

What Works Clearinghouse



Open Court Reading[®]

Program Description¹

Open Court Reading[®] is a core reading program for grades K–6 developed by SRA/McGraw-Hill that is designed to teach decoding, comprehension, inquiry, and writing in a logical progression. Part 1 of each unit, Preparing to Read, focuses on phonemic awareness, sounds and letters, phonics, fluency, and word knowledge. Part 2, Reading and Responding, emphasizes reading literature for understanding, comprehension, inquiry, and practical reading applications. Part 3, Language Arts, focuses on writing, spelling, grammar, usage, mechanics, and basic computer skills. SRA/McGraw-Hill revised *Open Court Reading*[®] and changed the name to *Imagine It!*[®] in 2007.² The study featured in this report evaluates *Open Court Reading*[®].

Research³

One study of *Open Court Reading*[®] that falls within the scope of the Adolescent Literacy review protocol meets What Works Clearinghouse (WWC) evidence standards without reservations. The study included more than 900 first-grade through fifth-grade students⁴ who attended five schools located in five states across the United States.

Based on this study, the WWC considers the extent of evidence for *Open Court Reading*[®] on adolescent readers to be small for one domain: comprehension. Three other domains are not reported in this intervention report. (See the Effectiveness Summary for further description of all domains.)

Effectiveness

Open Court Reading[®] was found to have potentially positive effects on comprehension for adolescent readers.

Report Contents	
Overview	p. 1
Program Information	p. 2
Research Summary	p. 3
Effectiveness Summary	p. 4
References	p. 5
Research Details for Each Study	p. 10
Outcome Measures for Each Domain	p. 12
Findings Included in the Rating for Each Outcome Domain	p. 13
Supplemental Findings for Each Outcome Domain	p. 14
Endnotes	p. 15
Rating Criteria	p. 16
Glossary of Terms	p. 17

Table 1. Summary of findings⁵

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of studies	Number of students	Extent of evidence
		Average	Range			
Comprehension	Potentially positive effects	+6	na	1	917	Small

na = not applicable

Program Information

Background

Open Court Reading® was developed for and is distributed by SRA/McGraw-Hill. SRA/McGraw-Hill revised the program and changed the name to *Imagine It!*® in 2007. Additions made to the program that came with the revision include increased instruction in vocabulary, writing and inquiry, stronger support for English language learners, and enhanced technology components. Address: McGraw-Hill Education, P.O. Box 182605, Columbus, OH 43218. Web: <https://www.mheonline.com/>. Telephone: (800) 334-7344.

Program details

Open Court Reading® materials are divided by grade and include the Reading, Phonemic Awareness and Phonics Kit (K); Sounds and Letters Workbook (K); Language Arts Skills Workbook (K); Big Books and Little Books (K–1); Language Arts Big Book (K–1); Pre-Decodable and Decodable Texts (K–3); Part 1 Lesson Cards (K–3); Desk Strips (K–3); Unit Assessment (K–6); Transparencies (K–6); Writer’s Workbook (K–6); Challenge Workbooks (K–6); Reteach Workbooks (K–6); Intervention Support (K–6); Phonics Skills Workbook (1); First and Second Readers (1–2); Reading and Phonics Package (1–3); Student Anthologies (1–6); Comprehension and Language Arts Workbook (1–6); Spelling and Vocabulary Skills Workbook (1–6); Inquiry Journal (2–6); and Language Arts Handbook (2–6). The Teacher’s Edition (K–6) contains information on providing systematic, explicit skills instruction centered on literature selections. Lesson plans indicate the goals and objectives for each lesson and detailed suggestions for carrying out the lessons.

Cost

The *Open Court Reading*® curriculum includes grade-specific student textbooks, workbooks, decodable books, and anthologies. *Open Court Reading*® *Online Professional Development* provides support for teachers, principals, reading specialists, and coaches. For details on specific product pricing, contact the program developer, SRA/McGraw-Hill.

Research Summary

Fifty-eight studies reviewed by the WWC investigated the effects of *Open Court Reading*® on adolescent readers. One study (Borman, Dowling, & Schneck, 2008) is a randomized controlled trial that meets WWC evidence standards without reservations. This study is summarized in this report. The remaining 57 studies do not meet either WWC eligibility screens or evidence standards. (See references beginning on p. 5 for citations for all 58 studies.)

