What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

Students with Learning Disabilities January 2014

Spelling Mastery

Program Description[1]

Spelling Mastery is designed to explicitly teach spelling skills to students in grades 1 through 6. One of several Direct Instruction curricula from McGraw-Hill that precisely specify how to teach incremental content, *Spelling Mastery* includes phonemic, morphemic, and whole-word strategies:

- Teachers use the phonemic strategy to teach beginning students to use sound–symbol correspondence to spell words and then generalize these skills to spell word segments that follow regular patterns.
- Teachers use the morphemic strategy to teach more advanced students meaningful prefixes, suffixes, and word bases, and how to apply consistent rules to merge words and word segments.
- Teachers use the whole-word strategy to teach all students to spell common words with irregular letter sounds and to memorize these more difficult words.

Spelling Mastery provides teachers with fully-scripted lessons organized according to students' skill development. Instruction is provided in scaffolded steps to help students finish each concept before a new one is presented.

Report Contents

Research[2]

The What Works Clearinghouse (WWC) identified two studies of *Spelling Mastery* that both fall within the scope of the Students with Learning Disabilities topic area and meet WWC evidence standards. These two studies meet standards without reservations. Together, these studies included 70 students with learning disabilities in grades 2 through 4 in three elementary schools or receiving instruction at a summer program.

The WWC considers the extent of evidence for *Spelling Mastery* on the writing achievement of students with learning disabilities to be small for one outcome domain—writing. There were no studies that meet standards in the eight other domains, so we do not report on the effectiveness of *Spelling Mastery* for those domains in this intervention report. (See the Effectiveness Summary on p. 4 for further descriptions of all domains.)

Effectiveness

Spelling Mastery was found to have potentially positive effects on writing for students with learning disabilities.

Table 1. Summary of findings[3]

		Improvement index *(percentile points)*				
Outcome domain	Rating of effectiveness	Average	Range	Number of studies	Number of students	Extent of evidence
Writing	Potentially positive effects	+28	+15 to +42	2	70	Small

Program Information

Background

Spelling Mastery is distributed by McGraw-Hill Education, P.O. Box 182605, Columbus, OH 43218. Email: SEG_customerservice@mcgraw-hill.com. Website: https://www.mheonline.com. Telephone: (800) 334-7344.

Program details

Spelling Mastery can be implemented with individuals or whole classes. The program is designed for all academic levels: advanced, average, and academically challenged. It can also be used in multi-age classrooms, with English language learners, and with students with learning disabilities.

Spelling Mastery instruction is delivered in daily 15–20 minute sessions. The program has six levels (A through F), each with 60 to 120 lessons. In each lesson, teachers present an exercise, listen to student responses, and provide immediate feedback. Following an instructional sequence, teachers introduce sound–spelling relationships, morphographs (the smallest unit of meaning in written language), rules, and principles. Teachers integrate concepts cumulatively, with all strategies used and practiced over time. Students' progress is assessed using regularly scheduled tests. Each level of *Spelling Mastery* provides a teacher presentation book, teacher/series guide, student workbook, and software.

Spelling Mastery instructors use three strategies. Using the phonemic strategy, teachers instruct beginning spellers on sound–symbol (phoneme–grapheme) relationships. This strategy is emphasized in *Spelling Mastery* levels A and B. Using the morphemic strategy, teachers instruct older students to spell and blend bases and affixes (morphographs) to form words. This strategy is most efficient for multisyllabic words that are not easy to spell phonemically. Finally, using the whole-word strategy, teachers teach common, irregularly spelled words using memorization procedures and drills.

Cost

The *Spelling Mastery* materials vary in cost, depending on the level (A–F) of the program being implemented. Student workbooks range from $11 to $16 per student. Teaching materials include a teaching guide, which ranges from $172 to $202. Instructional software can also be purchased for each level at a cost of $75 for the single instructor version and, for wider use, $262 for the network version.

