

Appendix

Appendix A1.1 Study characteristics: Dietsch, Bayha, & Zheng, 2005 (randomized controlled trial)

Characteristic	Description
Study citation	Dietsch, B., Bayha, J. L., & Zheng, H. (2005, April). <i>Short-term effects of a character education program among fourth grade students</i> . Paper presented at the American Educational Research Association annual meeting, Montreal, Quebec, Canada.
Participants	More than 400 fourth-grade students attending elementary schools in Louisiana and Florida. Eight schools whose principals expressed an interest in implementing the program were selected at random to participate. In each of the eight schools two fourth-grade teachers were randomly assigned to the intervention or control condition. Of the students in the study, 75% participated in the National School Lunch Program and about 50% met state reading and mathematics standards. The authors noted that all grade levels in the schools fell below state expectations in core academic areas.
Setting	All participating schools were located in rural, poor, and ethnically diverse communities in Louisiana and Florida.
Intervention	Teachers were encouraged to complete at least one program lesson per week for 15 weeks. The authors note that comprehensive character education would involve administrators, parents, and others on campus; this study used only the materials for classrooms. School officials were reluctant to have researchers enter the classroom, so fidelity and duration of student exposure to the intervention were not measured.
Comparison	Teachers in the comparison group were asked to teach as they normally would. They did not institute any new character education instruction during the study. After the posttests for the study, the teachers in the comparison group received copies of the classroom materials used in the program.
Primary outcomes and measurement	Outcome measures are based on student surveys and archival data (for attendance, citizenship, grades). The outcomes include mathematics, reading, and citizenship grades from report cards; attendance rates; positive and negative student behavior reflecting core values; and student attitudes toward core values. (See Appendices A2.1, A2.2, and A2.3.) ¹
Teacher training	The authors characterized teacher training as “minimal.” Teachers who implemented the intervention received lesson materials and training. Each classroom assigned to the intervention group received two visits from a trainer, who observed the students and offered assistance to the teacher.

1. Additional teacher-reported outcomes in the behavior domain (i.e., student behavior) could not be reviewed because the standard deviations reported by the study authors represented variations among teachers rather than variations among students. Therefore, a student-level effect size could not be estimated for this outcome. Further, the measure, teacher opinions of character education implementation, was not included in this review because it did not assess student outcomes.

Appendix A1.2 Study characteristics: DeVargas 1998 (quasi-experimental design)

Characteristic	Description
Study citation	DeVargas, R. C. (1998). A study of <i>Lessons in Character</i> : The effect of moral development curriculum upon moral judgment. <i>Dissertation Abstracts International</i> , 59 (11-A), 4042. (UMI No. 9913706)
Participants	Fifth-grade students selected from nine intervention and six comparison schools, with 21 classrooms and 31 students from the intervention schools and 17 classrooms and 30 students from the comparison schools. No information was reported on the sample's racial/ethnic composition or socioeconomic characteristics.
Setting	The study was conducted in Fort Worth Independent School District in Fort Worth, Texas.
Intervention	Teachers provided the instruction with the help of the school counselor, who acted as a trainer/facilitator. In some cases the counselor alone instructed the various classes in his/her school on a rotational basis. No information is reported about the fidelity of the implementation other than the author's assumption that the implementation of the <i>Lessons In Character</i> curriculum was similar across classrooms and across schools.
Comparison	The comparison schools were selected from the same geographic area as the intervention schools. Detailed information on the comparison condition was not provided.
Primary outcomes and measurement	The only outcome investigated in this study was moral judgment, as measured by the Sociomoral Reflection Measure—Short Form. (See Appendices A2.1, A2.2, and A2.3.)
Teacher training	School counselors acted as trainers or facilitators to the teachers during the implementation period.

Appendix A2.1 Outcome measures in the behavior domain

Outcome measure	Description
Citizenship grades	Proportion changing from below to above “satisfactory” on student report card (as cited in Dietsch, Bayha, & Zheng, 2005).

