

Appendix

Appendix A1 Study Characteristics: Dynarski, Gleason, Rangarajan, & Wood, 1998 (randomized controlled trial)

Characteristic	Description
Study citation	Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998). <i>Impacts of dropout prevention programs: Final report. A research report from the School Dropout Demonstration Assistance Program evaluation</i> . Princeton, NJ: Mathematica Policy Research, Inc.
Participants	<p>The <i>Middle College High School</i> study used a randomized controlled trial research design. The original study sample of 516 students was comprised of two cohorts. Cohort 1 included 199 students in the intervention group and 123 students in the control group who applied to attend the alternative high school at the beginning of the 1992–1993 school year. Cohort 2 included 123 students in the intervention group and 71 students in the control group who applied to attend the alternative high school at the beginning of the 1993–1994 school year.</p> <p>Participants were, on average, 18 years old at the time they applied to <i>Middle College High School</i>. About half of the sample was African-American; about one in five was Caucasian; about one in 12 was Hispanic; and slightly fewer than one in four was Asian or other ethnicities. Participants were evenly split between males and females. Most were behind grade level at baseline; two-thirds had had discipline problems at school in the past year; and just over half had dropped out of school in the past.</p> <p>Results summarized here are drawn from a follow-up survey administered two years after random assignment; 244 intervention-group students and 150 control-group students responded, for response rates of 76 and 77 percent, respectively.¹ Researchers compared the baseline characteristics of the two research groups on 13 demographic, socio-economic, and school performance measures. A statistical test of the overall difference between the follow-up survey respondents in the two research groups on the full set of 13 baseline characteristics found that the groups were not significantly different.</p> <p>An additional follow-up survey was conducted three years after random assignment with cohort 1 only. Because of relatively low response rates as well as evidence of substantial intervention-control baseline differences among respondents, these third-year results were not used for the WWC effectiveness rating of <i>Middle College High School</i>. These longer-term results are summarized in Appendices A4.2 and A4.3.</p>
Setting	The <i>Middle College High School</i> study was conducted at an alternative high school on the campus of the Seattle Central Community College in Seattle, Washington.
Intervention	<p>Seattle's <i>Middle College High School</i>, which opened in 1990 and continues to operate, targets students who are close to dropping out or have dropped out in the past. It is run as a collaboration between the Seattle Public Schools and the Seattle Central Community College. It offers students the opportunity to earn a high school diploma in a small alternative school located on the community college campus. The school emphasizes the development of critical thinking skills and focuses on experiential learning, internships, and support services. Services for students include individual counseling, peer support groups, attendance monitoring, and career awareness. In addition, students can take community college courses and use the college's academic and sports facilities. Students are active in school governance and operations. They participate in screening new applicants, running assembly programs, and formulating school policies on attendance, discipline, and other issues.</p> <p>At the time it was evaluated, the school enrolled about 300 students and its core academic curriculum focused on two core modules: (1) math/science and (2) integrated humanities. Each module was taught by a team of two teachers, supported by two or three paid, in-class tutors. Within these team-taught classes, students frequently participated in collaborative learning groups and worked on projects with other students. Each quarter, teachers developed a curriculum around a unifying theme, such as "rights and responsibilities" (Hershey, Adelman, & Murray, 1995).</p>

(continued)

Appendix A1 Study Characteristics: Dynarski, Gleason, Rangarajan, & Wood, 1998 (randomized controlled trial) (continued)

Characteristic	Description
Comparison	The control group did not receive <i>Middle College High School</i> services; however, they were free to participate in other regular and alternative education programs in the community. Most control group students participated in one of these other education options. During the first year after random assignment, control group members reported spending 63% of their time enrolled in a school or GED program, on average, compared with 60% for intervention students. According to student self-reports, over a third of the time control group members spent enrolled in school, they were attending one of Seattle's other alternative high schools.
Primary outcomes and measurement	Two relevant outcomes from the <i>Middle College High School</i> study are included in this review: whether participants dropped out of school and whether they earned a high school diploma or GED certificate. For a more detailed description of these outcome measures, see Appendices A2.1 and A2.2.
Staff/teacher training	<i>Middle College High School</i> teachers were regular high school teachers employed by the Seattle Public Schools. No additional information about their specific training was available.

1. Although the overall second-year response rates for the two cohorts combined were similar for intervention and control group students, there were substantial intervention–control differences in response rates within each cohort, particularly cohort 2. For cohort 1, the response rate for intervention group students was 82%, and for control group students, it was 75%. For cohort 2, the response rate for intervention group students was 65%, and for control group students, it was 82%.

