

What Works Clearinghouse



Science

May 2012

Astronomy Resources for Intercurricular Elementary Science (ARIES): Exploring Motion and Forces

Report Contents	
Overview	p. 1
Program Information	p. 2
Research Summary	p. 3
Effectiveness Summary	p. 4
References	p. 5
Research Details for Each Study	p. 6
Outcomes Measures for Each Domain	p. 8
Findings Included in the Rating for Each Outcome Domain	p. 9
Supplemental Findings for Each Outcome Domain	p. 10
Endnotes	p. 11
Rating Criteria	p. 12
Glossary of Terms	p. 13

Program Description¹

*ARIES: Exploring Motion and Forces*² is a physical science curriculum for students in grades 5–8 that employs 18 inquiry-centered, hands-on lessons called “explorations.” The curriculum draws upon students’ curiosity to explore phenomena, allowing for a discovery-based learning process. Group-centered lab work is designed to help students build an understanding of inertia, friction, gravity, speed, and acceleration. Students examine their prior ideas about the phenomena, formulate questions, build and use an apparatus to observe natural phenomena, make predictions, and gather data through structured experiments. *Exploring Motion and Forces* is part of the *ARIES* sequence of eight physical science units. The *ARIES* sequences can be used together for an overall curriculum or independently.

Research³

One study of *ARIES: Exploring Motion and Forces* that falls within the scope of the Science review protocol meets What Works Clearinghouse (WWC) evidence standards. This study included approximately 6,000 sixth-grade students from 20 schools in Maryland. Based on this study, the WWC considers the extent of evidence for *ARIES: Exploring Motion and Forces* on middle school students to be small for the general science achievement domain, the only domain identified by the review protocol.

Effectiveness

ARIES: Exploring Motion and Forces was found to have no discernible effects on general science achievement for middle school students.

Table 1. Summary of findings⁴

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of studies	Number of students	Extent of evidence
		Average	Range			
General science achievement	No discernible effects	+4 +4	–30 to +99	1	6,058	Small

Program Information

Background

ARIES: Exploring Motion and Forces was developed by the Science Education Department at the Harvard-Smithsonian Center for Astrophysics. The curriculum is available from the Charlesbridge Publishing Company. Address: 85 Main St., Watertown, MA 02472. Email: orders@charlesbridge.com. Web: <http://www.charlesbridge.com/school/html/aries.html>. Telephone: (800) 225-3214; (617) 926-0329. Fax: (800) 926-5775.

Program details

The *ARIES* curriculum includes eight self-contained units that cover key areas of physical science:

- *Exploring Time: Sundials, Waterclocks, and Pendulums*
- *Exploring Light and Color: Filters, Lenses, and Cameras*
- *Exploring the Earth in Motion: Daylight, Sun, and Shadow Patterns*
- *Exploring Waves: Water, Vibrations, and Sound*
- *Exploring Energy: Waterwheels, Windmills, and Sunlight*
- *Exploring Motion and Forces: Speed, Acceleration, and Friction*
- *Exploring Space: Sun, Moon, and Stars*
- *Exploring Navigation: Location, Direction, and Latitude*

This report examines *ARIES: Exploring Motion and Forces*, a six-week physical science curriculum designed for use in grades 5–8. The curriculum includes 18 “explorations” designed to help students understand forces (push and pull) and motions; inertia and friction; gravity; speed and acceleration; motions on horizontal surfaces and inclined planes; and falling, sliding, rolling, and wheeled motion. Each exploration asks students to examine their own ideas about motion or force by answering a series of questions about familiar concepts. Students then write their own questions about forces and motion and work in groups to conduct experiments, collect and summarize data, and discuss their results.

In this unit, students are asked to consider what causes an object to start moving, what affects the speed of an object, and how the object moves when different forces are used. Several activities show how pushes or pulls are examples of forces. Students slide several objects across a horizontal surface and throw a coin up in the air to observe examples of forces changing the speed and direction of objects. Students are asked questions such as “Why do you think the object started moving?”, “If your object is in constant motion, why do you think it keeps moving?”, “What evidence do you look for?”, and “How is the motion of a coin thrown into the air similar to the motion of a coin moving on a horizontal surface?”

Through Project SEDNet (<http://www.christa.org/SEDNet/index.htm>), *ARIES* offers one-day workshops for grade 3–8 teachers to learn how to implement the program through modeling and role playing.

Cost

The curriculum materials consist of a teacher manual and a student science journal. There is no traditional student textbook, but a substantial amount of lab materials is required to conduct the explorations (sliding disks, ramps, marbles, bells, rolling carts).