Appendix A2.2 Outcome measures in the knowledge, attitudes, and values domain

Outcome measure	Description
Student respect	Student survey: The WWC aggregated the nine item-level results for a modified version of the subscales of the School as a Caring Community Profile (SCCP-II) by averaging their means and effect sizes (as cited in Dietsch, Bayha, & Zheng, 2005).
Friendship and belonging	Student survey: The WWC aggregated the eight item-level results for a modified version of the subscales of the School as a Caring Community Profile (SCCP-II) by averaging their means and effect sizes (as cited in Dietsch, Bayha, & Zheng, 2005).
Shaping the environment	Student survey: The WWC aggregated the seven item-level results for a modified version of the subscales of the School as a Caring Community Profile (SCCP-II) by averaging their means and effect sizes (as cited in Dietsch, Bayha, & Zheng, 2005).
Support and care by teachers and staff	Student survey: The WWC aggregated the four item-level results for a modified version of the subscales of the School as a Caring Community Profile (SCCP-II) by averaging their means and effect sizes (as cited in Dietsch, Bayha, & Zheng, 2005).
Moral development	Gibbs' Sociomoral Reflection Measure–Short Form (SRM).

Appendix A2.3 Outcome measures in the academic achievement domain

Outcome measure	Description
Reading grades	Grade point average in reading (as cited in Dietsch, Bayha, & Zheng, 2005).
Mathematics grades	Grade point average in mathematics (as cited in Dietsch, Bayha, & Zheng, 2005).
Attendance (absences)	Number of absences over the course of one semester (as cited in Dietsch, Bayha, & Zheng, 2005).

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

Outcome measure ⁴	Study sample	Sample size (classrooms/ students)	Author's findings from the study		WWC calculations			
			Mean outcome ² (standard deviation ³)		Mean difference ⁵ (column 1– column 2)	Effect size ⁶	Statistical significance ⁷ (at $\alpha = 0.05$)	Improvement index ⁸
			Lessons in Character group (column 1)	Comparison group (column 2)				
Dietsch, Bayha, & Zheng, 2005 (randomized controlled trial)								
Citizenship grades	Grade 4	4/141	0.12 (0.32)	0.06 (0.23)	0.06	0.21	ns	+8
Domain average⁹ for behavior						0.21	ns	+8

1. This appendix reports findings considered for the effectiveness rating and the improvement index.
2. This appendix reports unadjusted posttest means reported by the study authors and the covariate-adjusted mean differences estimated by the WWC based on information from the study authors for the purpose of effect size calculation.
3. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes. The WWC requested and received standard deviations for all outcomes in the first study, Dietsch, Bayha, & Zheng (2005), because they were not reported in the paper.
4. Additional teacher-reported outcomes in the behavior domain (i.e., student behavior) could not be reviewed because the standard deviations reported by the study authors represented variations among teachers rather than variations among students. Therefore, a student-level effect size could not be estimated for this outcome. Further, the measure, teacher opinions of character education implementation, was not included in this review because it did not assess student outcomes. For a description of the types of outcome measures included in this topic review see the [Character Education Abstract](#).
5. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The measure, absences, was reversed so that a positive difference would favor the intervention group in Dietsch, Bayha, & Zheng (2005).
6. For an explanation of the effect size calculation, please see the [Technical Details of WWC-Conducted Computations](#).
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between groups. The level of statistical significance was calculated by the WWC and where necessary, corrects for clustering within classrooms or schools, and for multiple comparisons. For an explanation, see the [WWC Tutorial on Mismatch](#). See the [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate the statistical significance. In the case of the *Lessons in Character* report, a correction for clustering was needed.
8. 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.
9. This row provides the study average, which is also the domain average in this case. 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 A3.2 Summary of study findings included in the rating for the knowledge, attitudes, and values domain¹

