Appendix

Appendix A1.1 Study Characteristics: Kirby, 2006 (randomized controlled trial)

Characteristic	Description
Study citation	Kirby, P. C. (2006). *I CAN Learn® in Orleans Parish Public Schools: Effects on LEAP 8th grade math achievement, 2003–2004.* (New Orleans, LA: Ed-Cet, Inc.) *Additional source:* Kirby, P. C. (2004c). *I CAN Learn® in Orleans Parish Public Schools: Effects on LEAP 8th grade math achievement, 2003–2004.* (New Orleans, LA: Ed-Cet, Inc.)
Participants	The study sample included 2,400 eighth-grade students (1,082 *I CAN Learn®* students and 1,318 traditional mathematics students) and 57 math teachers (29 teachers taught only traditional classes; 15 taught only *I CAN Learn®* classes; 13 taught both classes) in 13 Orleans Parish Public Schools.[1] After school clerks entered student names and courses, the Pearson SASI® Basic Scheduling (3rd edition) software program used a random assignment algorithm to schedule the students into intervention and comparison classes. Teachers were not randomized to conditions. Only students with no special education classification were included in the analyses. About 96% of the students in the *I CAN Learn®* and traditional group were African-American, less than 1% were Hispanic, and less than 0.5% were Caucasian.[2] Of these students, approximately 50% of the *I CAN Learn®* group and 50% of the traditional class group were female.
Setting	The participating students were from the Orleans Parish public school system, which includes the city of New Orleans. The participating schools were selected for this study based on two criteria: the schools included both traditional and *I CAN Learn®* eighth-grade classes, and those classes included at least 20 students each.
Intervention	Students were taught using the *I CAN Learn®* mathematics curriculum. According to the author, *I CAN Learn®* students are expected to complete about 100 lessons during a school year. However, intervention students in this study began using *I CAN Learn®* in the second semester of the 2003–04 school year. Students completed 12.1 lessons on average, with time spent on *I CAN Learn®* ranging from using only the test prep module to completing 95 lessons. The intervention and evaluation occurred during the 2003–04 academic year.
Comparison	Comparison students were taught in traditional classes with the teacher serving as the primary deliverer of instruction. The author did not provide further information on the curriculum.
Primary outcomes and measurement	The primary outcome measure was the Louisiana Educational Assessment Program (LEAP) Grade 8 Mathematics Exam administered in spring 2004. For a more detailed description of this outcome measure, see Appendix A2.
Staff/teacher training	Information on teacher training was not provided.

1. The original sample included 14 schools. However, random assignment occurred in only 13 schools. For the purposes of this review, the study design and findings are reported for the subsample of 13 schools. Further, the study author conducted analyses of subsamples of students in addition to the total sample of 2,400 students in 13 schools. Those analyses focused on students of teachers who taught both conditions and students with matched pretest and posttest scores.
2. The study notes that these data were collected in the Orleans Parish Public Schools before Hurricane Katrina. The demographics in the school district have changed since the study was conducted.

Appendix A1.2 Study Characteristics: Kerstyn, 2001 (quasi-experimental design)

Characteristic	Description
Study citation	Kerstyn, C. (2001). *Evaluation of the I CAN Learn® mathematics classroom: First year of implementation (2000–2001 school year*). (Available from the Division of Instruction, Hillsborough County Public Schools, 901 East Kennedy Blvd., Tampa, FL 33602.)
Participants	The study sample[1] included 2,536 eighth-grade students in 116[2] classes (58 *I CAN Learn®* classes with 1,222 students, and 58 traditional classes with 1,314 students) in 36 Title I middle schools in the Hillsborough County Public School District in Florida. The study was limited to regular education students. All 116 classes were used in the analysis. Approximately 49% of the total sample were males (50% *I CAN Learn®*, 47% comparison) and 51% were females (50% *I CAN Learn®*, 53% comparison). Roughly 47% of the total sample qualified for the National School Lunch Program (49% *I CAN Learn®*, 44% comparison); 42% of the total sample were Caucasian (37% *I CAN Learn®*, 47% comparison); 27% were African-American (30% *I CAN Learn®*, 25% comparison); 31% were Hispanic (33% *I CAN Learn®*, 28% comparison) and 5% were of other racial classification (4% *I CAN Learn®*, 5% comparison). Within subgroups, a larger number of students in the *I CAN Learn®* classes were at the pre-algebra and algebra levels.
Setting	The participating students were from middle schools in the Hillsborough County Public School system in Florida, which includes the Tampa metro area.
Intervention	Students were taught using the *I CAN Learn®* mathematics curriculum. The author did not indicate how many of the lessons were required to be completed for the curriculum to be implemented as intended. The author indicated that the *I CAN Learn®* system was implemented in 45-, 50-, 80-, and 90-minute class periods. The intervention and evaluation occurred during the 2000–01 academic year.
Comparison	Comparison students were taught using a traditional instructional method. The author did not provide further information on the curriculum.
Primary outcomes and measurement	The primary outcome measure was the Florida Comprehensive Assessment Test (FCAT) Grade 8 Math Test administered in February 2001.[3] For a more detailed description of this outcome measure, see Appendix A2.
Staff/teacher training	Teachers in this study participated in training sessions on the use of the software and hardware, but not on use of the software in instruction.

