What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Dropout Prevention September 2010

National Guard Youth ChalleNGe Program

Program Description[1]

The *National Guard Youth ChalleNGe Program* is a residential education and training program designed for youth ages 16 to 18 who have dropped out of or been expelled from high school. During the 22-week residential period, participants are offered GED preparation classes and other program services intended to promote positive youth development, such as leadership, job skills, and service to the community. The residential period is quasi-military (youth live in barracks, wear uniforms, and experience military-style discipline), but there are no requirements for military service. After the residential period, trainees participate in a one-year structured mentoring program. Trainees select their own mentors who are then screened and trained by the program.

Research[2]

One study of the *National Guard Youth ChalleNGe Program* that falls within the scope of the Dropout Prevention review protocol meets What Works Clearinghouse (WWC) evidence standards, and no studies meet WWC evidence standards with reservations. This study included 1,196 youth in 10 states.[3]

Based on this one study, the WWC considers the extent of evidence for the *National Guard Youth ChalleNGe Program* on at-risk youth to be small for the completing school domain. The study that meets WWC evidence standards does not examine the effectiveness of the *National Guard Youth ChalleNGe Program* in the staying in school or progressing in school domains for at-risk youth.

Effectiveness

The *National Guard Youth ChalleNGe Program* was found to have potentially positive effects on completing school for at-risk youth.

	Staying in school	*Progressing in school*	*Completing school*
Rating of effectiveness	na	na	Potentially positive effects
Improvement index	na	na	+22 percentile points

na = not applicable

1. The descriptive information for this program was obtained from publicly available sources: the program's website (http://www.ngycp.org, downloaded June 2010) and Millenky, Bloom, and Dillon (2010). The WWC requests developers to review the program description sections for accuracy from their perspective. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review. The literature search reflects documents publicly available by May 2010.
2. The studies in this report were reviewed using WWC Evidence Standards, Version 2.0 (see the WWC Procedures and Standards Handbook, Chapter III), as described in protocol Version 2.0.
3. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

Additional program information

Developer and contact

The *National Guard Youth ChalleNGe Program* was developed by the National Guard Bureau in the U.S. Department of Defense; the pilot of the program was authorized by Congress in 1993. The national contact for the program is Joe Padilla. Address: Office of Athletics and Youth Development, Jefferson Plaza 1, Room 2456, 1411 Jefferson Davis Highway, Arlington, VA 22202-3231. Email: info@ngyf.org. Web: http://www.ngycp.org.

Scope of use

Over 92,000 youths graduated from the *National Guard Youth ChalleNGe Program* as of December 2009. The program is offered in Puerto Rico and the following 27 states: Alabama, Alaska, Arizona, Arkansas, California, Florida, Georgia, Hawaii, Illinois, Indiana, Kentucky, Louisiana, Maryland, Michigan, Mississippi, Montana, New Jersey, New Mexico, North Carolina, Oklahoma, Oregon, South Carolina, Texas, Virginia, West Virginia, Wisconsin, and Wyoming.

Description of intervention

The *National Guard Youth ChalleNGe Program* is offered to youth ages 16 to 18 who have dropped out of or been expelled from high school, and who are unemployed, not heavily involved with the criminal justice system (not currently on parole or probation for non-juvenile offenses, not serving time or awaiting sentencing, not convicted of a felony), drug free, and legal residents of the state in which the program is offered.

Each *National Guard Youth ChalleNGe Program* site accepts an average of 100 trainees in a cohort. The program consists of three phases. The first phase, the pre-*ChalleNGe* phase, is a two-week residential introduction to the program's rules and expectations. If trainees complete this phase, they move on to the 20-week residential phase. Trainees live on an active or closed National Guard base, a National Guard training center, or a school campus. This phase is highly structured and quasi-military. Trainees are called "cadets," live in barracks, have their hair cut short, wear uniforms, experience military-style discipline, and are given a highly regimented schedule. The program focuses on eight components: leadership, responsible citizenship, service to community, life-coping skills, physical fitness, health and hygiene, job skills, and academic excellence. The largest share of the day is spent on education, helping trainees prepare for the GED exam, or, in some cases, a high school diploma. The last phase, the post-residential phase, lasts one year. This phase involves structured mentoring with a mentor of the trainee's choosing. The mentor is screened and trained by the *National Guard Youth ChalleNGe Program*. The mentor is supposed to meet with the trainee a minimum of four hours per month during the post-residential phase. *National Guard Youth ChalleNGe Program* staff maintain at least monthly contact with the trainee and mentor during this phase.