Outcome measure ⁴	Study sample	Sample size (classrooms/ students)	Author's findings from the study		WWC calculations			
			Mean outcome ² (standard deviation ³)		Mean difference ⁵ (column 1– column 2)	Effect size ⁶	Statistical significance ⁷ (at $\alpha = 0.05$)	Improvement index ⁸
			Lessons in Character group (column 1)	Comparison group (column 2)				
Dietsch, Bayha, & Zheng, 2005 (randomized controlled trial)								
Student respect	Grade 4	21/372	3.30 (na)	3.11 (na)	0.19	0.13	ns	+5
Friendship and belonging	Grade 4	21/372	3.22 (na)	2.97 (na)	0.25	0.18	ns	+7
Shaping the environment	Grade 4	21/372	3.04 (na)	2.70 (na)	0.34	0.24	ns	+9
Support and care by teachers and staff	Grade 4	21/372	4.11 (na)	3.70 (na)	0.41	0.30	ns	+12
Average⁹ for knowledge, attitudes, values						0.21	ns	+8
DeVargas, 1998 (quasi-experimental study)								
Moral development	Grade 5	15 schools/ 61 students	2.20 (0.29)	2.17 (0.27)	0.03	0.11	ns	+4
Average⁹ for knowledge, attitudes, values						0.11	ns	+4
Domain average⁹ for knowledge, attitudes, and values across all studies						0.16	na	+6

na = not applicable

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the improvement index. The WWC averaged individual items to provide the scales reported here for four outcomes (student respect, friendship and belonging, shaping the environment, and support and care by teachers and staff). Findings on the individual item level are presented in Appendix A4.
2. This appendix reports posttest unadjusted means reported by the study authors and the covariate-adjusted mean differences estimated by the WWC based on information from the study authors for the purpose of effect size calculation.
3. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes. The WWC requested and received standard deviations for all outcomes in the first study, Dietsch, Bayha, & Zheng (2005), because they were not reported in the paper. These standard deviations were used to calculate effect size on the individual item level. (See Appendix A4.) Therefore, standard deviations on the scale level are presented as "not applicable" in this appendix.
4. Additional teacher-reported outcomes in the knowledge, attitudes, and values (i.e., student knowledge and reasoning; student values) domain could not be reviewed. The standard deviations reported by the study authors represented variations among teachers rather than variations among students. Therefore, a student effect size could not be estimated for this outcome. Further, the measure, teacher opinions of character education implementation, was not included in this review because it did not assess student outcomes. For a description of the types of outcome measures included in this topic review see the [Character Education Abstract](#).
5. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The measure, absences, was reversed so that a positive difference would favor the intervention group in Dietsch, Bayha, & Zheng (2005).
6. For an explanation on effect size calculation, please see the [Technical Details of WWC-Conducted Computations](#).
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between groups. The level of statistical significance was calculated by the WWC and where necessary, corrects for clustering within classrooms or schools, and for multiple comparisons. For an explanation, see the [WWC Tutorial on Mismatch](#). See the [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate the statistical significance. In the case of the *Lessons in Character* report, a correction for clustering and multiple comparisons was needed.
8. 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.
9. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.

Appendix A3.3 Summary of study findings included in the rating for the academic achievement domain¹

Outcome measure	Study sample	Sample size (classrooms/ students)	Author's findings from the study		WWC calculations			
			Mean outcome ² (standard deviation ³)		Mean difference ⁴ (column 1– column 2)	Effect size ⁵	Statistical significance ⁶ (at $\alpha = 0.05$)	Improvement index ⁷
			Lessons in Character group (column 1)	Comparison group (column 2)				
Dietsch, Bayha, & Zheng, 2005 (randomized controlled trial)								
Reading grades	Grade 4	4/141	2.95 (0.88)	2.66 (0.99)	0.29	0.31	ns	+12
Mathematics grades	Grade 4	4/141	2.95 (0.84)	2.56 (0.86)	0.39	0.46	Statistically significant	+18
Attendance (absences)	Grade 4	4/141	1.46 (1.93)	2.59 (2.72)	-1.13	0.48	Statistically significant	+18
Domain average⁸ for academic achievement						0.41	Statistically significant	+16