Appendix A2.1 Outcome measures for the staying in school domain

Outcome measure	Description
Dropped out	Percentage of students who dropped out of school by the end of the second follow-up year. These self-reported data were collected from follow-up surveys.

Appendix A2.2 Outcome measures for the completing school domain

Outcome measure	Description
Earned a high school diploma or GED	Percentage of students who received a high school diploma or GED certificate by the end of the second follow-up year. These self-reported data were collected from follow-up surveys.

Appendix A3.1 Summary of study findings included in the rating for the staying in school domain¹

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study			WWC calculations		
			Mean outcome		Mean difference ² (Middle College High School – comparison)	Effect size ³	Statistical significance ⁴ (at $\alpha = 0.05$)	Improvement index ⁵
			Middle College High School group	Comparison group				
Dynarski et al., 1998 (randomized controlled trial)⁶								
Dropped out (%)	Cohorts 1 and 2	394	36	33	–3	–0.08	ns	–3
Domain average for staying in school						–0.08	ns	–3

ns = not statistically significant

1. This appendix reports second year follow-up findings considered for the effectiveness rating and the average improvement index for the staying in school domain. Third-year follow-up findings, available for cohort 1 only, are not included in these ratings but are reported in Appendix A4.2.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. For the “dropped out” outcome, signs were reversed on the mean difference, effect size, and improvement index, since a reduction in dropping out is a favorable outcome. Means from Dynarski et al. (1998) are estimated using regression models that control for baseline characteristics.
3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
6. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Dynarski et al. (1998), no corrections for clustering or multiple comparisons were needed.

Appendix A3.2 Summary of study findings included in the rating for the completing school domain¹

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study			WWC calculations		
			Mean outcome		Mean difference ² <i>Middle College High School</i> – comparison)	Effect size ³	Statistical significance ⁴ (at $\alpha = 0.05$)	Improvement index ⁵
			<i>Middle College High School</i> group	Comparison group				
Dynarski et al., 1998 (randomized controlled trial)⁶								
Earned a high school diploma or GED (%)	Cohorts 1 and 2	394	40	38	2	0.05	ns	+2
Domain average for completing school						0.05	ns	+2

ns = not statistically significant

1. This appendix reports second year follow-up findings considered for the effectiveness rating and the average improvement index for the completing school domain. Third-year follow-up findings, available for cohort 1 only, are not included in these ratings but are reported in Appendix A4.3.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Means from Dynarski et al. (1998) are estimated using regression models that control for baseline characteristics.
3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
6. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Dynarski et al. (1998), no corrections for clustering or multiple comparisons were needed.

Appendix A4.1 Summary of additional findings for the completing school domain¹

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study					
			Mean outcome		WWC calculations			
			Middle College High School group	Comparison group	Mean difference ² (Middle College High School – comparison)	Effect size ³	Statistical significance ⁴ (at $\alpha = 0.05$)	Improvement index ⁵
Dynarski et al., 1998 (randomized controlled trial)⁶								
Earned a high school diploma (%)	Cohorts 1 and 2	394	21	18	3	0.12	ns	+5
Earned a GED certificate (%)	Cohorts 1 and 2	394	18	20	–2	–0.08	ns	–3

ns = not statistically significant

1. This appendix presents findings for the intervention's separate effects on high school diploma receipt and on GED certificate receipt. The intervention's effect on the combined measure of high school diploma or GED receipt was used for determining the effectiveness rating and is presented in Appendix A3.2.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Means from Dynarski et al. (1998) are estimated using regression models that control for baseline characteristics.
3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
6. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Dynarski et al. (1998), no corrections for clustering were needed.

Appendix A4.2 Summary of longer-term subgroup findings for the staying in school domain¹

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study					
			Mean outcome		WWC calculations			
			Middle College High School group	Comparison group	Mean difference ² (Middle College High School – comparison)	Effect size ³	Statistical significance ⁴ (at $\alpha = 0.05$)	Improvement index ⁵
Dynarski et al., 1998 (randomized controlled trial)⁶								
Dropped out at end of year 3 (%)	Cohort 1 only	217	31	38	7	0.19	ns	+7

ns = not statistically significant

1. This appendix presents the third-year follow-up findings for measures in the staying in school domain. These findings were not used for intervention rating purposes because the third-year survey was administered only to cohort 1 and because the survey had a relatively low response rate (67%). Moreover, substantial baseline differences existed between the intervention and control group members who responded to the third-year survey. The intervention's effect on staying in school was rated based on the second-year follow-up findings presented in Appendix A3.1.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. For the “dropped out” outcome, signs were reversed on the mean difference, effect size, and improvement index, since a reduction in dropping out is a favorable outcome. Means from Dynarski et al. (1998) are estimated using regression models that control for baseline characteristics.
3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
6. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Dynarski et al. (1998), no corrections for clustering were needed.

