

WWC Review of the Report “Evaluation of the College Possible Program: Results from a Randomized Controlled Trial”^{1,2}

The findings from this review do not reflect the full body of research evidence on the *College Possible* program.

What is this study about?

The study investigated the effect of the *College Possible* program, which is designed to serve low-income high school students. *College Possible* provides a 2-year after-school curriculum to high school juniors and seniors including SAT and ACT test preparation services, college admissions and financial aid consulting, and guidance in the transition to college.

Students apply as high school sophomores and enter the 2-year program as juniors. The program is limited to students from families below the median city/county household income with a suggested minimum GPA of 2.0. There are no costs to students, though students do agree to volunteer for 8 hours of community service per year in exchange for the services.

Over the course of 2 years, each participant in the *College Possible* program is scheduled to receive a total of 320 hours of direct program services. The sample was 91% minority, and most were potential first-generation college students.

Features of *College Possible* Program

Designed for low-income students, *College Possible* is a 2-year after-school curriculum that features:

- SAT/ACT test preparation
- Financial aid consulting
- College admissions guidance

The program is offered at no cost to partner high schools (the schools do provide classroom space for the program to operate). The program is free for participants, but participants do agree to provide 8 hours of community service each year in exchange for program services.

What did the study find?

The study author reported that there was a statistically significant effect of the *College Possible* program on the average number of college applications submitted (8.8 in the intervention group vs. 4.1 in the comparison group). Follow-up analyses revealed that students in the intervention group were more likely to apply to 4-year institutions (93% vs. 63%) and sent more applications to selective institutions (3.9 vs. 1.4) than students in the comparison group. The WWC confirmed the statistical significance of all of these effects.

Overall college enrollment rates were the same in both groups (about 64%). Students in the intervention group enrolled in selective institutions at a higher rate than students in the comparison group (12% vs. 5%), but this difference was not statistically significant. The author did report that the proportion of students enrolling in a 4-year institution was higher for the intervention group (45% vs. 34%), but the WWC found that this effect was not statistically significant after adjusting for multiple comparisons.

Finally, the author reported no statistically significant effect of the intervention on ACT scores (17.9 vs. 18.4).

This study also included an analysis that employed a regression discontinuity design, but that analysis does not meet WWC standards because the forcing variable does not have a sufficient number of unique values.

WWC Rating

The research described in this report meets WWC evidence standards with reservations³

This was a quasi-experiment for which the WWC could establish that the intervention and comparison groups were comparable at baseline.

The regression discontinuity (RD) design portion of this study did not meet pilot WWC RD standards because the forcing variable does not have a sufficient number of unique values.

Appendix A: Study details

Avery, C. (2013). *Evaluation of the College Possible program: Results from a randomized controlled trial* (NBER Working Paper 19562). Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org>

Setting	The study took place in eight high schools in St. Paul and Minneapolis, Minnesota.
Study sample	The study examined the effects of the <i>College Possible</i> program on 134 intervention students and 104 comparison students. Of these, 91% were racial/ethnic minorities, and 90% reported that they were potential first-generation college students. The average family income of the sample was \$25,000.
Intervention group	The intervention group was comprised of students who were randomly assigned to receive the <i>College Possible</i> program.
Comparison group	The comparison group was comprised of students who were randomly assigned not to receive the <i>College Possible</i> program.
Outcomes and measurement	The study reported findings for seven eligible outcomes in three domains—academic achievement in high school (ACT scores), college application submission (number of applications submitted, likelihood of applying to a 4-year institution, number of applications submitted to selective institutions—defined using Barron’s “most competitive,” “highly competitive,” and “very competitive” categories), and enrollment (enrolled in college, enrolled in a 4-year institution, enrolled in a competitive institution). The outcome measures were based on data from ACT, Inc., participating high schools, and the National Student Clearinghouse.
Support for implementation	The program is implemented by <i>College Possible</i> staff. There are no direct participation costs for the participating high schools, though they do provide office space for the full-time program coordinator and classroom space after school hours.
Reason for review	This study was reviewed by the WWC in response to receiving significant media attention.