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the improvement index.
2. This appendix reports unadjusted posttest means reported by the study authors and the covariate-adjusted mean differences estimated by the WWC based on information from the study authors for the purpose of effect size calculation.
3. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes. The WWC requested and received standard deviations for all outcomes in the first study, Dietsch, Bayha, & Zheng (2005), because they were not reported in the paper.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The measure, absences, was reversed so that a positive difference would favor the intervention group in Dietsch, Bayha, & Zheng (2005).
5. For an explanation on effect size calculation, please see the [Technical Details of WWC-Conducted Computations](#).
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between groups. The level of statistical significance was calculated by the WWC and where necessary, corrects for clustering within classrooms or schools, and for multiple comparisons. For an explanation, see the [WWC Tutorial on Mismatch](#). See the [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate the statistical significance. In the case of the *Lessons in Character* report, a correction for clustering and multiple comparisons was needed.
7. 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.
8. This row provides the study average, which is also the domain average in this case. 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 Summary of detailed study findings in the knowledge, attitudes, and values domain¹

Outcome measure	Study sample	Sample size (classrooms/ students)	Author's findings from the study		WWC calculations			
			Mean outcome ² (standard deviation ³)		Mean difference ⁴ (column 1– column 2)	Effect size ⁵	Statistical significance ⁶ (at $\alpha = 0.05$)	Improvement index ⁷
			Lessons in Character group (column 1)	Comparison group (column 2)				
Dietsch, Bayha, & Zheng, 2005 (randomized controlled trial)								
Student respect								
Treat classmates with respect	Grade 4	21/372	2.97 (1.36)	2.90 (1.29)	0.07	0.05	ns	+2
Respect personal property of others	Grade 4	21/372	3.09 (1.49)	3.28 (1.48)	–0.19	–0.13	ns	–5
Show respect for school property	Grade 4	21/372	3.13 (1.55)	2.98 (1.44)	0.15	0.10	ns	+4
Behave respectfully toward school staff	Grade 4	21/372	3.41 (1.45)	3.31 (1.33)	0.10	0.08	ns	+3
Show respect toward teachers	Grade 4	21/372	3.68 (1.54)	3.43 (1.38)	0.25	0.18	ns	+7
Do not pick on other students	Grade 4	21/372	3.20 (1.58)	2.64 (1.48)	0.56	0.37	ns	+14
Show good sportsmanship	Grade 4	21/372	3.87 (1.40)	3.39 (1.44)	0.48*	0.33	ns	+13
Show respect toward other students	Grade 4	21/372	3.73 (1.35)	3.27 (1.32)	0.46*	0.35	ns	+14
Do not put others down	Grade 4	21/372	2.63 (1.45)	2.77 (1.27)	–0.14	–0.11	ns	–4
Friendship and belonging								
Comfort others who are sad	Grade 4	21/372	3.10 (1.47)	2.90 (1.33)	0.20	0.14	ns	+6
Help each other even if not friends	Grade 4	21/372	2.95 (1.43)	2.68 (1.32)	0.27	0.19	ns	+8
Work well with each other	Grade 4	21/372	3.35 (1.33)	3.36 (1.31)	–0.01	0.00	ns	+0

(continued)

Appendix A4 Summary of detailed study findings in the knowledge, attitudes, and values domain¹ (continued)