Cost

The *National Guard Youth ChalleNGe Program* costs about $14,000 per individual (Bloom, Gardenhire-Crooks, & Mandsager, 2009). Since 1998, the federal government has paid 60% of the cost and states have paid 40%. In 2009, the U.S. Congress passed legislation that increased the federal share of program costs that can be paid by the U.S. Department of Defense to 75%. The program is offered free of charge to participants.

Research

Fourteen studies reviewed by the WWC investigated the effects of the *National Guard Youth ChalleNGe Program*. One study (Millenky, Bloom, & Dillon, 2010) is a randomized controlled trial that meets WWC evidence standards. The remaining 13 studies do not meet either WWC evidence standards or eligibility screens.

Millenky et al. (2010) used a randomized controlled trial to examine the effects of the *National Guard Youth ChalleNGe Program* in 10 sites located in 10 states. The study included 3,074 youth who were randomly assigned to participate in the *National Guard Youth ChalleNGe Program* or to a control group that was not eligible to receive services. A subset of 1,508 youth were

Research *(continued)*

randomly selected to be given the 21-month follow-up survey. The analysis sample included 1,196 youths who completed the follow-up survey.

Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or medium to large (see the WWC Procedures and Standards Handbook, Appendix G). The extent of evidence takes into account the number of studies and the total sample size across the studies that meet WWC evidence standards with or without reservations.[4]

The WWC considers the extent of evidence for the *National Guard Youth ChalleNGe Program* on at-risk youth to be small for the completing school domain. The study that meets WWC evidence standards does not examine the effectiveness of the *National Guard Youth ChalleNGe Program* in the staying in school and progressing in school domains for at-risk youth.

Effectiveness

Findings

The WWC review of interventions for Dropout Prevention addresses student outcomes in three domains: staying in school, progressing in school, and completing school. The study included in this report covers one domain: completing school. The findings below present the authors' estimates and WWC-calculated estimates of the size and the statistical significance of the effects of the *National Guard Youth ChalleNGe Program* on at-risk youth.[5]

Completing school. The Millenky et al. (2010) study found, and the WWC confirmed, a statistically significant positive difference between the *National Guard Youth ChalleNGe Program* group and the control group in their self-reported rates of having received a high school diploma or GED 21 months after random assignment. At follow-up, 61% of program group members reported having a diploma or GED, compared with 36% of control group members. In addition, the effect size of this difference was large enough to be considered substantively important according to WWC criteria (at least 0.25).

Rating of effectiveness

The WWC rates the effects of an intervention in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative. The rating of effectiveness takes into account four factors: the quality of the research design, the statistical significance of the findings, the size of the difference between participants in the intervention and the comparison conditions, and the consistency in findings across studies (see the WWC Procedures and Standards Handbook, Appendix E).

The WWC found the *National Guard Youth ChalleNGe Program* to have potentially positive effects on completing school for at-risk youth

Improvement index

The WWC computes an improvement index for each individual finding. In addition, within each outcome domain, the WWC computes an average improvement index for each study and an average improvement index across studies (see WWC Procedures and Standards Handbook, Appendix F). The improvement index represents the difference between the percentile rank of the average student in the intervention condition and the percentile rank of the average student in the comparison condition. Unlike the rating of effectiveness, the improvement

4. The extent of evidence categorization was developed to tell readers how much evidence was used to determine the intervention rating, focusing on the number and size of studies. Additional factors associated with a related concept (external validity, such as the students' demographics and the types of settings in which studies took place) are not taken into account for the categorization. Information about how the extent of evidence rating was determined for the *National Guard Youth ChalleNGe Program* is in Appendix A6.
5. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Millenky et al. (2010), no corrections for clustering or multiple comparisons were needed.

The WWC found the *National Guard Youth ChalleNGe Program* to have potentially positive effects on completing school for at-risk youth *(continued)*

index is entirely based on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analysis. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results for the intervention group.

The improvement index for completing school is +22 percentile points based on one finding in one study.

Summary

The WWC reviewed 14 studies on the *National Guard Youth ChalleNGe Program*. One of these studies meets WWC evidence standards; the remaining 13 studies do not meet either WWC evidence standards or eligibility screens. Based on the one study, the WWC found potentially positive effects on completing school for at-risk youth. The conclusions presented in this report may change as new research emerges.