1. The reported study sample was taken from Table 5 (page 9 of the study), while the reported demographic information was taken from Table 4 (page 8 of the study).
2. Originally the study included 59 matched pairs of *I CAN Learn®* and traditional classes, for a total of 118 classes; however, one of the pairs (MJ-3 #12) was dropped due to the traditional classroom transitioning to be an *I CAN Learn®* classroom in the second semester.
3. A Semester 1 exam was also administered to study participants at the end of the first semester. The results are not reported in Appendix A3 with the results of the FCAT, because the psychometric properties of the Semester 1 exam were not reported, and it measured the effect of the intervention at an earlier point than did the FCAT. Because the FCAT measured math achievement at the end of the school year, the WWC reasoned that the FCAT was a more appropriate and valid measure of the intervention's end-of-year effects.

Appendix A1.3 Study Characteristics: Kerstyn, 2002 (quasi-experimental design)

Characteristic	Description
Study citation	Kerstyn, C. (2002). *Evaluation of the I CAN Learn® mathematics classroom: Second year of implementation (2001–2002 school year*). (Available from the Division of Instruction, Hillsborough County Public Schools, 901 East Kennedy Blvd., Tampa, FL 33602.)
Participants	The study sample included 11,125 eighth-grade students in 597 classes (129 *I CAN Learn®* classes with 1,871 students and 468 traditionally taught mathematics classes with 9,254 students) enrolled in Algebra I, Algebra I Honors, MJ-3 pre-algebra, or MJ-3 Advanced math classes in the 36 middle schools in the Hillsborough County School District in Florida. Approximately 50% of the total sample were males (47% *I CAN Learn®*, 51% comparison) and 50% were females (53% *I CAN Learn®*, 49% comparison). Roughly 44% of the total sample qualified for the National School Lunch Program (65% *I CAN Learn®*, 39% comparison); 50% of the total sample were Caucasian (34% *I CAN Learn®*, 53% comparison); 21% were African-American (28% *I CAN Learn®*, 20% comparison); 24% were Hispanic (34% *I CAN Learn®*, 22% comparison) and 5% were of other racial classifications (4% *I CAN Learn®*, 5% comparison). Compared with classes using a traditional instruction method, a higher proportion of *I CAN Learn®* students were in the free and reduced-price lunch program and were from minority backgrounds (African-American and Hispanic).
Setting	The participating students were from middle schools in the Hillsborough County Public School system in Florida, which includes the Tampa metro area.
Intervention	Students were taught using the *I CAN Learn® Algebra* curriculum, which consists of 109 complete algebra lessons that students complete individually at their desks by accessing software through a number of DVD disks. The *I CAN Learn®* system is intended to be the primary source of instruction. The intervention and evaluation occurred during the 2001–02 academic year.
Comparison	Comparison students were taught using a traditional instruction method. The author did not provide further information on the method.
Primary outcomes and measurement	The primary outcome measure was the Florida Comprehensive Assessment Test (FCAT) Grade 8 Math Test administered in 2002. For a more detailed description of this outcome measure, see Appendix A2.
Staff/teacher training	Information on teacher training was not provided.

Appendix A1.4 Study Characteristics: Kirby, 2004a (randomized controlled trial with teacher-intervention confound problem)