Outcome measure	Study sample	Sample size (classrooms/ students)	Author's findings from the study		WWC calculations				
			Mean outcome ² (standard deviation ³)		Mean difference ⁴ (column 1– column 2)	Effect size ⁵	Statistical significance ⁶ (at $\alpha = 0.05$)	Improvement index ⁷	
			Lessons in Character group (column 1)	Comparison group (column 2)					
Help new students feel accepted	Grade 4	21/372	3.61 (1.42)	3.39 (1.37)	0.22	0.16	ns	+6	
Be willing to forgive each other	Grade 4	21/372	3.33 (1.52)	2.87 (1.36)	0.46	0.31	ns	+12	
Be patient with each other	Grade 4	21/372	3.04 (1.34)	2.66 (1.26)	0.38	0.29	ns	+12	
Listen to each other in the classroom	Grade 4	21/372	3.13 (1.49)	2.93 (1.39)	0.20	0.14	ns	+5	
Share what they have with others	Grade 4	21/372	3.25 (1.47)	2.93 (1.40)	0.32	0.22	ns	+9	
Shaping the environment									
Try to make it up if do something harmful	Grade 4	21/372	3.15 (1.45)	2.89 (1.39)	0.26	0.18	ns	+7	
Get other students to follow school rules	Grade 4	21/372	2.95 (1.53)	2.69 (1.29)	0.26	0.18	ns	+7	
Help to improve the school	Grade 4	21/372	3.14 (1.56)	2.99 (1.30)	0.15	0.11	ns	+4	
Have a positive influence on behavior	Grade 4	21/372	3.04 (1.43)	2.67 (1.31)	0.37	0.27	ns	+11	
Resolve conflicts without fighting or insults	Grade 4	21/372	2.87 (1.50)	2.48 (1.43)	0.39	0.27	ns	+10	
Stop it when see others being picked on	Grade 4	21/372	2.95 (1.53)	2.57 (1.31)	0.38	0.26	ns	+10	
Help to solve school problems	Grade 4	21/372	3.15 (1.56)	2.58 (1.39)	0.57	0.38	ns	+15	
Support and care by teachers and staff									
Can talk to teachers about their problems	Grade 4	21/372	3.84 (1.37)	3.52 (1.40)	0.32	0.23	ns	+9	

(continued)

Appendix A4 Summary of detailed study findings in the knowledge, attitudes, and values domain¹ (continued)

Outcome measure	Study sample	Sample size (classrooms/students)	Author's findings from the study		WWC calculations				
			Mean outcome ² (standard deviation ³)		Mean difference ⁴ (column 1–column 2)	Effect size ⁵	Statistical significance ⁶ (at $\alpha = 0.05$)	Improvement index ⁷	
			Lessons in Character group (column 1)	Comparison group (column 2)					
Teachers help students who need extra help	Grade 4	21/372	4.47 (1.26)	3.65 (1.45)	0.82	0.61	Statistically significant	+23	
Adults can be counted on to make sure that students are safe	Grade 4	21/372	4.11 (1.31)	3.93 (1.34)	0.18	0.14	ns	+5	
Teachers are fair in treatment of students	Grade 4	21/372	4.03 (1.38)	3.70 (1.43)	0.33	0.24	ns	+9	

1. This appendix presents item-level findings for four scales in the knowledge, attitudes, and values domain: student respect, friendship and belonging, shaping the environment, and support and care by teachers and staff. Aggregated scale scores were used for rating purposes and are presented in Appendix A3.2.
2. This appendix reports posttest unadjusted means reported by the study authors.
3. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
4. Covariate-adjusted mean differences estimated by the WWC based on information from the study authors for the purpose of effect size calculation. Positive differences favor the intervention group; negative differences favor the comparison group.
5. For an explanation on effect size calculation, please see the [Technical Details of WWC-Conducted Computations](#).
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between groups. The level of statistical significance was calculated by the WWC and where necessary, corrects for clustering within classrooms or schools. For an explanation, see the [WWC Tutorial on Mismatch](#). See the [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate the statistical significance. In the case of the *Lessons in Character* report, a correction for clustering was needed.
7. 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.

Appendix A5.1 Rating for the behavior 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 behavior, the WWC rated *Lessons in Character* as having no discernible effects. It did not meet the criteria for positive effects because it only had one study. In addition, it did not meet the criteria for other ratings (potentially positive effects, mixed effects, potentially negative effects, negative effects) because none of the studies showed statistically significant or substantively important effects.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Met. The WWC analysis found no statistically significant or substantively important 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. *Lessons in Character* had only one evaluation study meeting WWC evidence standards that reported findings on behavior, and so did not meet this criterion. Further, although that one study used a strong design, it did not show statistically significant positive effects.