Summary of study meeting WWC evidence standards without reservations

Borman et al. (2008) conducted a randomized controlled trial that examined the effects of *Open Court Reading*® on first-grade through fifth-grade students attending five schools from five states during the 2005–06 school year. At each school, classrooms were randomly assigned within each grade either to implement *Open Court Reading*® or to serve as the comparison group. The WWC based its effectiveness rating on findings from 917 students from grades 1–5 who participated in the study; 507 in the *Open Court Reading*® group and 410 in the comparison group. The study reported student outcomes after approximately seven months of program implementation.

Summary of studies meeting WWC evidence standards with reservations

No studies of *Open Court Reading*® meet WWC evidence standards with reservations.

Table 2. Scope of reviewed research

Grade	1, 2, 3, 4, 5
Delivery method	Whole class
Program type	Curriculum
Studies reviewed	58
Meets WWC standards without reservations	1 study
Meets WWC standards with reservations	0 studies

Effectiveness Summary

The WWC review of interventions for Adolescent Literacy addresses student outcomes in four domains: alphabets, reading fluency, comprehension, and general literacy achievement. The one study that contributes to the effectiveness rating in this report covers one domain: comprehension. The comprehension domain includes two constructs: reading comprehension and vocabulary development. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *Open Court Reading*® on adolescent readers.⁶ For a more detailed description of the rating and extent of evidence criteria, see the WWC Rating Criteria on p. 16.

Summary of effectiveness for the comprehension domain

One study reported findings in the comprehension domain.

Borman et al. (2008) found, and the WWC confirmed, statistically significant effects of *Open Court Reading*® on the Reading Composite score of the Comprehensive Test of Basic Skills, 5th edition (CTBS/5) Terra Nova test.

Thus, for the comprehension domain, one study showed statistically significant positive effects of *Open Court Reading*®. 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 comprehension domain

Rating of effectiveness	Criteria met
Potentially positive effects <i>Evidence of a positive effect with no overriding contrary evidence.</i>	The review of <i>Open Court Reading</i> ® in the comprehension domain had one study showing statistically significant positive effects and no studies showing statistically significant or substantively important negative effects or indeterminate effects.
Extent of evidence	Criteria met
Small	The review of <i>Open Court Reading</i> ® in the comprehension domain was based on one study that included five schools and 917 students.

References

Study that meets WWC evidence standards without reservations

Borman, G. D., Dowling, N. M., & Schneck, C. (2008). A multi-site cluster randomized field trial of Open Court Reading. *Educational Evaluation and Policy Analysis, 30*(4), 389–407.

Studies that do not meet WWC evidence standards

Mitchell, S., & Wile, N. (2001). *2001 literacy program evaluation: A report of the evaluation of literacy programs in elementary and middle schools*. Portland, OR: Portland Public Schools, Research and Evaluation Department. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Stockard, J. (2010). Promoting reading achievement and countering the “fourth-grade slump”: The impact of direct instruction on reading achievement in fifth grade. *Journal of Education for Students Placed at Risk, 15*(3), 218–240. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Studies that are ineligible for review using the Adolescent Literacy Evidence Review Protocol

Altwerger, B., Arya, P., Jin, L., Jordan, N. L., Laster, B., Martens, P., . . . Wiltz, N. (2004). When research and mandates collide: The challenges and dilemmas of teacher education in the era of NCLB. *English Education, 36*(2), 119. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Alvarez, L., & Corn, J. (2008). Exchanging assessment for accountability: The implications of high-stakes reading assessments for English learners. *Language Arts, 85*(5), 354–365. The study is ineligible for review because it does not use a comparison group design or a single-case design.

American Federation of Teachers. (1999). *Building on the best, learning from what works: Five promising remedial reading intervention programs*. Washington, DC: American Federation of Teachers. 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.

Additional source:

American Federation of Teachers. (1998). *Building on the best, learning from what works: Seven promising reading and English language arts programs*. Washington, DC: American Federation of Teachers.