Research Summary

The WWC identified seven studies that investigated the effects of *Spelling Mastery* on the writing achievement for students with learning disabilities.

Table 2. Scope of reviewed research

Grade	2–4
Delivery method	Whole class
Program type	Curriculum

The WWC reviewed three of those studies against group design evidence standards. Two studies (Darch & Simpson, 1990; Darch, Eaves, Crowe, Simmons, & Conniff, 2006) are randomized controlled trials that meet WWC evidence standards without reservations. These studies are summarized in this report. One study does not meet WWC evidence standards.

The WWC reviewed one additional study against the pilot single-case design standards. This study does not meet WWC pilot single-case design standards.

The remaining three studies do not meet WWC eligibility screens for review in this topic area. Citations for all seven studies are in the References section, which begins on p. 5.

Summary of studies meeting WWC evidence standards without reservations

Darch and Simpson (1990) examined the effects of *Spelling Mastery* on writing achievement. The study included 28 fourth-grade students in four classes at a university-based summer program located in the rural southwestern United States. All students in the study were identified by their local districts as learning disabled based on federal and state guidelines. The 28 students were randomly assigned either to the *Spelling Mastery* or *Visual Imagery* interventions. Both the *Spelling Mastery* and *Visual Imagery* groups received instruction for 25–30 minutes a day for a total of 25 days across a 6-week period. The same spelling words were taught to the two groups.

Darch et al. (2006) studied the effects of *Spelling Mastery* on writing achievement. The study included 42 students in grades 2 through 4 at three elementary schools in eastern Alabama. All students in the study were classified as learning disabled by the school district's special education review committee in accordance with state and federal guidelines. The 42 students were randomly assigned either to the *Spelling Mastery* intervention or to a comparison group using a basal spelling series that employed worksheets, scripted lectures, and discussion to instruct students on common spelling patterns. Intervention and comparison students each participated in sixteen 30-minute sessions delivered over a 4-week period, with each session focusing on six words. The same spelling words were taught to the two groups.

Summary of studies meeting WWC evidence standards with reservations

No studies of *Spelling Mastery* met WWC evidence standards with reservations.

Effectiveness Summary

The WWC review of *Spelling Mastery* for the Students with Learning Disabilities topic area includes student outcomes in nine domains: alphabetics, reading fluency, reading comprehension, general reading achievement, mathematics, writing, science, social studies, and progressing in school. The two studies of *Spelling Mastery* that meet WWC evidence standards reported findings in one of the nine domains: writing. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *Spelling Mastery* on students with learning disabilities. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 14.

Summary of effectiveness for the writing domain

Two studies that meet WWC standards without reservations reported findings in the writing domain.

Darch and Simpson (1990) found statistically significant positive effects of *Spelling Mastery* on three measures of writing, including the Predictable Words and Unpredictable Words subtests of the Test of Written Spelling and an author-created spelling test. WWC calculations confirm these findings. The WWC characterizes these study findings as having a statistically significant positive effect.

Darch et al. (2006) did not find a statistically significant positive effect of *Spelling Mastery* on any of the four measures of writing. However, both the study authors and the WWC found substantively important positive effects on all four measures of writing: the Test of Written Spelling, and three tests created for the study (a Generalization test, a Transfer test, and a Maintenance test). The WWC characterizes these study findings as having a substantively important positive effect.

Thus, for the writing domain, one study found a statistically significant positive effect and one study found substantively important positive effects. This results in a rating of potentially positive effects, with a small extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the writing domain

Rating of effectiveness	Criteria met
Potentially positive effects *Evidence of a positive effect with no overriding contrary evidence.*	In the two studies that reported findings, the estimated impact of the intervention on outcomes in the *writing* domain was statistically significant and positive in one study and substantively important and positive in the other study.
Extent of evidence	**Criteria met**
Small	Two studies that included 70 students in three schools and a university-based summer program reported evidence of effectiveness in the *writing* domain.