References

Meets WWC evidence standards

Millenky, M., Bloom, D., & Dillon, C. (2010). *Making the transition: Interim results of the* National Guard Youth ChalleNGe *evaluation.* New York: MDRC.

Additional source:

Bloom, D., Gardenhire-Crooks, A., & Mandsager, C. (2009). *Reengaging high school dropouts: Early results of the* National Guard Youth ChalleNGe Program *evaluation.* New York: MDRC.

Studies that fall outside the Dropout Prevention review protocol or do not meet WWC evidence standards

Fleming, A. R. (2005). *GED acquisition rates from an at-risk youth program's curriculum as influenced by three criteria: Multiple intelligence usage, class size, and its mentorship program.* Unpublished doctoral dissertation, Argosy University, Chicago, IL. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Garcia, F. E., Gasch, J. L., Wenger, J. W., & Ray, B. D. (2001). *Evaluation of the pilot program for home school and* ChalleNGe *program recruits.* Ft. Belvoir, VA: Defense Technical Information Center. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Handy, R. A. (2000). The quiet warriors: A qualitative study of high school dropouts, the Maryland National Guard, and Freestate *ChalleNGe* Academy. *Dissertation Abstracts International, 61*(12A), 212–4724. The study is ineligible for review because it does not use a comparison group design or a single-case design.

National Guard. (2003). National Guard Youth ChalleNGe Program: *2003 performance and accountability highlights.* Chantilly, VA: AOC Solutions, Inc. The study is ineligible for review because it does not use a comparison group design or a single-case design.

National Guard. (2004). National Guard Youth ChalleNGe Program: *2004 performance and accountability highlights.* Chantilly, VA: AOC Solutions, Inc. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

National Guard. (2005). National Guard Youth ChalleNGe Program: *2005 performance and accountability highlights.* Chantilly, VA: AOC Solutions, Inc. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

National Guard. (2006). National Guard Youth ChalleNGe Program: *2006 performance and accountability highlights.* Chantilly, VA: AOC Solutions, Inc. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

National Guard. (2007). National Guard Youth ChalleNGe Program: *2007 performance and accountability highlights.* Chantilly, VA: AOC Solutions, Inc. The study is ineligible for

References

review because it does not use a comparison group design or a single-case design.

National Guard. (2008). National Guard Youth ChalleNGe Program: *2008 performance and accountability highlights.* Chantilly, VA: AOC Solutions, Inc. The study is ineligible for review because it does not use a comparison group design or a single-case design.

National Guard. (2009). National Guard Youth ChalleNGe Program: *2009 performance and accountability highlights.* Chantilly, VA: AOC Solutions, Inc. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Price, H. B. (2007). *Demilitarizing what the Pentagon knows about developing young people: A new paradigm for educating students who are struggling in school and life.* Washington, DC: Brookings Institution, Center on Children and Families. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Smith, W. B., Jr. (2003). Efficacy of a post-secondary environmental science education program on the attitude toward science of a group of Mississippi *National Guard Youth ChalleNGe Program* students. *Dissertation Abstracts International, 64*(03A), 65–848. The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.

Wenger, J. W., & Hodari, A. K. (2004). *Final analysis of evaluation of homeschool and* ChalleNGe *program recruits.* Ft. Belvoir, VA: Defense Technical Information Center. The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.