Characteristic	Description
Study citation	Kirby, P. C. (2004a). *Comparison of I CAN Learn® and traditionally-taught 8th grade general math student performance on the California Standards Test, Spring 2004*. (New Orleans, LA: Ed-Cet, Inc.)
Participants	The study sample included 204 eighth-grade students enrolled in General Mathematics (91 in *I CAN Learn®* classes and 113 in traditionally-taught classes) in Bret Harte Middle School. The *I CAN Learn®* classes contained a higher proportion of African-American students, and a lower proportion of Hispanic and non-native English speaking students, than the comparison classes. Approximately 59% of the total sample were males (56% *I CAN Learn®*, 62% comparison) and 41% were females (44% *I CAN Learn®*, 38% comparison); 36% of the total sample were African-American (48% *I CAN Learn®*, 26% comparison), 29% were Hispanic (22% *I CAN Learn®*, 35% comparison), and 35% were reported as "all other" (30% *I CAN Learn®*, 39% comparison).
Setting	The participating students were from Bret Harte Middle School, which is one of five middle schools in Hayward Unified School District in Alameda County, California.
Intervention	Students were taught eighth-grade mathematics by one teacher using the *I CAN Learn® Pre-algebra* curriculum.[1] The *I CAN Learn®* lessons were chosen to align to the California Mathematics Standards. The teachers used the *I CAN Learn®* computer system as the primary system of instructional delivery and used the Classroom Manager class management/grade book system to identify students who needed additional instruction on specific concepts. The intervention and evaluation occurred during the 2003–04 academic year.
Comparison	Comparison students were taught in traditional classes, with the teacher as the primary deliverer of instruction, using a curriculum based on the state-adopted Glencoe pre-algebra textbook.
Primary outcomes and measurement	The primary outcome measure was the General Mathematics exam from the California Standards Test (CST) administered in 2004. For a more detailed description of this outcome measure, see Appendix A2.
Staff/teacher training	The intervention teacher received professional development in using the *I CAN Learn®* software and managing the instructional environment.

1. The study authors provided the WWC with the number of teachers in each condition.

Appendix A1.5 Study Characteristics: Kirby, 2004b (randomized controlled trial with teacher-intervention confound problem)

Characteristic	Description
Study citation	Kirby, P. C. (2004b). *Comparison of I CAN Learn® and traditionally-taught 8th grade student performance on the Georgia Criterion-Referenced Competency Test*. New Orleans, LA: Ed-Cet, Inc.
Participants	The study sample included 254 eighth-grade students (91 in *I CAN Learn®* classes and 163 in traditional classes) in Gilmer Middle School in Ellijay, Georgia who were not classified to receive special education or gifted services. School staff first stratified sample students based on prior year achievement and then randomly assigned the students to *I CAN Learn®* or traditional classes. Approximately 49% of the total sample were males (45% *I CAN Learn®*, 52% comparison) and 51% were females (55% *I CAN Learn®*, 48% comparison); 47% of the sample qualified for the National School Lunch Program (41% *I CAN Learn®*, 50% comparison); 93% were Caucasian (95% *I CAN Learn®*, 91% comparison), 0.5% were African-American (0% *I CAN Learn®*, 1% comparison), and 7% were Hispanic (6% *I CAN Learn®*, 8% comparison).
Setting	The participating students were from one middle school in the Gilmer County School District in northwestern Georgia.
Intervention	Students were taught by one teacher using the *I CAN Learn®* mathematics curriculum. The intervention teacher used the *I CAN Learn®* computer system as the primary system of instructional delivery and used the Classroom Manager class management/grade book system to identify students who needed additional instruction on specific concepts. The intervention and evaluation occurred during the 2003–04 academic year.
Comparison	Comparison students were taught in traditional classes, with the teacher serving as the primary deliverer of instruction. The author did not provide further information on the curriculum.
Primary outcomes and measurement	The primary outcome measure was the Georgia Criterion-Referenced Competency Test (GCRCT) Math Test administered in 2004. For a more detailed description of this outcome measure, see Appendix A2.
Staff/teacher training	The intervention teacher received professional development in using the *I CAN Learn®* software and managing the instructional environment.