Appendix A4.3 Summary of longer-term follow-up findings for the completing school domain¹

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study					
			Mean outcome		WWC calculations			
			Middle College High School group	Comparison group	Mean difference ² (Middle College High School – comparison)	Effect size ³	Statistical significance ⁴ (at $\alpha = 0.05$)	Improvement index ⁵
Dynarski et al., 1998 (randomized controlled trial)⁶								
Earned a high school diploma by end of year 3 (%)	Cohort 1 only	217	31	23	8	0.25	ns	+10
Earned a GED certificate by end of year 3 (%)	Cohort 1 only	217	24	37	-13	-0.38	Statistically significant	-15
Earned a high school diploma or GED by end of year 3 (%)	Cohort 1 only	217	55	61	-6	-0.15	ns	-6

ns = not statistically significant

1. This appendix presents the third-year follow-up findings for measures in the completing school domain. These findings were not used for intervention rating purposes because the third-year survey was administered only to cohort 1 and because the survey had a relatively low response rate (67%). Moreover, substantial baseline differences existed between the intervention and control group members who responded to the third-year survey. The intervention's effect on completing school was rated based on the second-year follow-up findings presented in Appendix A3.2.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Means from Dynarski et al. (1998) are estimated using regression models that control for baseline characteristics.
3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between -50 and +50, with positive numbers denoting results favorable to the intervention group.
6. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Dynarski et al. (1998), no corrections for clustering were needed.

Appendix A5.1 Middle College High School rating for the staying in school domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.¹ For the outcome domain of staying in school, the WWC rated *Middle College High School* as having no discernible effects.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: No studies showing a statistically significant or substantively important effect, either *positive* or *negative*.

Met. No studies showed statistically significant or substantively important positive or negative effects.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a *strong* design.

Not met. No studies showed statistically significant or substantively important positive effects.

AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

Met. No studies showed statistically significant or substantively important negative effects.

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

Not met. No studies showed statistically significant or substantively important positive effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

Met. No studies showed statistically significant or substantively important negative effects.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

Not met. No studies showed statistically significant or substantively important effects, either positive or negative.

OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

Not met. No studies showed statistically significant or substantively important effects.

(continued)

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. For a complete description, see the WWC Intervention Rating Scheme.

Appendix A5.1 **Middle College High School rating for the staying in school domain** *(continued)*

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *negative* effect.

Not met. No studies showed statistically significant or substantively important negative effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

Met. No studies showed statistically significant or substantively important positive effects.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a *strong* design.

Not met. No studies showed statistically significant or substantively important negative effects.

AND

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

Met. No studies showed statistically significant or substantively important positive effects.

Appendix A5.2 Middle College High School rating for the completing school domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.¹ For the outcome domain of completing school, the WWC rated *Middle College High School* as having no discernible effects.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: No studies showing a statistically significant or substantively important effect, either *positive* or *negative*.

Met. No studies showed statistically significant or substantively important positive or negative effects.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a *strong* design.

Not met. No studies showed statistically significant or substantively important positive effects.

AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

Met. No studies showed statistically significant or substantively important negative effects.

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

Not met. No studies showed statistically significant or substantively important positive effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

Not met. No studies showed statistically significant or substantively important negative effects.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

Not met. No studies showed statistically significant or substantively important effects, either positive or negative.

OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

Not met. No studies showed statistically significant or substantively important effects.

(continued)

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. For a complete description, see the WWC Intervention Rating Scheme.

Appendix A5.2 Middle College High School rating for the completing school domain (continued)

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *negative* effect.

Not met. No studies showed statistically significant or substantively important negative effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

Met. No studies showed statistically significant or substantively important positive effects.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a *strong* design.

Not met. No studies showed statistically significant or substantively important negative effects.

AND

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

Met. No studies showed statistically significant or substantively important positive effects.

Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence ¹
		Schools	Students	
Staying in school	1	>1	394	Small
Progressing in school	na	na	na	na
Completing school	1	>1	394	Small

na = not applicable/not studied

1. A rating of “medium to large” requires at least two studies and two schools across studies in one domain and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is “small.”