Appendix B: Outcome measures for each domain

Academic achievement	
<i>ACT scores (composite)</i>	This outcome is based on data from ACT, Inc.
Postsecondary applications	
<i>College applications (all)</i>	This outcome is based on data from the participating high schools and measures the total number of college applications students submitted. Of note is that half of the comparison group high schools did not keep formal records of college applications, so the analysis is restricted to intervention and comparison students in the four high schools that did keep these records.
<i>Applied to a 4-year institution</i>	This outcome is based on data from the participating high schools and measures whether students applied to a 4-year institution. Of note is that half of the comparison group high schools did not keep formal records of college applications, so the analysis is restricted to intervention and comparison students in the four high schools that did keep these records.
<i>College applications, selective institutions</i>	This outcome is based on data from the participating high schools and measures the total number of applications students submitted to selective institutions. Selectivity was defined using the Barron's "most competitive," "highly competitive," and "very competitive" categories. Of note is that half of the comparison group high schools did not keep formal records of college applications, so the analysis is restricted to intervention and comparison students in the four high schools that did keep these records.
Postsecondary enrollment	
<i>College enrollment (any)</i>	This outcome is based on data from the National Student Clearinghouse and measures whether students enrolled in college.
<i>College enrollment, 4-year institution</i>	This outcome is based on data from the National Student Clearinghouse and measures whether students enrolled in a 4-year institution.
<i>College enrollment, selective institution</i>	This outcome is based on data from the National Student Clearinghouse and measures whether students enrolled in a selective institution. Selectivity was defined using the Barron's "most competitive," "highly competitive," and "very competitive" categories.

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	
Academic achievement								
<i>ACT score (composite)</i>	Full sample	171	17.90 (4.04)	18.40 (4.26)	-0.50	-0.03	-1	> .05
Domain average for academic achievement						-0.03	-1	Not statistically significant
Postsecondary applications								
<i>College applications (all)</i>	Full sample	139	8.80 (nr)	4.10 (nr)	4.80	1.17	38	< .05
<i>Applied to a 4-year institution</i>	Full sample	139	93%	63%	30%	0.78	28	< .05
<i>College applications, selective institutions</i>	Full sample	139	3.90 (nr)	1.40 (nr)	2.50	0.68	25	nr
Domain average for postsecondary applications						+0.87	31	Statistically significant
Postsecondary enrollment								
<i>College enrollment (any)</i>	Full sample	238	64%	64%	0%	0.03	1	> .05
<i>College enrollment, 4-year institution</i>	Full sample	238	45%	34%	11%	0.29	11	< .05
<i>College enrollment, selective institution</i>	Full sample	238	12%	5%	7%	0.20	8	> .05
Domain average for postsecondary enrollment						+0.17	7	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 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. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. nr = not reported.

Study Notes: The p-values presented here were reported in the original study. A correction for multiple comparisons was needed and resulted in a WWC-computed critical p-value of 0.0167 for the College enrollment, 4-year institution outcome; therefore, the WWC does not find the result to be statistically significant. Data from the College applications, selective institutions outcome were obtained from Table 7 of the study report and from the study's author. The effect size for this outcome represents the number of applications sent to most, highly, and very competitive institutions (see Appendix B) and was computed by the WWC. Of note is that half of the comparison group high schools did not keep formal records of college applications, so the analysis is restricted to intervention and comparison students in the four high schools that did keep these records.

Endnotes

* On December 19, 2014, the WWC modified this report in response to new information provided by the study author. Specifically, the study author provided baseline standard deviations for GPA for the intervention and comparison groups, which allowed the WWC to assess the comparability of these groups on GPA at baseline. Based on the new information, the review team updated the study rating from *does not meet WWC evidence standards* to *meets WWC evidence standards with reservations*.

¹ Avery, C. (2013). *Evaluation of the College Possible program: Results from a randomized controlled trial* (NBER Working Paper 19562). Cambridge, MA: National Bureau of Economic Research.

² Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the author) to assess whether the study design meets WWC standards. The review reports the WWC's assessment of whether the study meets WWC standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the Postsecondary Education topic area review protocol, version 2.0. A quick review of this study was released on December 19, 2013, and this report is the follow-up review that replaces that initial assessment. The WWC rating applies only to the study outcomes that were eligible for review under this topic area. The reported analyses in this SSR are only for those eligible outcomes that either met WWC standards without reservations or met WWC standards with reservations, and do not necessarily apply to all results presented in the study.

³ Though the author describes the study as a randomized controlled trial, it is not a randomized controlled trial according to the WWC's definition of that term. To study the effects of the intervention, students were matched on high school GPA. Then, within each set of matched students, the study author alternated between assigning the highest and the lowest GPA student to the *College Possible* program. The WWC does not consider this assignment mechanism to be random or functionally random (see the WWC's *Procedures and Standards Handbook*, version 3.0, p. 9).

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2014, July). *WWC review of the report: Evaluation of the College Possible program: Results from a randomized controlled trial*. 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 individual 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 individuals, the improvement index represents the gain or loss of the average individual due to the intervention. As the average individual 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 study participants 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 eligible study participants are randomly assigned to 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.