



WWC Intervention Report

A summary of findings from a systematic review of the evidence



Adolescent Literacy

November 2017

Prentice Hall Literature[®] (1989–2005)

This intervention report presents findings from a systematic review of *Prentice Hall Literature[®]* (1989–2005) conducted using the WWC Procedures and Standards Handbook (version 3.0) and the Adolescent Literacy review protocol (version 3.0). No studies of *Prentice Hall Literature[®]* (1989–2005) that fall within the scope of the Adolescent Literacy 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 *Prentice Hall Literature[®]* (1989–2005) on students in grades 6–12. Research that meets WWC design standards is needed to determine the effectiveness or ineffectiveness of this intervention.

Program Description¹

Prentice Hall Literature[®] (1989–2005) is an English language arts curriculum designed for students in grades 6–12 that focuses on building reading, writing, listening, viewing, speaking, and language skills. Multiple editions of this curriculum were released between 1989 and 2005, including *Prentice Hall Literature[®]* (1989) and *Prentice Hall Literature: Timeless Voices, Timeless Themes[®]* (2000, 2002, 2005). The WWC refers to each of these editions as *Prentice Hall Literature[®]* (1989–2005) in this intervention report.

Prentice Hall Literature[®] (1989–2005) is based on a textbook with passages from fiction and nonfiction texts. Every reading selection in the curriculum's textbook includes pre-reading, active reading, and post-reading activities. The curriculum is organized by themes (such as cultural diversity or American individualism), and units focus on a specific genre (such as poetry, prose, or drama). Throughout each lesson, the curriculum describes related teaching techniques, including direct explanation, modeling, guided practice, feedback, and application. Additional materials are available to supplement the textbook, and are designed to enrich instruction or provide additional practice. *Prentice Hall Literature: Timeless Voices, Timeless Themes[®]* (2000, 2002, 2005) used the same instructional format as the original edition, and introduced new texts as well as new supplementary materials that teachers can use to differentiate instruction.

The publisher redesigned the curriculum with the release of *Prentice Hall Literature: Penguin Edition[®]* (2007). This revised edition integrated differentiated instruction into the core curriculum, introduced new readings from Penguin Publishing, and included new features such as writing workshops and expanded assessments. Evidence about the effectiveness of this and later editions – *Prentice Hall Literature: Language and Literacy[®]* (2010), *Prentice Hall Literature: Common Core Edition[®]* (2012), and *Pearson Literature[®]* (2015) – are included in a separate report.²

Research³

The WWC identified three studies of *Prentice Hall Literature[®]* (1989–2005) for students in grades 6–12 that were published or released between 1989 and 2017.

Two studies are within the scope of the Adolescent Literacy review protocol but do not meet WWC group design standards. Both studies (Bailey, 2005; Gatti, n.d.) do not meet WWC group design standards because they used quasi-experimental designs in which the equivalence of the analytic intervention and comparison groups at baseline was necessary, but not demonstrated.

One study (*Prentice Hall Literature*, n.d.) is out of the scope of the Adolescent Literacy review protocol because it used an ineligible study design.

References

Studies that do not meet WWC group design standards

Bailey, W. M. (2005). *Prentice Hall Literature: Timeless Voices, Timeless Themes. Program efficacy study*. Orlando, FL: WMB & Associates. Retrieved from <http://www.phschool.com/> The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Gatti, G. G. (n.d.). *Prentice Hall Literature: Timeless Voices, Timeless Themes. National effect size study*. Upper Saddle River, NJ: Prentice Hall. Retrieved from <http://www.phschool.com/> The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Study that is ineligible for review using the Adolescent Literacy Evidence Review Protocol

Prentice Hall Literature. (n.d.). *Prentice Hall Literature: Timeless Voices, Timeless Themes: Research summary*. Upper Saddle River, NJ: Author. The study is ineligible for review because it does not use an eligible design.

Endnotes

¹ The What Works Clearinghouse (WWC) requests developers to review the intervention description sections for accuracy from their perspective. The WWC provided the developer with the intervention description in January 2017, and the WWC incorporated feedback from the developer, which included a description of the earlier versions of *Prentice Hall Literature*[®]. Further verification of the accuracy of the descriptive information for this intervention is beyond the scope of this review.

² The WWC published a separate intervention report entitled *Prentice Hall/Pearson Literature*[®] (2007–15), which covers more recent editions of this intervention: <https://ies.ed.gov/ncee/wwc/InterventionReport/681>. This separate report was generated because the developer noted that the curriculum fundamentally changed with the introduction of *Prentice Hall Literature: Penguin Edition*[®] (2007). The WWC confirmed the decision to separate evidence on *Prentice Hall Literature*[®] into two intervention reports with the WWC literacy content expert. The companion report covers *Prentice Hall Literature: Penguin Edition*[®] (2007), *Prentice Hall Literature*[®] (2010), *Prentice Hall Literature: Common Core Edition*[®] (2012), and *Pearson Literature*[®] (2015).

³ The literature search reflects documents publicly available by January 2017. Reviews of studies in this report used the standards from the WWC Procedures and Standards Handbook (version 3.0) and the Adolescent Literacy 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

What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2017, November). *Adolescent Literacy intervention report: Prentice Hall Literature*[®] (1989–2005). Retrieved from <https://whatworks.ed.gov/>

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all subjects initially assigned to the intervention and comparison groups. If a randomized controlled trial (RCT) or regression discontinuity design (RDD) study has high levels of attrition, the validity of the study results can be called into question. An RCT with high attrition cannot receive the highest rating of *Meets WWC Group Design Standards without Reservations*, but can receive a rating of *Meets WWC Group Design Standards with Reservations* if it establishes baseline equivalence of the analytic sample. Similarly, the highest rating an RDD with high attrition can receive is *Meets WWC RDD Standards with Reservations*.