Arya, P., Laster, B., & Jin, L. (2005). An open look at the Open Court program. In B. Altwerger (Ed.), *Reading for profit: How the bottom line leaves kids behind* (pp. 128–141). Portsmouth, NH: Heinemann. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Association for Supervision and Curriculum Development, Council of Chief State School Officers, & McGraw-Hill Education. (2005). *Results with Open Court*. Columbus, OH: McGraw-Hill Education. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Barrett, A. J. (2007). Diffusion of innovation: A study of the impact of professional development practices on coherent implementation of Open Court Reading in the Los Angeles Unified School District. *Dissertation Abstracts International, 69*(01A), 122-84. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Barrett, T. J. (1995, November). *A comparison of two approaches to first grade phonics instruction in the Riverside Unified School District*. Paper presented at the annual meeting of the California Educational Research Association, Lake Tahoe, CA. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

- Briggs, K. L., & Clark, C. (1997). *Reading programs for students in the lower elementary grades: What does the research say?* Austin, TX: Texas Center for Educational Research. 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.
- Cannon, J. S., & Karoly, L. A. (1997). *Who is ahead and who is behind? Gaps in school readiness and student achievement in the early grades for California's children.* Santa Monica, CA: RAND Corporation. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Coles, G. (2000). "Direct, explicit, and systematic"—Bad reading science. *Language Arts*, 77(6), 543–545. 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.
- Colvin, L. (2002). Getting it right and making a difference: The instructional leadership strategies and programmatic materials that raise student achievement in low-socioeconomic urban elementary schools. *Dissertation Abstracts International*, 63(11A), 137. The study is ineligible for review because it does not examine the effectiveness of an intervention.
- Crowe, E. C., Connor, C. M., & Petscher, Y. (2009). Examining the core: Relations among reading curricula, poverty, and first through third grade reading achievement. *Journal of School Psychology*, 47(3), 187–214. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Florida Center for Reading Research. (2004). *Open Court Reading pre-K.* Tallahassee, FL: Author. 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.
- Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, 90(1), 37–55. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Gillam, R. B., & Carlile, R. M. (1997). Oral reading and story retelling of students with specific language impairment. *Language, Speech, & Hearing Services in Schools*, 28(1), 30–42. The study is ineligible for review because it does not examine the effectiveness of an intervention.
- Hayes, L. L. (2004). A comparison of two systematic approaches to phonics and spelling instruction in beginning reading: A basal phonics program and word study. *Dissertation Abstracts International*, 65(04A), 153. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Jordan, N. L. (2005). Basal readers and reading as socialization: What are children learning? *Language Arts*, 82(3), 204–213. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Land, R., & Moustafa, M. (2005). Scripted reading instruction: Help or hindrance? In B. Altwerger (Ed.), *Reading for profit: How the bottom line leaves kids behind* (pp. 63–77). Portsmouth, NH: Heinemann. The study is ineligible for review because it does not examine an intervention implemented in a way that falls within the scope of the review.
- Additional source:**
- Moustafa, M., & Land, R. (2001, November). *The effectiveness of "Open Court" on improving the reading achievement of economically-disadvantaged children.* Paper presented at the annual meeting of the National Council of Teachers of English, Baltimore, MD.
- Lauer, P., Akiba, M., Wilkerson, S., Apthorp, H., Snow, D., & Martin-Glenn, M. (2004). *Effectiveness of out-of-school-time strategies in assisting low-achieving students in reading and mathematics: A research synthesis.* Washington, DC: Institute of Education Sciences. 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.