Appendix A2 Outcome measures for the math achievement domain

Outcome measure	Description
Florida Comprehensive Assessment Test (FCAT) Grade 8 Math Test	The FCAT math test is a standardized measure that includes items related to all five content strands of Florida's *Sunshine State Standards* for mathematics: number sense, concepts, and operations; measurement; geometry and spatial sense; algebraic thinking; and data analysis and probability (as cited in Kerstyn, 2001; Kerstyn, 2002). Test content at grade 8 is evenly divided among these five content strands. Students are given 160 minutes to take the exam, which includes multiple-choice items, gridded-response items, and performance tasks. Test results are reported as scale scores, which range from 100 to 500.
General Mathematics exam from the California Standards Test (CST)	The General Mathematics CST for grade 8 is based on the California Mathematics Standards for grades 6 and 7 (as cited in Kirby, 2004a). The CST is administered to students in grades 8 and 9 who have not yet completed or are not enrolled in discipline-specific, standards-based math courses or who are enrolled in the first year of a multi-year Algebra I course. The Mathematics CST consists of 65 multiple-choice questions. Test results are reported as scale scores, which range from 150 to 600.
Louisiana Educational Assessment Program (LEAP) Grade 8 Mathematics Exam	The LEAP Mathematics test for grade 8 includes a multiple-choice section and an open-ended section for more complex tasks (as cited in Kirby, 2006). The test is aligned to the state's Comprehensive Curriculum and Grade Level Expectations. Six strands of the Louisiana Mathematics Framework are represented in LEAP: number and number relations; algebra; measurement; geometry; data analysis, probability, and discrete math; and patterns, relations, and functions. Test results are reported as scale scores, which range from 100 to 500.
Georgia Criterion-Referenced Competency Test (GCRCT) Math Test	The GCRCT is designed to measure how well students acquire the skills and knowledge described in the Georgia Performance Standards and the Quality Core Curriculum (as cited in Kirby, 2004b). The GCRCT Math Test contains 60 items in six areas: number sense and numeration; geometry and measurement; patterns and relationships/algebra; statistics and probability; computation and estimation; and problem solving. Test results are reported as scale scores, which range from 150 to 450.

Appendix A3 Summary of study findings included in the rating for the math achievement domain[1]

Outcome measure	Study sample	Sample size (clusters/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *I CAN Learn*® group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*I CAN Learn*® – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Kirby, 2006 (randomized controlled trial)[7]								
LEAP Math scale scores	Grade 8	13/2,400	295.30 (42.26)	278.82 (49.82)	16.48	0.35	Statistically significant	+14
Average for math achievement (Kirby, 2006)[8]						**0.35**	**Statistically significant**	**+14**
Kerstyn, 2001, Algebra 1 (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	16/350	351.38[9] (30.80)	344.46[9] (28.36)	6.92	0.23	ns	+9
Average for math achievement (Kerstyn, 2001, Algebra 1)[8]						**0.23**	**ns**	**+9**
Kerstyn, 2001, Algebra 1 Honors (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	16/336	372.99[9] (34.47)	373.73[9] (35.80)	–0.74	–0.02	ns	–1
Average for math achievement (Kerstyn, 2001, Algebra 1 Honors)[8]						**–0.02**	**ns**	**–1**
Kerstyn, 2001, MJ-3 pre-algebra (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	64/1,420	296.77[9] (38.23)	293.89[9] (38.09)	2.88	0.08	ns	+3
Average for math achievement (Kerstyn, 2001, MJ-3 pre-algebra)[8]						**0.08**	**ns**	**+3**
Kerstyn, 2001, MJ-3 Advanced (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	20/430	332.51[9] (31.19)	327.40[9] (29.60)	5.11	0.17	ns	+7
Average for math achievement (Kerstyn, 2001, MJ-3 Advanced)[8]						**0.17**	**ns**	**+7**

(continued)

Appendix A3 Summary of study findings included in the rating for the math achievement domain[1] *(continued)*

Outcome measure	Study sample	Sample size (clusters/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2]: *I CAN Learn*® group	Authors' findings from the study: Mean outcome (standard deviation)[2]: Comparison group	WWC calculations: Mean difference[3] (*I CAN Learn*® – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Kerstyn, 2002, Algebra 1 (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	66/1,195	347.43[10] (20.88)	349.79[11] (21.04)	–2.36	–0.11	ns	–5
Average for math achievement (Kerstyn, 2002, Algebra 1)[8]						**–0.11**	**ns**	**–5**
Kerstyn, 2002, Algebra 1 Honors (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	91/1,894	369.58[9] (27.84)	374.31[10] (27.61)	–4.73	–0.17	ns	–7
Average for math achievement (Kerstyn, 2002, Algebra 1 Honors)[8]						**–0.17**	**ns**	**–7**
Kerstyn, 2002, MJ-3 pre-algebra (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	328/5,957	293.81[9] (31.96)	289.90[10] (40.55)	3.91	0.10	Statistically significant	+4
Average for math achievement (Kerstyn, 2002, MJ-3 pre-algebra)[8]						**0.10**	**Statistically significant**	**+4**
Kerstyn, 2002, MJ-3 Advanced (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	112/2,079	329.93[9] (23.30)	331.04[10] (24.08)	–1.11	–0.05	ns	–2
Average for math achievement (Kerstyn, 2002, MJ-3 Advanced)[8]						**–0.05**	**ns**	**–2**
Kirby, 2004a (randomized controlled trial with teacher-intervention confound problem)[7]								
General Mathematics CST	Grade 8	1/204	315.58 (45.75)	299.73 (49.71)	15.85	0.33	Statistically significant	+13
Average for math achievement (Kirby, 2004a)[8]						**0.33**	**Statistically significant**	**+13**