References

Studies that meet WWC group design evidence standards without reservations

Darch, C., Eaves, R. C., Crowe, D. A., Simmons, K., & Conniff, A. (2006). Teaching spelling to students with learning disabilities: A comparison of rule-based strategies versus traditional instruction. *Journal of Direct Instruction, 6*(1), 1–16.

Darch, C., & Simpson, R. G. (1990). Effectiveness of visual imagery versus rule-based strategies in teaching spelling to learning disabled students. *Research in Rural Education, 7*(1), 61–70.

Studies that meet WWC group design evidence standards with reservations

None.

Study that does not meet WWC group design evidence standards

Darch, C., Kim, S., Johnson, S., & James, H. (2000). The strategic spelling skills of students with learning disabilities: The results of two studies. *Journal of Instructional Psychology, 27*(1), 15. The study does not meet WWC evidence standards because the measures of effectiveness cannot be attributed solely to the intervention—there was only one unit assigned to one or both conditions.

Studies that meet WWC pilot single-case design standards without reservations

None.

Studies that meet WWC pilot single-case design standards with reservations

None.

Study that does not meet WWC pilot single-case design standards

Owens, S. H., Fredrick, L. D., & Shippen, M. E. (2004). Training a paraprofessional to implement "spelling mastery" and examining its effectiveness for students with learning disabilities. *Journal of Direct Instruction, 4*(2), 153–172. The study does not meet WWC evidence standards because it is an ABAB or multiple baseline design with less than three data points in a phase.

Additional source:

Owens, S. H. (2001). Validated practices in spelling for students with learning disabilities. *Dissertation Abstracts International*, 62 (03A), 121-976.

Studies that are ineligible for review using the Students with Learning Disabilities Evidence Review Protocol

Kinder, D., Kubrina, R., & Marchand-Martella, N. (2005). Special education and direct instruction: An effective combination. *Journal of Direct Instruction, 5*(1), 1–36. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Preast, S. D. (2009). A study of direct instructional spelling strategies and their effect on students with special needs who are classified with mild mental disabilities. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 70*(1-A), 142. The study is ineligible for review because the WWC could not confirm that at least 50% of the sample was classified as students with learning disabilities.

Simonsen, F., & Gunter, L. (2001) Best practices in spelling instruction: A research summary. *Journal of Direct Instruction, 1*(2), 97–105. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Appendix A.1: Research details for Darch & Simpson (1990)

Darch, C., & Simpson, R. G. (1990). Effectiveness of visual imagery versus rule-based strategies in teaching spelling to learning disabled students. *Research in Rural Education, 7*(1), 61–70.

Table A1. Summary of findings

Meets WWC evidence standards without reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Writing	28 students	+40	Yes

Setting

The study was conducted in a university-based summer program in the rural southwestern United States.

Study sample

The sample for this study included 28 fourth-grade students with learning disabilities who had a history of low academic achievement. Twenty-one students were White, seven students were Black, and 16 were male. The mean age of the entire sample was 10 years 6 months. The mean full scale IQ for the entire sample was 92. The students were identified by their local districts as learning disabled based on both federal and state guidelines. The students were randomly assigned either to the *Spelling Mastery* or *Visual Imagery* interventions. The study does not specify the number of students in each condition, and the authors did not reply to a request for this information.[4] Four teachers were randomly assigned to teach the interventions, two to each intervention. The authors did not report any attrition and did not respond to a request for this information.

Intervention group

Students in the intervention group received instruction through lesson 40 of the Level C *Spelling Mastery* program (the full program has 137 lessons). The teacher followed scripted lessons included with the program to teach students how to recognize the meaning of the smallest word segment that has meaning (morphograph) and to identify these segments within words. Students then practiced spelling words composed of learned segments by identifying the segments and then spelling the whole words. Students were also taught spelling rules and practiced using the rules with relevant examples. Finally, teachers instructed on several spelling rules that enabled a spelling strategy to apply to many words, such as dropping the 'e' when adding 'ing' to the end of a word. Teachers provided *Spelling Mastery* instruction for 25–30 minutes per day for 25 days during a 6-week period.