- Lee, S. K., Ajayi, L., & Richards, R. (2007). Teachers' perceptions of the efficacy of the Open Court program for English proficient and English language learners. *Teacher Education Quarterly*, 34(3), 19–33. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Levin, J., Haertel, E., Kirst, M., & Williams, T. (2006). *Similar students, different results: Why do some schools do better? Additional findings: Elementary school curriculum program and API: A more detailed examination*. Palo Alto, CA: EdSource, Inc. The study is ineligible for review because it does not examine the effectiveness of an intervention.
- Maddahian, E. (2002). *A comparative study of second grade students' reading, language, and spelling gain scores for the Los Angeles Unified School District reading programs. Planning, assessment and research division publication*. Los Angeles, CA: Los Angeles Unified School District, Program Evaluation and Research Branch. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Manzo, K. K. (2004). L. A. students get reading by the book. *Education Week*, 24(3), 1–18. 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.
- Manzo, K. K. (2004). Page turner. *Teacher Magazine*, 16(3), 9–10. 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.
- Manzo, K. K. (2006). Reading program benefits some Calif. schools. *Education Week*, 25(34), 9. 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.
- Martens, P., & Goodman, Y. (2005). The influence of decodable texts on readers' strategies. In B. Altwerger (Ed.), *Reading for profit: How the bottom line leaves kids behind* (pp. 78–95). Portsmouth, NH: Heinemann. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- McGraw-Hill Companies, Business Roundtable, & National Association of Elementary School Principals. (2002). *Results with Open Court Reading*. Columbus, OH: McGraw-Hill Companies. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- McRae, D. J. (2002). *Research findings 2002*. DeSoto, TX: SRA/McGraw-Hill. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- McRae, D. J. (2002). *Test score gains for Open Court schools in California*. DeSoto, TX: SRA/McGraw-Hill. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Miners, Z. (2007). Open Court reading program: A Florida district NCLB success. *District Administration*, 43(3), 24. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Mora, E. R. (2002). Spanish speakers learning to read in English only classrooms: Language policy, beginning reading instruction, instructional strategies, and perseverance. *Dissertation Abstracts International*, 63(03A), 206. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Morris, J. B. (2002). The role of literacy coaches in implementing research based reading programs at low achieving schools. *Dissertation Abstracts International*, 63(03A), 153. The study is ineligible for review because it does not examine the effectiveness of an intervention.
- National Center for Education Research. (2008). *Effects of preschool curriculum programs on school readiness: Report from the Preschool Curriculum Evaluation Research initiative*. Washington, DC: Institute of Education Sciences. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

- Newkirk, T. (2002). Reading and the limits of science. *Education Week*, 21(32), 39. 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.
- O'Brien, D. M., & Ware, A. M. (2002). Implementing research-based reading programs in the Fort Worth Independent School District. *Journal of Education for Students Placed at Risk*, 7(2), 167–195. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Oliver, D., & Maddahian, E. (2002). *K–3 district reading plan evaluation: Year 2 report*. Los Angeles, CA: Los Angeles Unified School District, Program Evaluation and Research Branch. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Parkman, L. L. (2011). An investigation of the impact of early intervention-reading programs on the academic achievement of third-grade students in reading. *Dissertation Abstracts International*, 71(10A), 3601. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Pascopella, A. (2004). Struggling English language learner might be at advantage. *District Administration*, 40(4), 21. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Pease-Alvarez, L., & Samway, K. D. (2008). Negotiating a top-down reading program mandate: The experiences of one school. *Language Arts*, 86(1), 32–41. The study is ineligible for review because it does not examine the effectiveness of an intervention.
- Peck, S., & Serrano, A. M. (2002, April). “Open Court” and English language learners: Questions and strategies. Based on paper presented at the annual meeting of the American Association for Applied Linguistics, Salt Lake City, UT. The study is ineligible for review because it does not examine the effectiveness of an intervention.
- Rosenshine, B. (2002). Helping students from low-income homes read at grade level. *Journal of Education for Students Placed at Risk*, 7(2), 273–283. 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.
- Salasin, E. K. (2005). “Someone should know where we’re going”: Teacher negotiation of mandated literacy reform. *Dissertation Abstracts International*, 66(10A), 174. The study is ineligible for review because it does not examine the effectiveness of an intervention.
- Samway, K. D., & Pease-Alvarez, L. (2005). Teachers’ perspectives on Open Court. In B. Altwerger (Ed.), *Reading for profit: How the bottom line leaves kids behind* (pp. 142–155). Portsmouth, NH: Heinemann. The study is ineligible for review because it does not examine the effectiveness of an intervention.
- Simmons, D. C., Coyne, M. D., O’man, K., McDonagh, S., Harn, B. A., & Kame’enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*, 41(2), 158–173. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Skindrud, K., & Gersten, R. (2006). An evaluation of two contrasting approaches for improving reading achievement in a large urban district. *Elementary School Journal*, 106(5), 389–407. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- SRA/McGraw-Hill. (2008). *Oakland school district receives national recognition from SRA/McGraw-Hill for achievement in reading*. Columbus, OH: The McGraw-Hill Companies. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- SRA/McGraw-Hill. (2008). *Orlando elementary schools posting highest FCAT reading gains rely on SRA/McGraw-Hill literacy programs*. Columbus, OH: The McGraw-Hill Companies. The study is ineligible for review because it does not use a comparison group design or a single-case design.