For single-case design research, attrition occurs when an individual fails to complete all required phases or data points in an experiment, or when the case is a group and individuals leave the group. If a single-case design does not meet minimum requirements for phases and data points within phases, the study cannot receive the highest rating of *Meets WWC Pilot Single-Case Design Standards without Reservations*.

Baseline A point in time before the intervention was implemented in group design research and in regression discontinuity design studies. When a study is required to satisfy the baseline equivalence requirement, it must be done with characteristics of the analytic sample at baseline. In a single-case design experiment, the baseline condition is a period during which participants are not receiving the intervention.

Clustering adjustment An adjustment to the statistical significance of a finding when the units of assignment and analysis differ. When random assignment is carried out at the cluster level, outcomes for individual units within the same clusters may be correlated. When the analysis is conducted at the individual level rather than the cluster level, there is a mismatch between the unit of assignment and the unit of analysis, and this correlation must be accounted for when assessing the statistical significance of an impact estimate. If the correlation is not accounted for in a mismatched analysis, the study may be too likely to report statistically significant findings. To fairly assess an intervention's effects, in cases where study authors have not corrected for the clustering, the WWC applies an adjustment for clustering when reporting statistical significance.

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 method by which intervention and comparison groups are assigned (group design and regression discontinuity design) or the method by which an outcome measure is assessed repeatedly within and across different phases that are defined by the presence or absence of an intervention (single-case design). Designs eligible for WWC review are randomized controlled trials, quasi-experimental designs, regression discontinuity designs, and single-case designs.

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 analytic sample groups are similar on observed characteristics defined in the review area protocol.

Glossary of Terms

Extent of evidence An indication of how much evidence from group design studies supports the findings in an intervention report. The extent of evidence categorization for intervention reports focuses on the number and sizes of studies of the intervention in order to give an indication of how broadly findings may be applied to different settings. There are two extent of evidence categories: small and medium to large.

- **small:** includes only one study, or one school, or findings based on a total sample size of less than 350 students and 14 classrooms (assuming 25 students in a class)
- **medium to large:** includes more than one study, more than one school, and findings based on a total sample of at least 350 students or 14 classrooms

Gain scores The result of subtracting the pretest from the posttest for each individual in the sample. Some studies analyze gain scores instead of the unadjusted outcome measure as a method of accounting for the baseline measure when estimating the effect of an intervention. The WWC reviews and reports findings from analyses of gain scores, but gain scores do not satisfy the WWC's requirement for a statistical adjustment under the baseline equivalence requirement. This means that a study that must satisfy the baseline equivalence requirement and has baseline differences between 0.05 and 0.25 standard deviations *Does Not Meet WWC Group Design Standards* if the study's only adjustment for the baseline measure was in the construction of the gain score.

Group design A study design in which outcomes for a group receiving an intervention are compared to those for a group not receiving the intervention. Comparison group designs eligible for WWC review are randomized controlled trials and quasi-experimental designs.

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.

Intervention An educational program, product, practice, or policy aimed at improving student outcomes.

Intervention report A summary of the findings of the highest-quality research on a given program, product, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against design standards, and summarizes the findings of those that meet WWC design standards.

Multiple comparison adjustment An adjustment to the statistical significance of results to account for multiple comparisons in a group design study. The WWC uses the Benjamini-Hochberg (BH) correction to adjust the statistical significance of results within an outcome domain when study authors perform multiple hypothesis tests without adjusting the p -value. The BH correction is used in three types of situations: studies that tested multiple outcome measures in the same outcome domain with a single comparison group; studies that tested a given outcome measure with multiple comparison groups; and studies that tested multiple outcome measures in the same outcome domain with multiple comparison groups. Because repeated tests of highly correlated constructs will lead to a greater likelihood of mistakenly concluding that the impact was different from zero, in all three situations, the WWC uses the BH correction to reduce the possibility of making this error. The WWC makes separate adjustments for primary and secondary findings.

Outcome domain A group of closely-related outcomes. A domain is the organizing construct for a set of related outcomes through which studies claim effectiveness.

Glossary of Terms

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	For group design research, the WWC rates the effectiveness of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. For single-case design research, the WWC rates the effectiveness of an intervention in each domain based on the quality of the research design and the consistency of demonstrated effects.
Regression discontinuity design (RDD)	A design in which groups are created using a continuous scoring rule. For example, students may be assigned to a summer school program if they score below a preset point on a standardized test, or schools may be awarded a grant based on their score on an application. A regression line or curve is estimated for the intervention group and similarly for the comparison group, and an effect occurs if there is a discontinuity in the two regression lines at the cutoff.
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% ($p < .05$).
Study rating	The result of the WWC assessment of a study. The rating is based on the strength of the evidence of the effectiveness of the educational <i>intervention</i> . Studies are given a rating of <i>Meets WWC Design Standards without Reservations</i> , <i>Meets WWC Design Standards with Reservations</i> , or <i>Does Not Meet WWC Design Standards</i> , based on the assessment of the study against the appropriate design standards. The WWC has design standards for group design, single-case design, and regression discontinuity design studies.
Substantively important	A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.
Systematic review	A review of existing literature on a topic that is identified and reviewed using explicit methods. A WWC systematic review has five steps: 1) developing a review protocol; 2) searching the literature; 3) reviewing studies, including screening studies for eligibility, reviewing the methodological quality of each study, and reporting on high quality studies and their findings; 4) combining findings within and across studies; and, 5) summarizing the review.

Please see the [WWC Procedures and Standards Handbook \(version 3.0\)](#) for additional details.



An **intervention report** summarizes the findings of high-quality research on a given program, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against evidence standards, and summarizes the findings of those that meet standards.

This intervention report was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.