(continued)

Appendix A3 Summary of study findings included in the rating for the math achievement domain[1] *(continued)*

Outcome measure	Study sample	Sample size (clusters/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *I CAN Learn®* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*I CAN Learn®* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Kirby, 2004b (randomized controlled trial with teacher-intervention confound problem)[7]								
GCRCT[12]	Grade 8	1/254	333.54 (35.72)	319.89 (31.73)	13.65	0.41	Statistically significant	+16
Average for math achievement (Kirby, 2004b)[8]						**0.41**	**Statistically significant**	**+16**
Domain average for math achievement across all studies[8]						**0.12**	**na**	**+5**

ns = not statistically significant

na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the math achievement domain. Subgroup findings from the same studies are not included in these ratings, but are reported in Appendix A4.
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.
4. For an explanation of the effect size calculation, see 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 the groups.
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 results favorable to the intervention group.
7. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the cases of Kerstyn (2001) and Kirby (2006), corrections for clustering were needed, so the significance levels may differ from those reported in the original studies. No other studies required corrections for clustering or multiple comparisons.
8. 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.
9. For Kerstyn (2001), the author reported both classroom-level and student-level posttest means and standard deviations. In order to compute a student-level effect size that could be considered for rating purposes, the WWC used the student-level data and corrected for clustering. For further details, please see Technical Details of WWC-Conducted Computations.
10. The intervention group value from Kerstyn (2002) is the intercept from the hierarchical linear modeling (HLM) model plus the program coefficient from the HLM analysis.
11. The control group mean from Kerstyn (2002) is the intercept from the HLM model.
12. The author reported results from the Georgia Criterion-Referenced Competency Test as scale scores and as criterion scores (that is, the percentage that passed the criterion score compared with the percentage that failed), but the WWC focused on the results from the scale scores because they contain more information than categorical scores.

Appendix A4 Summary of subgroup findings for the math achievement domain[1]

Outcome measure	Study sample	Sample size (clusters/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *I CAN Learn*® group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*I CAN Learn*® – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Kerstyn, 2002, MJ-3 pre-algebra standard curriculum[7] (quasi-experimental design)[8]								
FCAT mathematics	Grade 8	258/4,045	294.16[9] (33.46)	295.06[10] (33.91)	–0.90	–0.03	ns	–1
Kerstyn, 2002, MJ-3 pre-algebra FCAT-exempt[7] (quasi-experimental design)[8]								
FCAT mathematics	Grade 8	249/888	299.53[9] (37.62)	284.38[10] (52.87)	15.15	0.29	Statistically significant	+12

ns = not statistically significant

1. This appendix presents subgroup findings for measures that fall in math achievement. The standard curriculum and FCAT-exempt subgroup scores from Kerstyn (2002) are presented in Appendix A4, and these combined scores can be found in Appendix A3. Total group scores 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.
4. For an explanation of the effect size calculation, see 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 the groups.
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 results favorable to the intervention group.
7. Standard curriculum and FCAT-exempt groups are categorized by the state in terms of their impact on school accountability grades. The standard curriculum students are used in the school accountability grade, while the FCAT-exempt students (including ESE, LEP, and Homebound students) are not.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Kerstyn (2002), no correction for clustering was needed.
9. The intervention group value from Kerstyn (2002) is the intercept from the HLM model plus the program coefficient from the HLM analysis.
10. The control group mean from Kerstyn (2002) is the intercept from the HLM model.

Appendix A5 *I CAN Learn® Pre-Algebra* and *Algebra* rating for the math 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 math achievement, the WWC rated *I CAN Learn® Pre-Algebra* and *Algebra* as positive. The remaining ratings (potentially positive, mixed, no discernible effects, potentially negative, and negative) were not considered, as *I CAN Learn®* was assigned the highest applicable rating.

Rating received
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.
Met. Four studies of *I CAN Learn®* showed statistically significant positive effects. Of those, one study had a strong design according to WWC standards.
AND
• Criterion 2: No studies showing statistically significant or substantively important *negative* effects.
Met. Four studies of *I CAN Learn®* showed statistically significant positive effects. The remaining study showed indeterminate effects. No studies showed statistically significant or substantively important negative effects.

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

Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size[1]		Extent of evidence[2]
		Schools	Students	
Math achievement	5	51	16,519	Medium to large

1. The samples in Kerstyn (2001) and Kerstyn (2002) comprised distinct samples of students within the same set of 36 schools.
2. A rating of "medium to large" requires at least two studies and two schools across studies in one domain, and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is "small."