Educators may buy all of the necessary materials in a prepared kit for \$699 (which includes one teacher manual, 30 science journals, and one apparatus bin), purchase student lab materials (apparatus bin) separately for \$449 (rates as of March 2012), or make their own sets of student lab materials. Additional information can be found on the Charlesbridge Publishing Company website (<http://www.charlesbridge.com>).

Research Summary

Two studies reviewed by the WWC investigated the effects of *ARIES: Exploring Motion and Forces* on middle school students. One study (Pyke, Lynch, Kuipers, Szesze, & Watson, 2004, 2005, 2006) is a randomized controlled trial that meets WWC evidence standards. The remaining study does not meet WWC evidence standards. (See references beginning on p. 5 for citations for both studies.)

Table 2. Scope of reviewed research

Grade	6
Delivery method	Small group/Whole class
Program type	Curriculum
Studies reviewed	2
Meets WWC standards	1 study
Meets WWC standards with reservations	0 studies

Summary of studies meeting WWC evidence standards without reservations

Pyke et al. (2004, 2005, 2006) conducted a randomized controlled trial that examined the effects of *ARIES: Exploring Motion and Forces* on middle school students' knowledge and understanding of physical science. Three separate cohorts were formed, and the results were presented in three research reports. These reports have been combined into a single study for this review. The total study sample included 6,058 sixth-grade students attending 20 middle schools in Maryland.

The study used a school-level random assignment design that took place in a three-step process. First, Pyke et al. (2004) grouped all district schools into five school profile categories, each having similar demographic and achievement characteristics. Next, the authors selected one pair of schools from each of the five school profile categories. Finally, one school from each pair was randomly assigned either to implement the intervention or to serve as a comparison school. This process was initially performed to identify 10 study schools for Cohort 1. The same 10 schools identified for Cohort 1 were also analyzed for Cohort 2. The randomization process was repeated for Cohort 3, with the exception that the original 10 schools chosen for Cohort 1 were removed from the set of potential schools. This resulted in an additional 10 schools that were included in the study, for a total of 20 middle schools.

Cohort 1 was formed in the 2003–04 school year and consisted of 1,266 sixth-grade students who received *ARIES: Exploring Motion and Forces* and 1,115 sixth-grade students in the comparison group. In Pyke et al. (2005), Cohort 2 was formed in the 2004–05 school year and consisted of 910 sixth-grade students who received the intervention and 1,005 sixth-grade students in the comparison group. Note that the WWC considers Cohort 2 to be part of the original design.⁵ In Pyke et al. (2006), Cohort 3 was formed in the 2005–06 school year and consisted of 902 sixth-grade students who received the intervention and 860 sixth-grade students in the comparison group. All comparison group students were taught using the regular science curriculum. The study reported student outcomes after approximately six weeks of program implementation. The WWC based its effectiveness rating on findings from the three cohorts of students.⁶

Summary of studies meeting WWC evidence standards with reservations

No studies of *ARIES: Exploring Motion and Forces* meet WWC evidence standards with reservations.

Effectiveness Summary

The WWC review of interventions for Science addresses student outcomes in one domain: general science achievement. The domain includes three outcome constructs: life science, earth/space science, and physical science. The study that contributes to the findings in this report covers one construct: physical science. The findings below present the authors' estimates and WWC-calculated estimates of the size and the statistical significance of the effects of *ARIES: Exploring Motion and Forces* on middle school students.⁷ For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 12.

Summary of effectiveness for the general science achievement domain

One study reported findings in the general science achievement domain.

Pyke et al. (2004) reported statistically significant positive effects of *ARIES: Exploring Motion and Forces* on the Motion and Forces Assessment for Cohort 1, and Pyke et al. (2006) found statistically significant positive effects for Cohort 3. For Cohort 2 students, the authors did not find statistically significant effects of *ARIES: Exploring Motion and Forces* on the Motion and Forces Assessment. According to WWC calculations, the effects for Cohort 1, Cohort 2, and Cohort 3 were not statistically significant when adjusted for clustering. The average effect across the three cohorts was not large enough to be considered substantively important according to WWC criteria (i.e., an effect size of at least 0.25).⁸

Thus, for the general science achievement domain, one study showed indeterminate effects. This results in a rating of no discernible effects, with a small extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the general science achievement domain

Rating of effectiveness	Criteria met
No discernible effects <i>No affirmative evidence of effects.</i>	The review of <i>ARIES: Exploring Motion and Forces</i> for the general science achievement domain had one study showing indeterminate effects.
Extent of evidence	Criteria met
Small	The review of <i>ARIES: Motion and Forces</i> for the general science achievement domain was based on one study that included 20 schools and 6,058 students.

