

Houghton Mifflin Reading®

No studies of *Houghton Mifflin Reading*® that fall within the scope of the Beginning Reading review protocol meet What Works Clearinghouse (WWC) group design standards. Because no studies meet WWC group design standards at this time, the WWC is unable to draw any conclusions based on research about the effectiveness or ineffectiveness of *Houghton Mifflin Reading*® on beginning readers in grades K–3. Research that meets WWC standards is needed to determine the effectiveness or ineffectiveness of this intervention.

Program Description¹

Houghton Mifflin Reading® is a reading program designed for grades K–6. The program provides step-by-step instruction in reading using Big Books (fiction and nonfiction literature), anthologies, Read Aloud books, and audio compact discs. The product is designed to be used as a full-year curriculum program with instruction on developing oral language, comprehension, phonemic awareness, decoding skills (phonics, analogy, context, and word recognition), fluency, reading comprehension, writing, spelling, and grammar. Instruction is organized by a set of themes (10 for grades K–1, and six for grades 2–6) with selected Big Books and other classroom activities to highlight each theme. Themes for grade 1, for example, include “Family and Friends,” “All Together Now,” and “Let’s Look Around!” This review of the program for the Beginning Reading topic area focuses on beginning readers in grades K–3.

Research²

The WWC identified 10 studies of *Houghton Mifflin Reading*® for beginning readers in grades K–3 that were published or released between 1983 and 2014.

Five studies are within the scope of the Beginning Reading review protocol but do not meet WWC group design standards.

- Three studies used a quasi-experimental design where measures of effectiveness could not be attributed solely to the intervention.
- Two studies used a quasi-experimental design and did not establish that the intervention group was comparable to the comparison group prior to the start of the intervention.

Five studies are out of the scope of the Beginning Reading review protocol for reasons other than study design.

- Three studies did not use a sample aligned with the protocol.
- Two studies did not examine the effectiveness of the intervention in a way that falls within the scope of the protocol.

References

Studies that do not meet WWC group design standards

- Barnes, S. (2007). *The effects of curriculum structure on the achievement of grade 3 and grade 5 mobile students as measured by the Maryland school assessment* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3257409) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Crummett, T. C. C. (2008). *A comparison of the efficacy of a traditional reading program versus a full time intervention reading program for at-risk second and third graders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3292162) The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.
- Di Leo, M. A. (2012). *Reading First/Bay State Reading Initiative: Public vs. private implementation—which produces the best results?* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3465194) The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.
- Ryder, R. J., Burton, J. L., & Silberg, A. (2006). Longitudinal study of direct instruction effects from first through third grades. *Journal of Educational Research*, 99(3), 180–191. The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.
- Swartz, J., & Johnston, K. (2003). *Efficacy study of Houghton Mifflin Reading: A legacy of literacy, final report*. Cambridge, MA: Abt Associates. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Studies that do not meet WWC pilot single-case design standards

None.

Studies that are ineligible for review using the Beginning Reading Evidence Review Protocol

- EDSTAR. (2002). *Analysis of the effects of using Houghton Mifflin Reading programs on reading test scores in Chicago Public Schools*. Raleigh, NC: EDSTAR, Inc. The study is ineligible for review because it does not use a sample aligned with the protocol.
- Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology*, 31(1), 1–29. The study is ineligible for review because it is out of scope of the protocol.
- Haager, D., Dhar, R., Moulton, M., & McMillan, S. (2009). *The California Reading First year 7 evaluation report*. Morgan Hill, CA: Educational Data Systems. The study is ineligible for review because it is out of scope of the protocol.

Additional sources:

- Haager, D., Dhar, R., Heimbichner, C., Moulton, M., & McMillan, S. (2008). *The California Reading First year 6 evaluation report*. Morgan Hill, CA: Educational Data Systems.
- Haager, D., Dhar, R., Moulton, M., & McMillan, S. (2006). *The California Reading First year 4 evaluation report*. Morgan Hill, CA: Educational Data Systems.
- Haager, D., Dhar, R., Moulton, M., & McMillan, S. (2008). *The California Reading First year 5 evaluation report*. Morgan Hill, CA: Educational Data Systems.
- Haager, D., Dhar, R., Moulton, M., & Varma, S. (2005). *The California Reading First year 3 evaluation report*. Morgan Hill, CA: Educational Data Systems

Joseph, K. C. (2011). *Comparison of Houghton Mifflin core reading program and corrective reading program effects on English language learners' reading fluency and comprehension in grades 4–6* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3449689) The study is ineligible for review because it does not use a sample aligned with the protocol.

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*. Mountain View, CA: Edsource. <http://files.eric.ed.gov/fulltext/ED491576.pdf>. The study is ineligible for review because it does not use a sample aligned with the protocol.

Endnotes

¹ The descriptive information for this program was obtained from a publicly available source: the program's website (<http://www.hmhco.com>, downloaded August 2014). The WWC requests developers to review the program description sections for accuracy from their perspective. The program description was provided to the developer in August 2014; 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.

² The literature search reflects documents publicly available by April 2014. This report has been updated to include reviews of four studies that have been released since 2008. (The previous report was released in September 2008.) Of the additional studies, two were not within the scope of the protocol, and two were within the scope of the protocol but did not meet WWC group design standards. A complete list and disposition of all studies reviewed are provided in the references. The studies in this report were reviewed using the Standards from the WWC Procedures and Standards Handbook (version 3.0), along with those described in the Beginning Reading review protocol (version 3.0). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse (2015, February). *Beginning Reading intervention report: Houghton Mifflin Reading*[®]. Retrieved from <http://whatworks.ed.gov>

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 Procedures and Standards Handbook (version 3.0).
Improvement index	Along a percentile distribution of individuals, the improvement index represents the gain or loss of the average individual due to the intervention. As the average individual 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 study participants 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 eligible study participants are randomly assigned to 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 Procedures and Standards Handbook (version 3.0).
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% ($p < .05$).
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 \(version 3.0\)](#) for additional details.