- Stockard, J., & Englemann, K. (2010). The development of early academic success: The impact of Direct Instruction's "Reading Mastery." *Journal of Behavioral Assessment and Intervention for Children*, 1(1), 2–24. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- The Business Roundtable Education Initiative, & the National Association of Elementary School Principals. (2005). *Open Court results*. Columbus, OH: McGraw-Hill Education. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Torgesen, J. K., & King, R. (2000). *FCRR technical report #3: Improving the effectiveness of reading instruction in one elementary school: A description of the process*. Tallahassee, FL: Florida Center for Reading Research. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Twining, L. L. (2008). Raising student achievement at Eberman Elementary School with effective teaching strategies. *Dissertation Abstracts International*, 69(6A). The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Walk, R. S. (2008). Associations involving Open Court reading in kindergarten and student performance on standardized assessments in reading in a Tennessee school system. *Dissertation Abstracts International*, 69(3A), 923. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Webster, R. E., & Braswell, L. A. (1991). Curriculum bias and reading achievement test performance. *Psychology in the Schools*, 28(3), 193–199. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Wehby, J. H., Falk, K. B., Barton-Arwood, S., Lane, K. L., & Cooley, C. (2003). The impact of comprehensive reading instruction on the academic and social behavior of students with emotional and behavioral disorders. *Journal of Emotional & Behavioral Disorders*, 11(4), 225. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample includes less than 50% general education students.
- Wills, H., Kamps, D., Abbott, M., Bannister, H., & Kaufman, J. (2010). Classroom observations and effects of reading interventions for students at risk for emotional and behavioral disorders. *Behavioral Disorders*, 35(2), 103–119. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Appendix A: Research details for Borman et al., 2008

Borman, G. D., Dowling, N. M., & Schneck, C. (2008). A multi-site cluster randomized field trial of Open Court Reading. *Educational Evaluation and Policy Analysis, 30*(4), 389–407.

Table A. Summary of findings

Meets WWC evidence standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
Comprehension	49 classrooms/917 students	+6	Yes

Setting The study initially included six schools—one each in Florida, Georgia, Idaho, Indiana, North Carolina, and Texas. Two schools were from rural areas, two from suburban areas, and two from urban areas. The Georgia school dropped out of the study.

Study sample SRA/McGraw-Hill recruited a group of schools that had not previously used *Open Court Reading*® to participate in the study. The six schools that initially participated were given free *Open Court Reading*® materials as well as a training program for teachers and implementation support. At each school, classrooms were randomly assigned within each grade either to be enrolled in *Open Court Reading*® or to serve as the comparison group.

The initial sample consisted of 57 grade 1–5 classrooms containing a total of 1,099 students. The Georgia school dropped out of the study, which resulted in a loss of four classrooms in both the intervention and comparison groups. Some students were absent during the administration of the posttest. However, the resulting attrition rates of schools and classrooms were low.⁷

The analysis sample consisted of students in grades 1–5 and included 507 students in the 27 *Open Court Reading*® classrooms and 410 students in the 22 comparison classrooms. Participating students were more than 70% minority, and more than 75% were eligible for free or reduced price lunches. Fewer than 15% were English as a Second Language (ESL) students, and fewer than 10% were special education students.

Intervention group *Open Court Reading*® is a curriculum that includes textbooks, workbooks, decodable books, and anthologies. The curriculum consists of three main components: (a) Preparing to Read, (b) Reading and Responding, and (c) Language Arts. Teachers were given a teacher’s edition of the curriculum that included scripted direct instruction lessons; and diagnostic and assessment packages. The program is designed to be used for 2.5 hours per day with grades 1–2 and for two hours per day with grades 4–6. However, the authors report that external consultants observed that some teachers provided only 90 minutes of daily instruction. The intervention was implemented from fall to spring during the 2005–06 school year.

