

Appendix

Appendix A1 Study characteristics: Eisen, Zellman, & Murray, 2003 (randomized controlled trial)

Characteristic	Description
Study citation	Eisen, M., Zellman, G. L., & Murray, D. M. (2003). Evaluating the Lions-Quest “Skills for Adolescence” drug education program: Second-year behavior outcomes. <i>Addictive Behaviors, 28</i> , 883–897.
Participants	Participants of this study included 7,426 students who were followed from sixth through eighth grade. Female students comprised 52% of the sample. The distribution of minority students was as follows: 34% Hispanic, 18% African-American, and 7% Asian-American.
Setting	The study took place in 34 middle schools from four school districts located in three large metropolitan areas (Los Angeles–Long Beach, California; Detroit–Wayne County, Michigan; and Washington, DC–Baltimore, Maryland).
Intervention	A one-semester, 40-session <i>Skills for Adolescence</i> curriculum was implemented in English or Spanish in the intervention schools. Each session lasted 35–45 minutes. No information was provided on implementation fidelity other than that teachers were required to schedule and teach 8 of the 40 sessions that included drug prevention components, knowing that they may be observed by project staff and consenting to this observation. ¹
Comparison	The comparison group “received their usual drug education programming” and were left to the discretion of the teachers at each school. A range of drug prevention programs and related practices were implemented in the comparison schools, including school assemblies, teacher-devised classroom curricula, and exposure to the Drug Abuse Resistance Education (DARE) program. The comparison schools were promised implementation of <i>Skills for Adolescence</i> once the final follow-up data were collected.
Primary outcomes and measurement	Outcome measures examined student self-reported cigarette smoking and alcohol, marijuana, and other illicit drug use prevalence rates. (See Appendix A2 for more detailed descriptions of outcome measures.) ²
Teacher training	The teachers attended a three-day workshop conducted by Quest International certified trainers and received teacher manuals and workbooks for each of their students.

1. The WWC requested information from the first study author about the timing of intervention delivery. The study author reported that in some schools the intervention was implemented during the first semester, while in other schools the intervention was implemented later in the academic year.
2. Four additional student outcomes measuring attitudes were examined in this study: intentions to use tobacco, alcohol, and other drugs; perceived harm of tobacco, alcohol, and other drugs; refusal self-efficacy (to peer pressure); and perceived peer (close friend) use of tobacco, alcohol, and other drugs. These student outcomes were not reviewed because of lack of statistical information for computing effect sizes.

Appendix A2 Outcome measures in the behavior domain

Outcome measure	Description
Alcohol use—lifetime	One item on which students indicated if they had ever drunk alcohol. This item was adapted from the Monitoring the Future Study (MTF; as cited in Eisen, Zellman, & Murray, 2003).
Alcohol use—last 30 days	One item on which students indicated if they had drunk alcohol in the last 30 days. This item was adapted from the Monitoring the Future Study (MTF; as cited in Eisen, Zellman, & Murray, 2003).
Binge drinking	One item on which students indicated whether they had engaged in excessive drinking for three or more days during the last 30 days. This item was adapted from the Monitoring the Future Study (MTF; as cited in Eisen, Zellman, & Murray, 2003).
Cigarette smoking—lifetime	One item on which students indicated if they had ever smoked cigarettes. This outcome measure was among the items established by the National Cancer Institute as a standard for surveying prevalence rates (as cited in Eisen, Zellman, & Murray, 2003).
Cigarette smoking—last 30 days	One item on which students indicated if they had smoked cigarettes in the last 30 days. This outcome measure was among the items established by the National Cancer Institute as a standard for surveying prevalence rates (as cited in Eisen, Zellman, & Murray, 2003).
Marijuana use—lifetime	One item on which students indicated if they had ever used marijuana. This item was adapted from the Monitoring the Future Study (MTF; as cited in Eisen, Zellman, & Murray, 2003).
Marijuana use—last 30 days	One item on which students indicated if they had used marijuana in the last 30 days. This item was adapted from the Monitoring the Future Study (MTF; as cited in Eisen, Zellman, & Murray, 2003).
Other illicit drug use—lifetime	One item on which students indicated if they had ever used other illicit drugs. This item was adapted from the Monitoring the Future Study (MTF; as cited in Eisen, Zellman, & Murray, 2003).
Other illicit drug use—last 30 days	One item on which students indicated if they had used other illicit drugs in the last 30 days. This item was adapted from the Monitoring the Future Study (MTF; as cited in Eisen, Zellman, & Murray, 2003).