References

Study that meets WWC evidence standards without reservations

Pyke, C., Lynch, S., Kuipers, J., Szesze, M., & Watson, W. (2004). *Implementation study of Exploring Motion and Forces (2003–2004)* (SCALE-uP Report No. 5). Washington, DC: George Washington University, SCALE-uP.

Additional sources:

Pyke, C., Lynch, S., Kuipers, J., Szesze, M., & Watson, W. (2005). *Implementation study of Exploring Motion and Forces (2004–2005)* (SCALE-uP Report No. 8). Washington, DC: George Washington University, SCALE-uP.

Pyke, C., Lynch, S., Kuipers, J., Szesze, M., & Watson, B. (2006). *Implementation study of Exploring Motion and Forces (2005–2006)* (SCALE-uP Report No. 13). Washington, DC: George Washington University, SCALE-uP.

Rethinam, V., Pyke, C., & Lynch, S. (2008). Using multi-level analyses to study the effectiveness of science curriculum materials. *Evaluation & Research in Education*, 21(1), 18–42.

Studies that meet WWC evidence standards with reservations

No studies meet WWC evidence standards with reservations.

Study that does not meet WWC evidence standards

Ward, R. B., Sadler, P. M., & Shapiro, I. I. (2008). Learning physical science through astronomy activities: A comparison between constructivist and traditional approaches in grades 3–6. *Astronomy Education Review*, 6(2), 1–19. 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.

Appendix A: Research details for Pyke et al., 2004

Pyke, C., Lynch, S., Kuipers, J., Szesze, M., & Watson, W. (2004). *Implementation study of Exploring Motion and Forces (2003–2004)* (SCALE-uP Report No. 5). Washington, DC: George Washington University, SCALE-uP.

Table A1. Summary of findings

Meets WWC evidence standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
General science achievement	20 schools/6,058 students	+4	No

Setting The study took place in 20 schools in Maryland’s Montgomery County School District (10 schools for Cohorts 1 and 2 and 10 schools for Cohort 3). The student population of this large suburban district is 43% White, 22% African American, 14% Asian American, and 20% Hispanic. Self-contained classrooms of English language learners and special education students were excluded from this study. The study is part of a multiyear research project called “Scaling Up Curriculum for Achievement, Learning, and Equity Project” (SCALE-uP).⁹

Study sample In this randomized study, researchers followed three cohorts of sixth-grade students. A sampling frame of five school-profile categories was created based on achievement and demographic factors. Each profile contained approximately seven schools. The authors randomly selected two schools from each of the five profiles, and one school from each pair was randomly selected to implement *ARIES: Exploring Motion and Forces*. Cohorts 1 and 2 were sixth-grade students from the same set of implementing and non-implementing schools. For the analysis of Cohort 3, researchers excluded schools in the previous analyses and repeated the randomization scheme on the remaining schools.¹⁰ Cohorts 1 and 2 consisted of sixth-grade students in the 2003–04 and 2004–05 school years, respectively. The Cohort 1 analysis sample included 1,266 sixth-grade students who received the intervention and 1,115 sixth-grade students who did not receive the intervention. Cohort 2 included 910 sixth-grade students who received the intervention and 1,005 who did not receive the intervention. For these cohorts, students were randomly dropped from each of the matched pairs of schools to create balanced sample sizes across schools. Cohort 3 consisted of sixth-grade students in the 2005–06 school year and included 902 students who received the intervention and 860 students who did not receive the intervention. Overall and differential attrition rates for Cohorts 1, 2, and 3 were low.¹¹ Students’ outcomes were presented by cohort; these findings can be found in Appendix C. Additional findings for subgroups by gender, race/ethnicity, and students eligible for Special Education (SPED) can be found in Appendix D.

Intervention group

Exploring Motion and Forces: Speed, Acceleration, and Friction (Harvard-Smithsonian Center for Astrophysics, 2001) is an inquiry-based middle school science curriculum that received an acceptable rating by the American Association for the Advancement of Science Project 2061, a curriculum analysis project funded by the Interagency Educational Research Initiative of the National Science Foundation. *ARIES: Exploring Motion and Forces*, a six-week physical science unit, is intended for fifth- to eighth-grade students and is broken into four parts and 18 “explorations.” Its focus is on inquiry-centered and activity-based student-centered material. The materials contain a teacher manual, a student science journal, and exploration materials. Cohort 3 students also received a 114-page notebook. The same notebook was given to teachers in the first two cohorts, with the expectation that the pages would be copied and distributed to students. The material covered in *ARIES: Exploring Motion and Forces* was aligned with the district’s sixth-grade curriculum standards, but it did not cover all of the topics. The unit was implemented over a period of approximately six weeks.

