to provide information about a student's achievement in mathematics. The test includes multiple-choice items, as well as short and extended constructed-response items.

Appendix C: Study findings for each domain

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	
General reading achievement								
NJ ASK: Language Arts	Full sample, grades 3–8	59 charter schools/16,486 charter students	0.07 (0.60)	0.02 (0.44)	0.05	0.05	+2	< 0.01
Domain average for general reading achievement					**0.05**	**0.05**	**+2**	**Statistically significant**
General math achievement								
NJ ASK: Mathematics	Full sample, grades 3–8	60 charter schools/16,547 charter students	0.09 (0.60)	0.02 (0.43)	0.07	0.07	+3	< 0.01
Domain average for general math achievement					**0.07**	**0.07**	**+3**	**Statistically significant**

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. 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. Because this study examined year-to-year gains in achievement, the effect sizes and improvement indices presented are not directly comparable to effect sizes that would be estimated in an analysis of achievement levels. The statistical significance of the study's domain average was determined by the WWC. The study is characterized as having statistically significant positive effects in reading and math because univariate statistical tests are reported for each outcome measure, the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant. NJ ASK = New Jersey Assessment of Skills and Knowledge.

Study Notes: No corrections for clustering or multiple comparisons were needed. Means and standard deviations were provided by the authors at the request of the WWC. Therefore, WWC calculations presented in this table differ from those reported by the study authors because the former are based on the simple means and standard deviations provided to the WWC by the authors, while the latter are based on estimates from the regression model, which included covariates. The *p*-values presented here were reported in the original study and are based on covariate-adjusted models.

Appendix D: Supplemental findings by domain

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	
General reading achievement								
NJ ASK: Language Arts	Newark, grades 3–8	13 charter schools/3,853 charter students	0.10 (0.62)	–0.05 (0.43)	0.15	0.15	+6	< 0.01
NJ ASK: Language Arts	Other major cities, grades 3–8	20 charter schools/4,728 charter students	0.06 (0.58)	0.04 (0.44)	0.01	0.01	0	< 0.01
General math achievement								
NJ ASK: Mathematics	Newark, grades 3–8	13 charter schools/3,887 charter students	0.15 (0.61)	–0.05 (0.45)	0.20	0.20	+8	< 0.01
NJ ASK: Mathematics	Other major cities, grades 3–8	20 charter schools/4,719 charter students	0.09 (0.58)	0.04 (0.43)	0.05	0.05	+2	> 0.05

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. 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. Because this study examined year-to-year gains in achievement, the effect sizes and improvement indices presented are not directly comparable to effect sizes that would be estimated in an analysis of achievement levels. NJ ASK = New Jersey Assessment of Skills and Knowledge.

Study Notes: No corrections for clustering or multiple comparisons were needed. Means and standard deviations were provided by the authors at the request of the WWC. Therefore, WWC calculations presented in this table differ from those reported by the study authors because the former are based on the raw means and standard deviations provided to the WWC by the authors, while the latter are based on estimates from the regression model, which included covariates. The *p*-values presented here were reported in the original study and are based on covariate-adjusted models.

Endnotes

[1] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the author[s]) 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. The WWC rating applies only to the summarized results, and not necessarily to all results presented in the study. This study was reviewed using the single study review protocol, version 2.0. A quick review of this study was released on December 21, 2012, 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. (2013, October). *WWC review of the report: Charter school performance in New Jersey*. 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 2.1) for additional details.