Comparison group The comparison classrooms used a “business as usual” curriculum and were instructed not to use *Open Court Reading*® or any of its materials. Principals mentioned that curricula currently in use by the comparison classrooms consisted of *Reading Street* by Scott Foresman, *Literacy Place* by Scholastic, *McGraw-Hill Reading* by MacMillan/McGraw-Hill, *Collections* by Harcourt, and *Trophies* by Harcourt. Consultants visited comparison classrooms and verified that they were not using *Open Court Reading*®.

Outcomes and measurement

For both the pretest (October 2005) and the posttest (May 2006), students took the Comprehensive Test of Basic Skills, 5th edition (CTBS/5) Terra Nova Reading Comprehension and Vocabulary subtests. A Reading Composite score also is reported, which is the average of these two subtest measures. For a more detailed description of these outcome measures, see Appendix B. Findings for the combined student sample on the Reading Composite score can be found in Appendix C. Additional findings reflecting subtest outcomes for the combined student sample and separately for grades 4 and 5 can be found in Appendix D.

Support for implementation

Teachers were provided training opportunities with external consultants, which consisted of two- to three-day summer workshops. In addition, the consultants, who had teaching experience and detailed knowledge of *Open Court Reading*® and were trained by SRA/McGraw-Hill, visited and observed classrooms and provided feedback to teachers.

Appendix B: Outcome measures for each domain

Comprehension	
<i>Comprehensive Test of Basic Skills (CTBS/5) Terra Nova Reading Composite score</i>	This assessment consists of two subtests, Reading Comprehension and Vocabulary, and combines selected-response items with constructed-response items that allow students to produce short and extended responses. The Reading Composite score is a simple average of the CTBS/5 Reading Comprehension and Vocabulary subtests described below (as cited in Borman et al., 2008).
Reading comprehension construct	
<i>Comprehensive Test of Basic Skills (CTBS/5) Terra Nova Reading Comprehension subtest</i>	This assessment combines selected-response items with constructed-response items that allow students to produce short and extended responses. The Reading Comprehension subtest items focus on five objectives: (a) oral comprehension of passages read aloud, (b) basic understanding of literal meanings of passages, (c) analyzing text, (d) evaluating and extending meaning, and (e) identifying reading strategies (as cited in Borman et al., 2008).
Vocabulary development construct	
<i>Comprehensive Test of Basic Skills (CTBS/5) Terra Nova Vocabulary subtest</i>	This assessment combines selected-response items with constructed-response items that allow students to produce short and extended responses. The Vocabulary subtest focuses on three objectives: (a) understanding word meaning, (b) identifying multi-meaning words, and (c) inferring words in context (as cited in Borman et al., 2008).

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

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Borman et al., 2008^a								
<i>CTBS/5: Reading Composite score</i>	Grades 1–5	49 classrooms/ 917 students	612.77 (49.36)	604.82 (28.55)	7.95	0.16	+6	< 0.05
Domain average for comprehension (Borman et al., 2008)						0.16	+6	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 change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. The statistical significance of the study’s domain average was determined by the WWC. The statistical significance of the domain average is the statistical significance of the hierarchical linear modeling (HLM) coefficient reported by the study authors for the Reading Composite score of the Comprehensive Test of Basic Skills, 5th edition (CTBS/5).

^a For Borman et al. (2008), no corrections for clustering or multiple comparisons were needed. The p-value presented here was reported in the original study. The effect size in the table is based on the HLM-adjusted results for combined grades 1–5 reported in Table 4 of Borman et al. (2008) (rather than the unadjusted means and standard deviations reported in Table 3 of Borman et al. [2008], since this effect size does not control for pre-intervention test scores, and the p-value does not account for the clustered sample design). The intervention mean equals the comparison mean plus the intervention coefficient from the HLM analysis.

