Accelerated Reader as a Mechanism for Independent Reading Assessment:
Reading Gains in Ninth Graders

Abstract

The relationship of Accelerated Reader (AR) practice tests to improved reading outcomes has not been established. Examining the utilization of AR by a particular subset of ninth graders can suggest whether those students perform differently at reading comprehension tasks than do their counterparts participating in other independent reading assessments. The opportunity for repeated, summative assessment has particular appeal in today’s climate of high-stakes testing, but can multivariate analysis of experimental and control groups could indicate factors associated with optimal gains.

Introduction

Reading is a fundamental skill that underpins all other academic success. Because grade-level reading ability is central to the accountability measures imposed on schools by the federal No Child Left Behind (NCLB) Act, many schools are looking for proven ways to raise reading scores on standardized tests.

Accelerated Reader, a computerized reading practice assessment developed and marketed by Advantage Learning Systems, now Renaissance Learning, has been used extensively in schools over the last two decades to encourage schoolchildren at all levels to read a variety of fiction and nonfiction titles.

The multiple-choice reading practice tests delivered via AR software give schoolchildren both latitude in the selection of independent reading materials and immediate feedback regarding their comprehension and recall of the text. The software also serves teachers by offering labor-saving, student-centered assessments, which provide information as to the ongoing success of instructional interventions. Upon completion of designated titles, students gain points within the class and school database by answering at least sixty percent of the questions correctly. Points earned are based on the number of correct answers and are related to the reading level and length of the book. Teachers can create their own quizzes for delivery through the AR interface. The reading practice tests purchased through Accelerated Reader and third-party vendors are designed to focus on literal comprehension rather than reader inference. They are tested to eliminate cultural specificity and are reviewed to ensure psychometric integrity (Topping and Sanders, 2000).

The testing interface reminds students exactly how many points they have accumulated that school year, praising passing performance, or, in the event of quiz failure, convey encouragement through statements designed to boost morale and confidence. Settings within the program are delivered so that teacher action is required for students to take the same practice quiz another time after failing a quiz, a default setting to discourage students from testing on a book they have not completed.

Adoption of the Accelerated Reader program affects school libraries. AR becomes a budgetary commitment as the library must purchase quizzes for titles added to the collection. At schools with high levels of Accelerated Reader integration, students may limit their reading to books with quizzes available, effectively limiting the resources in use to the subset of Accelerated Reader titles. In a best-case scenario, circulation of Accelerated Reader titles may cause overall
circulation to increase in response to competition for points. But examining the relationship between the use of the AR program and improved reading comprehension and test-taking can suggest whether that use of resources is instructionally justified.

**Review of the Literature**

Existing research analyzing the relationship between use of the Accelerated Reader program and reading gains finds different outcomes related to the grade levels and setting where AR is used. Many participating schools and districts report no significant difference between the reading gains of students using Accelerated Reader when contrasted with those students not using the program or even reflect less improvement among certain subgroups (Ross, Nunnery, & Goldfeder, 2004, Melton, Smothers, Anderson, Fulton, Replogle, & Thomas, 2004). Others report students using Accelerated Reader as displaying statistically significant gains when contrasted with paired counterparts at another school or when used by at-risk reading intervention students (Vollands et al. 2000). Using a longitudinal statewide dataset, Topping and Sanders (2000) could not establish a causal relationship between Accelerated Reader participation and improved reading skill independent from measures of teacher effectiveness. They did confirm that continued use of AR leads to an increase in the overall volume of reading, a variable related to improved reading test scores, but found that reading gains were significant only in students consistently testing above the 85th percentile on the practice quizzes, at what Renaissance Learning defines as "mastery" level.

Most extant work carried out by academics evaluating the efficacy of AR, including that conducted by Topping and Sanders (2000), has been underwritten by Renaissance Learning. Ross et al. (2004) attempt to structure their study to meet the rigorous experimental guidelines recently outlined by federal government, but Renaissance Learning's provision of Accelerated Reader materials and monthly in-school consulting free-of-charge on the part of the software company makes this use of the software far from typical.