Comparison group

Comparison group teachers used regular curriculum materials normally available to Montgomery County Public Schools’ teachers that addressed the same target benchmarks. There was no single comparison group curriculum, and teachers were not restricted to any specific material. In the final report only, for Cohort 3, Pyke et al. (2006) clarified that teachers in the comparison group used a wide variety of sources, including textbooks, handouts, the Internet, and videos.

Outcomes and measurement

For both the pretest and posttest, students took the Motion and Forces Assessment (MFA), which consisted of six constructed responses and four selected responses. For a more detailed description of this outcome measure, see Appendix B.

Support for implementation

The study does not discuss the training of teachers.

Appendix B: Outcome measures for each domain

General science achievement

Physical science construct

Motion and Forces Assessment (MFA)

The MFA is aligned with the Benchmarks for Science Literacy. It consists of six constructed responses and four selected responses that are valued differently to create a weighted score. The MFA was designed by researchers at George Washington University for use in two projects: Scaling Up Curriculum for Achievement, Learning, and Equity Project (SCALE-uP) and the American Association for the Advancement of Science (AAAS) Project 2061. Cronbach's alpha estimated on the entire sample was 0.52. For Cohort 1 and Cohort 2, inter-rater reliability on the six constructed response questions had average kappa statistics ranging from 0.67 to 0.88 across all items and ranges. Changes were made to the MFA for Cohort 3. One of the constructed response questions was found to be too difficult, so the researchers excluded this problem in the analysis of Cohort 3. Average inter-rater reliability kappa statistics for Cohort 3 were based on a 5% sample of the assessments and ranged from 0.77 to 0.91 across all items and ranges (as cited in Pyke et al., 2004).

Appendix C: Findings included in the rating for the general science achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Pyke et al., 2004^a								
<i>Motion and Forces Assessment</i>	Grade 6/ Cohort 1	10 schools/ 2,381 students	60.91 (22.63)	57.51 (22.88)	3.40	0.15	+6	< 0.05
<i>Motion and Forces Assessment</i>	Grade 6/ Cohort 2	10 schools/ 1,915 students	53.76 (23.37)	55.42 (23.35)	-1.66	-0.07	-3	> 0.05
<i>Motion and Forces Assessment</i>	Grade 6/ Cohort 3	10 schools/ 1,762 students	56.09 (22.18)	50.84 (22.49)	5.25	0.23	+9	< 0.05
Domain average for general science achievement (Pyke et al., 2004)						0.10	+4	Not statistically significant

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. 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 a study’s domain average was determined by the WWC. Outcomes for each cohort were provided in three separate reports. Reports for Cohort 2 (Pyke et al., 2005) and Cohort 3 (Pyke et al., 2006) are included as additional sources of information for the Pyke et al. (2004) study in the references section.

^a For Pyke et al. (2004), a correction for clustering was needed and resulted in significance levels for Cohort 1 and Cohort 3 ($p > 0.05$) that differ from those in the original study. The p -values presented in Appendix C were reported in the original study. For Cohort 1 and Cohort 2, the intervention and comparison group mean outcome values are ANCOVA-adjusted posttest scores, with pretest scores being treated as a covariate. For Cohort 3, the intervention and comparison group mean outcome values are the unadjusted posttest means. The data reported in the table for Cohort 1 and Cohort 2 were provided by the author to the WWC and are not the data included in the original reports. The data for Cohort 1 and Cohort 2 were requested from the author because the analyses included in the original reports involved methods that are not acceptable by the WWC: adjustments to pretest scores (Cohort 1) and data imputation for missing observations (Cohort 2).