Appendix

Appendix A1 Study characteristics: Millenky, Bloom, and Dillon, 2010

Characteristic	Description
Study citation	Millenky, M., Bloom, D., & Dillon, C. (2010). *Making the transition: Interim results of the* National Guard Youth ChalleNGe *evaluation.* New York: MDRC.
Participants	Between 2005 and 2007, a total of 3,074 high school dropouts between the ages of 16 and 18 were randomly assigned to the *National Guard Youth ChalleNGe Program* in 10 states; 2,320 to the intervention group and 754 to the control group. There was a single site in each state. A random subsample of 1,508 participants (916 in the intervention group and 592 in the control group) was selected to be given the 21-month follow-up survey. The analysis sample included 1,196 youths (736 in the intervention group and 460 in the control group) who responded to the follow-up survey; 80% of the intervention group and 78% of the control group responded to the survey. Most sample members (84%) were male. They ranged in age from 16 to 18 years old at program entry. The sample was racially and ethnically diverse: 41% were Caucasian; 40% were African-American; 14% were Hispanic; and 5% were from other racial and ethnic groups. Sample members had performed poorly in school before entering the program. About half reported receiving mostly Ds and Fs before dropping out.
Setting	The study took place in 10 sites in California, Florida, Georgia, Illinois, Michigan, Mississippi, New Mexico, North Carolina, Texas, and Wisconsin.
Intervention	The *National Guard Youth ChalleNGe Program* attempts to promote positive youth development for high school dropouts through a residential education and training program followed by a structured mentoring program. The intervention consisted of three phases. Phase 1: Two weeks of program orientation and physical and psychological assessment in a residential, quasi-military setting. Phase 2: Twenty weeks of education and training in a residential, quasi-military setting. During this phase, the majority of participants' time was spent on educational activities, including work toward a GED or high school diploma. Education and training covered the following eight areas: leadership, responsible citizenship, service to community, life-coping skills, physical fitness, health and hygiene, job skills, and academic excellence. Phase 3: One year postresidential phase. Participants worked with program staff to arrange a postresidential placement in employment, continued education, or military service. In addition to participating in the placement activity, each participant was supposed to maintain monthly contact with a mentor, who was chosen by the participant and trained by *ChalleNGe* program staff.
Comparison	Control group members were not eligible to receive *National Guard Youth ChalleNGe Program* services but could receive other services available in the community. Contrary to the study design, eight control group members enrolled in the program (1% of the full control group). These enrollees were still assigned to the control group in all analyses.
Primary outcomes and measurement	The relevant study outcome included in this review is whether students earned a high school diploma or GED, based on student follow-up interviews. For a more detailed description of this outcome measure, see Appendix A2. The study also examined a number of other outcomes that are not within the scope of the Dropout Prevention protocol.
Staff/teacher training	Program staff included team leaders who directly supervised the students, teachers who provided classroom instruction, and counselors who provided individual, group, and career counseling. It was common for staff members, particularly team leaders, to be National Guard members or to have military experience. Teachers often came to the program from the local school district or from community colleges. Counselors typically had bachelor's or advanced degrees in psychology, social work, or other relevant fields.

Appendix A2 Outcome measure for the completing school domain

Outcome measure	Description
Earned a diploma or GED certificate	This binary outcome was measured based on a follow-up survey conducted 21 months after random assignment. As a part of this survey, youth reported whether they had earned a high school diploma or a GED certificate.

Appendix A3 Summary of study findings included in the rating for the completing school domain[1]

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome — *ChalleNGe* group	Authors' findings from the study: Mean outcome — Comparison group	WWC calculations: Mean difference[2] (*ChalleNGe* – comparison)	WWC calculations: Effect size[3]	WWC calculations: Statistical significance[4] (at $\alpha = 0.05$)	WWC calculations: Improvement index[5]
Millenky, Bloom, & Dillon, 2010[6]								
Earned a Diploma or GED Certificate	Survey respondents	1,196	0.61	0.36	0.24	0.60	Statistically significant	+22
Domain average for completing school[7]						**0.60**	**Statistically significant**	**+22**

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the completing school domain. Subgroup findings from the same study are not included in these ratings but are reported in Appendices A4.1 and A4.2.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
3. For an explanation of the effect-size calculation, see WWC Procedures and Standards Handbook, Appendix B. In the case of binary outcome variables, effect sizes are calculated using the Cox effect-size index. The Cox index is based on logged odds ratios; therefore, standard deviations are not needed for the calculation of effect sizes for binary outcome measures, such as whether youth have earned a high school diploma or GED.
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 favorable results for the intervention group.
6. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. No corrections for clustering or multiple comparisons were needed.
7. This row provides the study average, which in this case is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A4.1 Summary of study findings for separate outcomes for the completing school domain[1]

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome: *ChalleNGe* group	Authors' findings from the study: Mean outcome: Comparison group	WWC calculations: Mean difference[2] (*ChalleNGe* – comparison)	WWC calculations: Effect size[3]	WWC calculations: Statistical significance[4] (at $\alpha = 0.05$)	WWC calculations: Improvement index[5]
Millenky, Bloom, & Dillon, 2010[6]								
Earned a Diploma	Survey respondents	1,196	0.22	0.16	0.06	0.22	Statistically significant	+9
Earned a GED Certificate	Survey respondents	1,196	0.48	0.22	0.26	0.73	Statistically significant	+27

1. This appendix presents findings for outcome measures that disaggregate whether a respondent earned a high school diploma or a GED certificate. The combined outcome measures were used for rating purposes and are presented in Appendix A3.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
3. For an explanation of the effect-size calculation, see WWC Procedures and Standards Handbook, Appendix B. In the case of binary outcome variables, effect sizes are calculated using the Cox effect-size index. The Cox index is based on logged odds ratios; therefore, standard deviations are not needed for the calculation of effect sizes for binary outcome measures, such as whether youth have earned a high school diploma or GED.
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, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. No corrections for clustering or multiple comparisons were needed.