Comparison group

Students in the comparison group received *Visual Imagery* and were presented with the same practice words as the students in the *Spelling Mastery* group. When a word was presented, students were directed to look at the word and apply a four-step *Visual Imagery* model:

- after covering the word, the teacher asked the students if they could see the image of the word in their minds;
- students were directed to imagine the word displayed on a large outdoor screen;
- students were asked to imagine each letter of the word pasted onto the screen; and
- students were told to remember the word by visualizing themselves nailing the letters of the word onto the screen.

Students were then asked to apply the *Visual Imagery* strategy to several other words without teacher assistance. This independent practice typically required 5–8 minutes. Teachers provided *Visual Imagery* instruction for 25–30 minutes per day for 25 days during a 6-week period.

Outcomes and measurement

Three tests in the writing domain were administered after the intervention was completed. The measures included a spelling test developed by the study authors and the Test of Written Spelling (TWS) Predictable Words and Unpredictable Words subtests (the authors do not report a composite TWS score). Three unit tests were also administered but are not included in this report.[5] For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

The senior author met with each of the four study teachers twice for roughly one hour each time. During these training sessions, the correct instructional procedures for the appropriate spelling program were modeled. Through role-playing, the teachers practiced the instructional procedures and were critiqued by the senior author. All teachers were judged to have mastered their respective instructional strategies.

Appendix A.2: Research details for Darch et al. (2006)

Darch, C., Eaves, R. C., Crowe, D. A., Simmons, K., & Conniff, A. (2006). Teaching spelling to students with learning disabilities: A comparison of rule-based strategies versus traditional instruction. *Journal of Direct Instruction, 6*(1), 1–16.

Table A2. Summary of findings

Meets WWC evidence standards without reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Writing	42 students	+16	No

Setting

Participants were attending specialized programs for students with learning disabilities in special education classrooms in eastern Alabama.

Study sample

Forty-four students were randomly assigned, and the analysis sample for this study included 42 students with learning disabilities at three elementary schools.

The students were classified as learning disabled by the school district's special education review committee in accordance with state and federal guidelines, including a one standard deviation discrepancy between tested intelligence and achievement. Students were between 8 and 12 years of age and in grades 2 through 4. Thirty-two students were male and 10 were female. Full-scale IQs ranged from 80 to 116, with an average of 87.

Within each school, the students were randomly assigned to instructional groups of three to six students each. Four of the eight groups included a total of 21 intervention students, and the other four groups included a total of 21 comparison students. The study does not specify the distribution of the eight intervention and comparison instructional groups across schools, and the authors did not respond to a request for this information. The analysis sample of 42 students excluded two students (one intervention student and one comparison student) who were randomly assigned but who were absent for much of the study period.

Intervention group

Students in the intervention group received instruction from Level D of the *Spelling Mastery* program (students did not receive the entire Level D program). The teacher followed scripted lessons, and each week's sessions focused on a different type of word (e.g., phonetically regular words, phonetically regular words with a prefix or suffix, irregular words). Lessons were organized around three instructional activities: (1) introduction of the spelling rule in whole group instruction time (5–7 minutes); (2) application of the spelling rule, also in whole-group instruction (10–12 minutes); and (3) independent worksheet practice while the teacher circulated to assist as necessary (5–7 minutes). The intervention was implemented in sixteen 30-minute sessions over a 4-week period with each session focusing on six words. All words taught had a fourth-grade level of difficulty.

Comparison group

The comparison group was taught the same words as the intervention group using traditional basal instruction (HBJ Spelling and Laidlaw Spelling). Lessons were organized around three instructional activities: (1) a pretest, self-corrected by students using whole-group instruction (5–7 minutes); (2) independent worksheet practice with the teacher circulating to assist as necessary (10–12 minutes); and (3) follow-up activities, including dictionary and handwriting skill training activities (5–7 minutes).

