

What Works Clearinghouse



August 2012

WWC Review of the Report “Milwaukee Parental Choice Program Longitudinal Educational Growth Study Fifth Year Report”¹

The findings from this review do not reflect the full body of research evidence on the *Milwaukee Parental Choice Program*.

What is this study about?

The study examined the effectiveness of the *Milwaukee Parental Choice Program (MPCP)*, which provides vouchers for low-income students to attend private schools.

The study analyzed data on about 600 students who were given *MPCP* vouchers in the 2006–07 school year. The authors created a comparison group of similar size by matching the *MPCP* students to students who attended Milwaukee Public Schools.

The study estimated the effect of the voucher program by comparing *MPCP* students’ standardized test scores in 2010—four years after they were given a voucher—with those of comparison group students.

What did the study find?

The study found that the reading test scores of *MPCP* students in 2010 were 0.15 standard deviations higher than they were for the matched comparison sample, a statistically significant difference. There was no significant difference in math test scores in 2010.

WWC Rating

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

Strengths: The intervention and comparison groups were matched on neighborhood locations, math and reading test scores in 2006, grade level, and demographic characteristics using a statistical procedure.

Cautions: Because it is not a randomized controlled trial, the study cannot ensure that the differences in achievement between the *MPCP* students and the comparison students are due solely to the *MPCP*. In addition, in the 2010–11 school year, the district implemented a new policy to publicize the test scores of *MPCP* students at each school. This policy could have affected the academic achievement of *MPCP* students in that year.

Features of the *Milwaukee Parental Choice Program (MPCP)*

The *MPCP* is a program that provides vouchers for children in low-income families to attend participating secular or religious private schools of their choice.

In 2010, about 21,000 children participated in the *MPCP*, and the maximum voucher amount was nearly \$6,500.

Appendix A: Study details

Witte, J. F., Carlson, D., Cowen, J. M., Fleming, D. J., & Wolf, P. J. (2012). *Milwaukee Parental Choice Program longitudinal educational growth study fifth year report*. Report of the School Choice Demonstration Project, University of Arkansas, Fayetteville. SCDP Milwaukee Evaluation Report #29.

Setting	The study was conducted in Milwaukee, Wisconsin.
Study sample	<p>The intervention group consisted of students who participated in the <i>Milwaukee Parental Choice Program (MPCP)</i> to attend private schools in the 2006–07 school year. The comparison group consisted of a matched set of students who attended Milwaukee Public Schools in the 2006–07 school year. The intervention and comparison groups were matched based on neighborhood location, test scores in 2006–07, grade level, and demographic information using a statistical matching process.</p> <p>The matching process is detailed in the baseline evaluation report. Specifically, each <i>MPCP</i> student was matched with a comparison student within the same census tract and test score group. Test score groups were created in five-percentile intervals based on the mean of a student’s math and reading test scores in 2006. For example, one test score group contained everyone whose mean score in math and reading fell between the 15th and 20th percentiles. If more than one potential comparison student was identified within tract and within test score group, the researchers identified the comparison student whose estimated propensity for <i>MPCP</i> participation was closest to the propensity of the <i>MPCP</i> student. Propensity of <i>MPCP</i> participation was predicted based on the mean of students’ math and reading test scores, gender, race, and English language learner status. If the propensity for the <i>MPCP</i> student could not be estimated because of missing data, then the comparison student was chosen at random. If the <i>MPCP</i> student was missing a test score, then the comparison student was selected at random within the census tract. The analysis of math scores was based on 633 <i>MPCP</i> students and 674 comparison students, and the analysis of reading scores was based on 636 <i>MPCP</i> students and 672 comparison students.²</p>
Intervention group	Students in the intervention group were selected from a list of students who had applied to and been accepted into the <i>MPCP</i> in the 2006–07 school year.
Comparison group	Students in the comparison group attended a Milwaukee Public School in the 2006–07 school year. Comparison students may have participated in the <i>MPCP</i> in other years.
Outcomes and measurement	The eligible outcomes in this study were math and reading scores from the Wisconsin Knowledge and Concepts Examination (WKCE) measured in November 2010. The intervention and comparison students in the analysis were in grades 3, 4, and 6 in the 2006–07 school year and grades 7, 8, and 10 in the 2010–11 school year. The test scores were standardized into z-scores based on the school district’s means and standard deviations of WKCE math and reading test scores. For a more detailed description of these outcome measures, see Appendix B.
Reason for review	This study was identified for review by the WWC because it received significant media attention.

Appendix B: Outcome measures for each domain

General mathematics achievement

Wisconsin Knowledge and Concepts Examination (WKCE)—Math

Students' math test scores in 2010 on the WKCE were standardized to have a mean of zero and standard deviation of one across the district and grade.

General literacy achievement

WKCE—Reading

Students' reading test scores in 2010 on the WKCE were standardized to have a mean of zero and standard deviation of one across the district and grade.

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 mathematics achievement								
<i>WKCE—Math</i>	All students	636 MPCP students and 672 matched comparison students	nr	nr	0.07	0.07	+3	> 0.05
Domain average for general mathematics achievement						0.07	+3	Not statistically significant
General literacy achievement								
<i>WKCE—Reading</i>	All students	633 MPCP students and 674 matched comparison students	nr	nr	0.15	0.15	+6	< 0.05
Domain average for general literacy achievement						0.15	+6	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. 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. The statistical significance of the study's domain average was determined by the WWC; the study is characterized as having a statistically significant positive effect because univariate statistical tests are reported for each outcome measure, at least half of the effects are positive and statistically significant, and no effects are negative and statistically significant. nr = not reported. WKCE = Wisconsin Knowledge and Concepts Examination.

Study Notes: No corrections for clustering or multiple comparisons were needed. Intervention and comparison group means were not reported. The mean difference is the estimated impact from the regression reported for Model 2, Table 4 of the paper. Because the outcomes were standardized, the effect size equals the mean difference. The improvement index was calculated by the WWC, and the p-values presented here were reported by the authors.

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. This study was reviewed using the Single Study review protocol, version 2.0. A quick review of this study was released on April 6, 2012, and this report is the follow-up review that replaces that initial assessment. The WWC rating applies only to the results that were eligible under this topic area and met WWC standards without reservations or met WWC standards with reservations, and not necessarily to all results presented in the study.

² Citation for baseline report: Witte, J. F., Wolf, P. J., Cowen, J. M., Fleming, D. J., & Lucas-McLean, J. (2008). *MPCP Longitudinal Educational Growth Study Baseline Report*. Report of the School Choice Demonstration Project, University of Arkansas, Fayetteville. SCDP Milwaukee Evaluation Report #5.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, August). *WWC review of the report: Milwaukee Parental Choice Program longitudinal educational growth study fifth year report*. 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 tend to be 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.