Appendix D: Supplemental subtest findings for the comprehension domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Borman et al., 2008^a								
<i>CTBS/5: Vocabulary subtest</i>	Grades 1–5	49 classrooms/ 918 students	604.52 (55.85)	593.73 (55.49)	10.79	0.19	+8	< 0.05
<i>CTBS/5: Comprehension subtest</i>	Grades 1–5	49 classrooms/ 923 students	621.04 (48.11)	615.18 (47.92)	5.86	0.12	+5	< 0.05
<i>CTBS/5: Vocabulary subtest</i>	Grade 4	7 classrooms/ 133 students	631.08 (41.98)	631.34 (50.65)	–0.26	–0.01	0	> 0.05
<i>CTBS/5: Comprehension subtest</i>	Grade 4	7 classrooms/ 138 students	641.42 (41.79)	647.96 (42.76)	–6.54	–0.15	–6	> 0.05
<i>CTBS/5: Vocabulary subtest</i>	Grade 5	6 classrooms/ 105 students	660.18 (40.25)	636.17 (28.46)	24.01	0.69	+25	> 0.05
<i>CTBS/5: Comprehension subtest</i>	Grade 5	6 classrooms/ 106 students	662.53 (33.04)	650.84 (38.18)	11.69	0.32	+13	> 0.05

Table Notes: The supplemental findings presented in this table are additional subtest findings from the study 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 change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. CTBS/5 = Comprehensive Test of Basic Skills, 5th edition.

^a For the Borman et al. (2008) combined sample results (grades 1–5), a correction for multiple comparisons was needed, but the WWC could not apply this correction because exact p-values were not reported by the authors. An author query was not performed to request exact p-values for the subtest results because the significance level of the findings in this study is based on the Reading Composite score (reported in Appendix C) for which no correction for multiple comparisons was needed. The first two effect sizes in this table are based on the HLM-adjusted results for combined grades 1–5 reported in Table 4 of Borman et al. (2008). The intervention mean equals the comparison mean plus the intervention coefficient from the HLM analysis. The p-value levels presented for the first two outcomes were reported in the original study.

For the individual grade results (grade 4, grade 5), corrections for clustering and multiple comparisons were needed. When adjusted for clustering of students within classrooms, the WWC-calculated effects on the outcome measures listed in the table were not statistically significant. For the individual grade results, Borman et al. (2008) reported the unadjusted posttest means and standard deviations. The p-values were not reported in the original study. An author query was not performed to request these p-values because they would not have affected the significance level of the main results used in the WWC rating of the study. The reported p-values were computed by the WWC.

Endnotes

¹ The descriptive information for this program was obtained from a publicly available source: the program's website (<https://www.sraonline.com/productsamples.html?show=2&gid=342&tid=1>, downloaded June 2011). 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, and we incorporated feedback from the developer. 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 August 2011.

² This information was obtained from a publicly available source: <http://www.mcgraw-hill.com/releases/education/20070502.shtml> (downloaded August 2011). Additions made to the program that came with the 2007 revision and renaming to *Imagine It!*[®] include increased instruction in vocabulary, writing and inquiry, stronger support for English language learners, and enhanced technology components.

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

⁴ The Adolescent Literacy topic area reviews studies of interventions administered to students in grades 4–12 (or 9–18 years of age). For studies that include samples of students that span both the Adolescent Literacy (grades 4–12) and Beginning Reading (grades K–3) topic areas and cannot be adequately disaggregated by grade level, the Adolescent Literacy topic area will review any studies that include fifth-grade students or higher (for example, a combined sample of students from grades 3–6). The Beginning Reading topic area will review any studies that include only fourth-grade students and lower (for example, a combined sample of students from grades K–4). Findings in this Adolescent Literacy report are based on results from the HLM analysis for a combined sample of students from grades 1–5 (Borman et al., 2008) because the HLM coefficients that were used to calculate effect sizes on the outcome measures in Appendix C were not disaggregated by grade level in the original study.

⁵ For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p.16. These improvement index numbers show the average and range of student-level improvement indices for all findings across the study.

⁶ The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Borman et al. (2008), a correction for multiple comparisons was needed for the two subtests included in Appendix C, but the WWC could not apply this correction because exact *p*-values were not reported by the authors. The statistical significance of the domain average is the statistical significance of the HLM coefficient reported by the study authors for the Reading Composite score of the Comprehensive Test of Basic Skills, 5th edition.

⁷ Student attrition rates are unknown, as study authors did not attempt to match individual students from pretest to posttest.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, August). *Adolescent Literacy intervention report: Open Court Reading*[®]. 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. 16.
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. 16.
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.