Outcomes and measurement

Four tests were administered after the intervention was completed, all in the writing domain. The measures included the Test of Written Spelling (TWS) including Predictable Words and Unpredictable Words subtests, a Generalization test, a Transfer test, and a Maintenance test. Four unit tests were also administered but are not included in this report (see endnote 5). For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Each of three teachers delivered spelling instruction to both the intervention and comparison groups. The teachers were graduate students enrolled in a masters program in learning disabilities from a mid-sized southeastern university. The study authors provided three 1-hour training sessions to each teacher. The first author supervised, critiqued, and evaluated the teachers as they role-played teaching the scripted lessons. Prior to implementation, teachers were required to demonstrate adequate performance during role-playing situations.

Appendix B: Outcome measures for the writing domain

Writing	
Spelling test (created by Darch & Simpson, 1990)	The researchers created an end-of-intervention spelling test by randomly selecting 25 of the words that were taught to all students. Students were given as much time as needed to complete the test (as cited in Darch & Simpson, 1990).
Test of Written Spelling (TWS)	The TWS is a 20-minute paper and pencil, norm-referenced test of spelling for students in grades 1–12 (raw scores). This test was administered on the day following the last day of instruction. The TWS includes two subtests: (1) the Predictable Words subtest measures spelling performance on phonetically regular words, and (2) the Unpredictable Words subtest measures spelling performance on phonetically irregular words. Darch and Simpson (1990) used TWS, and Darch et al. (2006) used TWS-3; both report the number of correctly spelled words (as cited in Darch & Simpson, 1990 and Darch et al., 2006).
Generalization test (created by Darch et al., 2006)	The Generalization test was created by the researchers for this study. It required students to write sentences including spelling words taught during the intervention period. A randomly-selected subset of words included in the curriculum was dictated to the students, who were then asked to write a sentence using each word. The reported measure is the percentage of correct words spelled multiplied by 100. This test was administered on the day following the last day of instruction (as cited in Darch et al., 2006). The reliability of this measure was not reported; however, the principal investigator for the Students with Learning Disabilities topic area determined that the overall test is likely to meet WWC reliability standards.[a]
Transfer test (created by Darch et al., 2006)	The Transfer test was created by the researchers for this study. It required students to spell 24 words not taught during intervention but with similar patterns to intervention words. The reported measure is the percentage of correct words spelled multiplied by 100. This test was administered on the last day of instruction (as cited in Darch et al., 2006). The reliability of this measure was not reported; however, the principal investigator for the Students with Learning Disabilities topic area determined that the overall test is likely to meet WWC reliability standards.[a]
Maintenance test (created by Darch et al., 2006)	The Maintenance test was created by the researchers for this study. It required students to spell 24 words randomly selected from lessons taught throughout the intervention. The reported measure is the percentage of correct words spelled multiplied by 100. The test was administered 1 week after the end of the intervention (as cited in Darch et al., 2006). The reliability of this measure was not reported; however, the principal investigator (PI) for the Students with Learning Disabilities topic area determined that the overall test is likely to meet WWC reliability standards.[a] Although this measure was administered about a week after the other measures, the PI determined that this test was administered soon enough after the conclusion of the intervention to be considered an immediate posttest.

[a] The Students with Learning Disabilities topic area requires each outcome to meet at least one of three reliability requirements: internal consistency of at least 0.60, temporal stability/test-retest reliability of at least 0.40, and inter-rater reliability of at least 0.50. Because each outcome measures whether written words are spelled correctly, the PI determined that coders could grade the measure with little subjectivity providing inter-rater reliability that would exceed the requirements.