Appendix D: Subgroup findings for the general science achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Pyke et al., 2004^a								
Grade 6, Cohort 1								
<i>Motion and Forces Assessment</i>	Females	10 schools/ 1,064 students	54.21 (22.71)	53.05 (22.35)	1.16	0.09	+4	> 0.05
<i>Motion and Forces Assessment</i>	Hispanic	10 schools/ 394 students	46.74 (22.83)	47.04 (21.67)	-0.30	-0.01	-1	> 0.05
<i>Motion and Forces Assessment</i>	SPED	10 schools/ 195 students	42.69 (21.83)	48.26 (23.12)	-5.57	-0.18	-7	> 0.05

Table Notes: The supplemental findings presented in this table are additional subgroup findings for Cohort 1 from the study in this report that do not factor into the determination of the intervention rating. The WWC reports findings only for subgroups of interest that are equivalent with regard to pretest scores or include statistical adjustment for characteristics relevant to equating the groups as given in the protocol. Subgroup findings for Cohort 2 (Pyke et al., 2005) are not included in this report because the researchers imputed test score outcomes for a substantial proportion of students. Subgroup findings for Cohort 3 (Pyke et al., 2006) are not included in this report because evidence of baseline equivalence for the analysis samples was not provided in the study. Student subgroups include gender, ethnicity, and eligibility for special education services (SPED). 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.

^a For Pyke et al. (2004), a correction for clustering was needed and resulted in significance levels that differ from those in the original study. The intervention group mean outcome values are the unadjusted comparison group posttest means plus the difference in mean gains between the intervention and comparison groups. The p-values presented here were computed by the WWC. The results for Cohort 2 are excluded from the table because the researchers imputed test score outcomes for a substantial proportion of students.

Endnotes

¹ The descriptive information for this program was obtained from publicly available sources: the program's website (<http://www.cfa.harvard.edu/sed/projects/sednet.html>, downloaded June 2011), and Ochsendorf, Chen Deprey, Lopez, Roudebush, Lynch, & Pyke (2003). *Using the Project 2061 Curriculum Analysis to Rate a Middle School Science Curriculum Unit: ARIES: Exploring Motion and Forces* (SCALE-uP Report No. 10). Washington, DC: George Washington University and Montgomery County Public Schools. The program description was provided to the developer in June 2011; however, the WWC received no response. 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 June 2011.

² Harvard-Smithsonian Center for Astrophysics. (2001). *ARIES—Exploring Motion and Forces: Speed, Acceleration, and Friction*. Watertown, MA: Charlesbridge Publishing.

³ The studies in this report were reviewed using WWC Evidence Standards, Version 2.1, as described in the Science review protocol, Version 2.0. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

⁴ For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 12. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies.

⁵ Because science was not an elective course, the WWC review team downplayed selection bias for Cohort 2 students that could arise as a result of the selective enrollment and non-enrollment of participants into a study in the light of the upcoming intervention assignment.

⁶ Outcomes for each cohort were provided in three separate reports. Reports for Cohort 2 (Pyke et al., 2005) and Cohort 3 (Pyke et al., 2006) are included as additional sources of information for the Pyke et al. (2004) study in the references section.

⁷ 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 the WWC Procedures and Standards Handbook, Appendix C for clustering and the WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. For Pyke et al. (2004), a correction for clustering was needed, so the significance levels differ from those reported in the original study.

⁸ The WWC computes an average effect size as a simple average of the effect sizes across all individual findings within the study domain.

⁹ SCALE-uP is funded by the Interagency Education Research Initiative and administered by the National Science Foundation. This intervention report presents results from Year 2, Year 3, and Year 4, when authors reported the results of the first, second, and third years of implementation of *Exploring Motion and Forces: Speed, Acceleration, and Friction* (Harvard-Smithsonian Center for Astrophysics, 2001) for sixth-grade students in the Montgomery County School District. In project Year 0 and Year 1, authors reported results of *Chemistry That Applies* (State of Michigan, 1993) on eighth-grade students in the same school district. Their findings are presented in the Science intervention report for *Chemistry That Applies*. In Year 2 and Year 3, authors reported the results of the first and second years of implementation of the curriculum unit *GEMS® The Real Reasons for Seasons* study (Lawrence Hall of Science, 2000) for seventh-grade students in the same district.

¹⁰ One of the five matched pairs of schools for Cohort 3 was excluded from the analysis because the comparison school in that pair did not assess students.

¹¹ For Cohorts 1 and 2, attrition rates were based on counts provided by the author after an inquiry by the WWC.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, May). *Science intervention report: ARIES: Exploring Motion and Forces*. Retrieved from <http://whatworks.ed.gov>.

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC evidence standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC evidence standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows 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, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

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 treatment 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 and inclusion in this report 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.
Extent of evidence	An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria earlier in this report.
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 treatment 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 treatment and comparison groups.
Rating of effectiveness	The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria earlier in this report.
Single-case design	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.