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

Met. The WWC analysis found no statistically significant or substantively important negative effects in this domain.

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, thus qualifying as a *positive* effect.

Not met. The WWC analysis found no statistically significant or substantively important positive effects in this domain.

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

Not met. No studies showed a statistically significant or substantively important negative effect. Because one study showed indeterminate effects and no studies showed statistically significant or substantively important positive effects, *Lessons in Character* did not meet this criterion.

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 effects. See the [WWC Intervention Rating Scheme](#) for a complete description.

(continued)

Appendix A5.1 Rating for the behavior domain *(continued)*

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

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect. 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. The WWC analysis found no statistically significant or substantively important positive or negative effects in this domain.

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. The WWC analysis found no statistically significant or substantively important effects in this domain.

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. The WWC analysis found no statistically significant or substantively important negative effects in this domain.

- 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. The WWC analysis found no statistically significant or substantively important effects in this domain.

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 is based on a strong design.

Not met. The WWC analysis found no statistically significant negative effects in this domain.

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

Met. The WWC analysis found no statistically significant or substantively important positive effects in this domain.

Appendix A5.2 Rating for the knowledge, attitudes, and values 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 knowledge, attitudes, and values, the WWC rated *Lessons in Character* as having no discernible effects. It did not meet the criteria for other ratings (positive effects, potentially positive effects, mixed effects, potentially negative effects, and negative effects) because none of the studies showed statistically significant or substantively important effects.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important or statistically significant effect, either positive or negative.

Met. The WWC analysis found no statistically significant 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. *Lessons in Character* had two evaluation studies meeting WWC evidence standards that reported findings on knowledge, attitudes, and values. One of those studies used a strong design. But none of the studies showed a statistically significant or substantively important positive effect, so *Lessons in Character* did not meet this criterion.

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

Met. The WWC analysis found no statistically significant or substantively important negative effects in this domain.

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, thus qualifying as a *positive* effect.

Not met. The WWC analysis found no statistically significant or substantively important positive effects in this domain.

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

Not met. No studies showed a statistically significant or substantively important negative effect. Because one study showed indeterminate effects and no studies showed statistically significant or substantively important positive effects, *Lessons in Character* did not meet this criterion.

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 effects. See the [WWC Intervention Rating Scheme](#) for a complete description.

(continued)

Appendix A5.2 Rating for the knowledge, attitudes, and values domain *(continued)*

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

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect. 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. The WWC analysis found no statistically significant or substantively important positive or negative effects in this domain.

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. The WWC analysis found no statistically significant or substantively important effects in this domain.

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. The WWC analysis found no statistically significant or substantively important negative effects in this domain.

- 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. The WWC analysis found no statistically significant or substantively important positive effects in this domain.

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 is based on a strong design.

Not met. The WWC analysis found no statistically significant negative effects in this domain.

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

Met. The WWC analysis found no statistically significant or substantively important positive effects in this domain.

Appendix A5.3 Rating for the academic achievement 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 academic achievement, the WWC rated *Lessons in Character* as having no discernible effects. It did not meet the criteria for positive effects because none of the studies showed statistically significant or substantively important effects. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Lessons in Character* 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, thus qualifying as a *positive* effect.

Met. In the one study on *Lessons in Character* that examined student academic outcomes, the average effect size was positive and statistically significant. Further, effects on two student outcomes were positive and statistically significant.²

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

Met. No studies showed a statistically significant or substantively important negative effect. Because one study showed statistically significant positive effects and no studies showed indeterminate effects, *Lessons in Character* met this criterion.

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. *Lessons in Character* had only one evaluation study meeting WWC evidence standards that reported findings on academic achievement. But that one study used a strong design and statistically significant positive effects.

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

Met. The WWC analysis found no 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 effects. See the [WWC Intervention Rating Scheme](#) for a complete description.

2. Although the study author reported three statistically significant effects, the WWC analysis confirmed the significance of only two of those findings.