Appendix C: Findings included in the rating for the writing domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Darch & Simpson, 1990[a]								
Spelling test (author created)	Grade 4	28 students	17.50 (3.80)	11.70 (4.10)	5.80	1.42	+42	< 0.01
TWS Predictable Words subtest	Grade 4	28 students	29.20 (4.20)	24.00 (4.20)	5.20	1.20	+39	< 0.01
TWS Unpredictable Words subtest	Grade 4	28 students	15.20 (4.20)	11.20 (2.00)	4.00	1.18	+38	< 0.01
Domain average for writing (Darch & Simpson, 1990)						**1.27**	**+40**	**< 0.01**
Darch et al., 2006[b]								
TWS-3	Grades 2–4	42 students	18.33 (7.24)	14.33 (9.82)	4.00	0.39	+15	0.14
Generalization test (author created)	Grades 2–4	42 students	7.19 (3.51)	5.14 (4.57)	2.05	0.43	+17	0.10
Transfer test (author created)	Grades 2–4	42 students	9.76 (3.74)	7.33 (6.06)	2.43	0.41	+16	0.12
Maintenance test (author created)	Grades 2–4	42 students	11.24 (5.16)	8.00 (6.75)	3.24	0.46	+18	0.08
Domain average for writing (Darch et al., 2006)						**0.42**	**+16**	**Not statistically significant**
Domain average for writing across all studies						**0.85**	**+28**	**Not statistically significant**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). 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 the study's domain average was determined by the WWC. na = not applicable. TWS = Test of Written Spelling.

[a] For Darch and Simpson (1990), the *p*-values presented here were reported in the original study. A correction for multiple comparisons was needed and confirmed the significance levels reported in the original study. This study is characterized as having a statistically significant positive effect because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant, accounting for multiple comparisons. For more information, please refer to the WWC Standards and Procedures Handbook, version 2.1, page 96.

[b] For Darch et al. (2006), the *p*-values presented here were reported in the original study. The WWC calculated the intervention group mean using a difference-in-differences approach by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest mean. This study is characterized as having a substantively important positive effect because the effect for all measures within the domain is positive and substantively important, and no effects are negative or statistically significant, accounting for multiple comparisons. For more information, please refer to the WWC Standards and Procedures Handbook, version 2.1, page 96.

Appendix D: Description of subtest findings for the writing domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Darch et al., 2006[a]								
TWS-3 Predictable Words subtest	Grades 2–4	42 students	10.57 (4.52)	7.33 (5.42)	3.24	0.57	+22	0.04
TWS-3 Unpredictable Words subtest	Grades 2–4	42 students	7.71 (3.08)	7.00 (4.74)	0.71	0.11	+5	0.16

Table Notes: The supplemental findings presented in this table are additional findings from the studies in this report that do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). 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. TWS-3 = Test of written spelling (third edition).

[a] For Darch et al. (2006), the *p*-values presented here were reported in the original study. A difference-in-differences adjustment was needed and resulted in a WWC-computed *p*-value of 0.07 for the *TWS Predictable Words subtest*; therefore, the WWC does not find the result to be statistically significant. The WWC calculated the intervention group mean using a difference-in-differences approach by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest mean.

Endnotes

[1] The descriptive information for this program was obtained from a publicly available source: the program's website (www.mheonline.com, downloaded December 2013). The WWC requests developers review the program description sections for accuracy from their perspective. 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 April 2013.

[2] The studies in this report were reviewed using the Evidence Standards from the WWC Procedures and Standards Handbook (version 2.1), along with those described in the Students with Learning Disabilities review protocol (version 2.2). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[3] For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 14. These improvement index numbers show the average and range of student-level improvement indices for all findings in the study.

[4] The WWC assumes a total of 14 students in each intervention group and that there was no student attrition.

[5] The unit tests administered during the study period do not meet Students with Learning Disabilities protocol requirements that outcomes be measured at the end of the intervention. The principal investigator determined that the final unit test, measuring only the final unit of instruction, was likely to be less representative of the intervention impacts than the overall spelling test and the Test of Written Spelling, and thus, this outcome is not reported.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2014, January). *Students with Learning Disabilities intervention report: Spelling Mastery*. 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 intervention 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 on p. 14.

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 intervention 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 intervention 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 on p. 14.

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.