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

Outcome measure	Study sample	Sample size ³ (students/ schools)	Author's findings from the study					
			Mean outcome ²		Mean difference ⁴ (Skills for Adolescence – comparison)	WWC calculations		
			Skills for Adolescence group	Comparison group		Effect size ⁵	Statistical significance ⁶ (at $\alpha = 0.05$)	Improvement index ⁷
Eisen, Zellman, & Murray, 2003 (randomized controlled trial)								
Alcohol use—lifetime	Grade 8	5316 to 5610/34	33.03	33.67	–0.64	–0.02	ns	–1
Alcohol use—last 30 days	Grade 8	5316 to 5610/34	77.15	76.82	0.33	0.01	ns	+0
Cigarettes smoking—lifetime	Grade 8	5316 to 5610/34	72.00	72.50	–0.50	–0.02	ns	–1
Cigarettes smoking— last 30 days	Grade 8	5316 to 5610/34	87.53	88.52	–0.99	–0.06	ns	–2
Marijuana use—lifetime	Grade 8	5316 to 5610/34	72.76	69.50	3.26	0.10	ns	+4
Marijuana use—last 30 days	Grade 8	5316 to 5610/34	88.68	86.21	2.47	0.14	ns	+5
Other illicit substances use—lifetime	Grade 8	5316 to 5610/34	81.05	81.56	–0.51	–0.02	ns	–1
Other illicit substances use—last 30 days	Grade 8	5316 to 5610/34	93.11	93.02	0.09	0.01	ns	+0
Binge drinking—last 30 days	Grade 8, baseline binge drinkers	5078 to 5359/34	73.00	63.00	10.00	0.28	Statistically significant	+11
Binge drinking—last 30 days	Grade 8, baseline nonbinge drinkers	237 to 251/34	88.00	88.00	0.00	0.00	ns	+0
Domain average⁸ for behavior						0.04	ns	+2

ns = not statistically significant

1. This appendix reports follow-up findings (measured toward the end of eighth grade) considered for effectiveness and the improvement index. Findings reported from the same study for the same students toward the end of seventh grade are not included in the ratings, but are reported in Appendix A4.1. An additional follow-up finding (at grade 8) for binge drinking for the entire sample is presented in Appendix A4.2.
2. The mean outcome for each study group is the percentage of participants who, based on self reporting, did not engage in a given type of problem behavior.
3. The number of students in the analytic sample varied across outcomes, ranging from 5,316 to 5,610 students. The calculation of effect size was based on the middle point of this range (5,463). The number of baseline binge drinkers and nonbinge drinkers was estimated for statistical significance calculations based on the reported proportions and sample size for eighth-grade students.

Appendix A3 Summary of study findings included in the rating for the behavior domain^a (continued)

4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The study presents the percentage of students who reported involvement in each type of problem behavior. The percentages reported in this appendix show the proportion of students with favorable outcomes (nonusers of tobacco, alcohol, and other illicit drugs). So the signs of the mean differences presented here are the reverse of those reported by the study authors.
5. Effect sizes were calculated using the odds ratio formula for binary outcomes, which were then converted to standardized mean differences (Hedges's g). For an explanation of the 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 about the clustering correction, see the [WWC Tutorial on Mismatch](#). See the [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate statistical significance. In the case of *Skills for Adolescence*, corrections for multiple comparisons were 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. The WWC-computed domain effect sizes for each study and for each domain across studies are a simple average rounded to two decimal places. The improvement indices are calculated from the average effect sizes.

Appendix A4.1 Summary of findings of short-term effects for the behavior domain¹

Outcome measure ³	Study sample	Sample size ⁴ (students/ schools)	Author's findings from the study		WWC calculations			
			Mean outcome ²		Mean difference ⁵ (<i>Skills for Adolescence</i> – comparison)	Effect size ⁶	Statistical significance ⁷ (at $\alpha = 0.05$)	Improvement index ⁸
			<i>Skills for Adolescence</i> group	Comparison group				
Eisen, Zellman, & Murray, 2003 (randomized controlled trial)								
Alcohol—lifetime	Grade 7	4,106 to 5,644/34	70.39	69.81	0.58	0.02	ns	+1
Alcohol—last 30 days	Grade 7	4,106 to 5,644/34	92.83	92.75	0.08	0.01	ns	+0
Binge drinking—last 30 days	Grade 7	4,106 to 5,644/34	96.85	96.42	0.43	0.08	ns	+3
Cigarettes—lifetime	Grade 7	4,106 to 5,644/34	71.78	73.31	-1.53	-0.05	ns	-2
Cigarettes—last 30 days	Grade 7	4,106 to 5,644/34	97.12	96.11	1.01	0.19	ns	+7
Marijuana—lifetime	Grade 7	4,106 to 5,644/34	90.57	88.24	2.33	0.15	ns	+6
Marijuana—last 30 days	Grade 7	4,106 to 5,644/34	95.72	94.56	1.16	0.15	ns	+6
Other illicit drugs—lifetime	Grade 7	4,106 to 5,644/34	92.97	93.64	-0.67	-0.07	ns	-3
Other illicit drugs—last 30 days	Grade 7	4,106 to 5,644/34	96.64	96.45	0.19	0.03	ns	+1
Alcohol—lifetime	Grade 7, Hispanic	nr/29	67.31	63.42	3.89	0.10	ns	+4
Alcohol—lifetime	Grade 7, non-Hispanic	nr/29	71.43	74.03	-2.6	-0.08	ns	-3
Alcohol—last 30 days	Grade 7, Hispanic	nr/29	93.08	89.64	3.44	0.27	Statistically significant	+11
Alcohol—last 30 days	Grade 7, non-Hispanic	nr/29	92.82	94.41	-1.59	-0.16	ns	-6