Appendix A4.2 Summary of under age 17 and age 17 and older subgroup findings for the completing school domain[1]

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome: *ChalleNGe* group	Authors' findings from the study: Mean outcome: Comparison group	WWC calculations: Mean difference[2] (*ChalleNGe* – comparison)	WWC calculations: Effect size[3]	WWC calculations: Statistical significance[4] (at $\alpha = 0.05$)	WWC calculations: Improvement index[5]
Millenky, Bloom, & Dillon, 2010[6]								
Earned a Diploma or GED Certificate	Under age 17	435	0.58	0.27	0.30	0.78	Statistically significant	+28
Earned a Diploma or GED Certificate	Age 17 and older	761	0.62	0.40	0.22	0.53	Statistically significant	+20

1. This appendix presents under 17 and 17 and older subgroup findings for measures that fall in the completing school domain. Total group outcomes were used for rating purposes and are presented in Appendix A3.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
3. For an explanation of the effect-size calculation, see WWC Procedures and Standards Handbook, Appendix B. In the case of binary outcome variables, effect sizes are calculated using the Cox effect-size index. The Cox index is based on logged odds ratios; therefore, standard deviations are not needed for the calculation of effect sizes for binary outcome measures, such as whether youth have earned a high school diploma or GED.
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, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. No corrections for clustering or multiple comparisons were needed.

Appendix A4.3 Summary of low-achieving and higher-achieving subgroup findings for the completing school domain[1]

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome: *ChalleNGe* group	Authors' findings from the study: Mean outcome: Comparison group	WWC calculations: Mean difference[2] (*ChalleNGe* – comparison)	WWC calculations: Effect size[3]	WWC calculations: Statistical significance[4] (at $\alpha = 0.05$)	WWC calculations: Improvement index[5]
Millenky, Bloom, & Dillon, 2010[6]								
Earned a Diploma or GED Certificate	Higher-achieving	621	0.65	0.38	0.28	0.68	Statistically significant	+25
Earned a Diploma or GED Certificate	Low-achieving	540	0.55	0.33	0.22	0.55	Statistically significant	+21

1. This appendix presents low-achieving and higher-achieving subgroup findings for measures that fall in the completing school domain. Low-achieving is a binary indicator for students who earned mostly Ds and Fs when in school. Higher-achieving is a binary indicator for students who earned mostly better than Ds and Fs when in school. Total group outcomes were used for rating purposes and are presented in Appendix A3.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
3. For an explanation of the effect-size calculation, see WWC Procedures and Standards Handbook, Appendix B. In the case of binary outcome variables, effect sizes are calculated using the Cox effect-size index. The Cox index is based on logged odds ratios; therefore, standard deviations are not needed for the calculation of effect sizes for binary outcome measures, such as whether youth have earned a high school diploma or GED.
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, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. No corrections for clustering or multiple comparisons were needed.

Appendix A5 *National Guard Youth ChalleNGe Program* 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 the *National Guard Youth ChalleNGe Program* as having potentially positive effects for at-risk youth. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, negative effects) were not considered, as the *National Guard Youth ChalleNGe Program* was assigned the highest applicable rating.

Rating received

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.

  Met. One study of the *National Guard Youth ChalleNGe Program* reported a statistically significant and substantively important positive effect in this domain.

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 found statistically significant or substantively important negative effects or indeterminate effects in this domain.

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. Only one study of the *National Guard Youth ChalleNGe Program* reported a statistically significant and substantively important positive effect in this domain.

AND

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

  Met. No studies found statistically significant or substantively important negative effects in this domain.

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 Procedures and Standards Handbook, Appendix E.

Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[1]
		Schools	Students	
Staying in school	na	na	na	na
Progressing in school	na	na	na	na
Completing school	1	>2	1,196	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." For more details on the extent of evidence categorization, see the WWC Procedures and Standards Handbook, Appendix G.