(continued)

Appendix A4.1 Summary of study findings included in the rating for the knowledge, attitudes, and values domain¹ (continued)

Outcome measure ³	Study sample	Sample size ⁴ (students/ schools)	Author's findings from the study					Statistical significance ⁷ (at $\alpha = 0.05$)	Improvement index ⁸
			Mean outcome ²		Mean difference ⁵ (<i>Skills for Adolescence</i> – comparison)	Effect size ⁶			
			<i>Skills for Adolescence</i> group	Comparison group					
Binge drinking—last 30 days	Grade 7, Hispanic	nr/29	96.79	93.77	3.02	0.42	Statistically significant	+16	
Binge drinking—last 30 days	Grade 7, non-Hispanic	nr/29	97.08	97.69	–0.61	–0.15	ns	–6	
Alcohol—last 30 days	Grade 7, past binge drinkers	nr/29	83.02	79.55	3.47	0.14	ns	+6	
Alcohol—last 30 days	Grade 7, past cigarette smokers	nr/29	91.98	87.21	4.77	0.32	Statistically significant	+12	
Alcohol—last 30 days	Grade 7, past marijuana users	nr/29	83.19	76.48	6.71	0.25	ns	+10	

ns = not statistically significant

nr = not reported

1. This appendix presents findings that fall in the behavior domain measured toward the end of seventh grade. Follow-up findings (measured toward the end of eighth grade) were used for rating purposes and are presented in Appendix A3.
2. The mean outcome for each group is the percentage of participants who, based on self reporting, did not engage in a given type of problem behavior.
3. The study also reported student outcomes for binge drinking by past cigarette smokers and binge drinkers and recent cigarette use by past marijuana users. These analyses were not reviewed due to severe attrition of schools from the analysis sample.
4. The number of students varied by outcome, ranging from 4,106 to 5,644 students. The effect size calculation was based on the middle point of this range (4,875).
5. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The study presents the percentage of students who reported involvement in each type of problem behavior. So the signs of the mean difference presented here are the reverse of those reported by the study authors.
6. Effect size was calculated using the odds ratio formula for binary outcomes, which were then converted to standardized mean differences (Hedges's *g*). 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. For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). See the [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate statistical significance. In the case of *Skills for Adolescence*, no corrections for clustering were 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.

Appendix A4.2 Summary of other findings for the behavior domain¹

Outcome measure	Study sample	Sample size (students/schools)	Author's findings from the study		WWC calculations			
			Mean outcome (standard deviation ²)		Mean difference ³ (<i>Skills for Adolescence</i> – comparison)	Effect size ⁴	Statistical significance ⁵ (at $\alpha = 0.05$)	Improvement index ⁶
			<i>Skills for Adolescence</i> group	Comparison group				
Eisen, Zellman, & Murray, 2003 (randomized controlled trial)								
Binge drinking—last 30 days	Grade 8	5316 to 5610/34	87.33	86.89	0.44	0.02	ns	+1

ns = not statistically significant

1. This appendix presents an additional finding, binge drinking, that falls in the behavior domain. This finding is presented here for the entire sample, while findings for binge drinking behavior of past binge drinkers and nondrinkers were used for rating purposes and are presented in Appendix A3.
2. 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.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The study presents the percentage of students who reported involvement in each type of problem behavior. The percentages reported in this appendix show the proportion of students with favorable outcomes (nonusers of tobacco, alcohol, and other illicit drugs). So the signs of the mean differences presented here are the reverse of those reported by the study authors.
4. For an explanation of the effect size calculation, please see the [Technical Details of WWC-Conducted Computations](#).
5. 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 about the clustering correction, see the [WWC Tutorial on Mismatch](#). See the [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate statistical significance. In the case of *Skills for Adolescence*, no corrections for clustering were needed.
6. 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 Skills for Adolescence rating for the behavior domain

The WWC rates interventions as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.¹

For the outcome domain of behavior, the WWC rated *Skills for Adolescence* as having potentially positive effects. It did not meet the criteria for positive effects, because both studies that met WWC standards with reservations did not use a strong design, according to WWC criteria. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Skills for Adolescence* 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. The one study on *Skills for Adolescence* that met WWC standards showed a statistically significant positive effect on one student outcome.

- 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.

Met. The WWC analysis found no indeterminate effects or statistically significant or substantively important negative effects.

Other ratings considered

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

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

Not met. *Skills for Adolescence* had only one study showing a statistically significant positive effect.

- 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.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain level effects. 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.