There is also doubt as to whether AR use creates a culture of literacy. Topping and Fisher (2002) investigated how local use of the computerized reading practice relates to student perceptions of reading as measured by a motivation-to-read questionnaire. Many researchers (Topping and Sanders, 2000, Everhart, 2005, Ross et al. 2004) make use of Renaissance Learning benchmarks for categorizing implementation of the AR in three levels of intensity. Topping and Sanders, 2000 found schools with mid- and high-level implementations lead to improved "aesthetic enjoyment" as well as social motivation and recognition for reading skill evidenced through point accumulation, but that there was no correlation between AR points accumulated and an increase in either breadth of reading or student opinion about the reading.

**Methods**

The proposed methodology contrasts reading gains made by ninth-graders over the course of an academic year, comparing students using the Accelerated Reader quizzes to gauge comprehension with their counterparts not using the computerized reading practice. Students are also studied as an aggregate as well as subgroups based on demographic and performance history; Topping and Sanders (2000) and Melton et al (2004) sort students into quartiles by pre-test scores to examine efficacy within subgroups.

**Procedures**

This model presents a static group comparison, "a design in which a group which has experienced X is compared with one which has not, for the purpose of establishing the effect of
(Campbell & Stanley, p. 12). With the exception of honors students and students assigned to intervention English classes which are excluded from this model, ninth grade students are assigned by a randomized computer-scheduling program to language arts sections, providing a control group composed of students whose teachers used other assessment to gauge students' completion of independent reading requirements.

Analysis should provide some indication as to whether being required to complete the computer-delivered reading practice test correlates with increased gains in reading as measured by achievement test scores.

\[
\begin{align*}
X & \quad \quad \quad \quad 0 \\
& \quad \quad \quad \quad 0 \\
\end{align*}
\]

**Principle:** Analysis of covariance (ANCOVA) using regression will control for mediating variables to investigate the relationship between student background variables, a particular classroom computerized reading practice program, and reading gains.

**Type of Design:** Bivariate and multivariate data analysis using previously collected data.

**Independent variables:** At the public suburban high school studied, one ninth grade teacher makes use of Accelerated Reader as a mechanism for assessing independent reading. The overarching independent variable in the proposed study is the assignment of ninth grade students \((n > 1100)\) to that faculty member over the last seven years. For students in the AR classrooms, a further series of independent variables, also used by Topping and Sanders (2000) with regard to Accelerated Reader, will indicate whether there is a critical or optimum level of usefulness with regard to particular student populations depending on:

- the number of books read
- mean level of books read
- reading volume
- percent correct (p. 318).

The demographic factors most often used in educational research -- race, gender, and socioeconomic background -- will be used as mediating variables.

Expansion of the study beyond a single school would introduce variables like schools and school systems as well.

Topping and Sanders (2000) also created a composite variable they call "challenge," which is created by the relationship between students' reading levels as assigned by STAR testing and the reading level of the books for which students earn points. They found increased reading challenge was associated with decreased teacher effectiveness at all levels of percent correct, even when reading volume is high. The utility of this type of information is dubious, however, given that the majority of books are written at grade levels below the chronological designations for high schoolers, and students typically read below grade.

**Dependent variables:** Scores derived using another Renaissance Learning product, the STAR reading assessment, will be used as pre- and post-tests to determine the outcome measure, or dependent variable. Improvement, rather than students' overall end reading ability, is the measure under consideration.
At the public suburban high school studied, only one ninth-grade teachers makes use of Accelerated Reader as a mechanism for assessing independent reading. As Topping and Fisher (2002) note, Accelerated Reader implementations vary, and some schools complicate the relationship between AR points and extrinsic reading motivation through the controversial use of rewards and other incentives for participation. In this design, students in three of ten or so sections of ninth-grade Language Arts are held responsible for earning 15 AR points every nine weeks, or 60 points total over the course of the school year. Students that read on a lower level must successfully complete more quizzes in order to gain the same number of AR points and meet the requirement, which comprises ten percent of their quarterly grade. Other ninth grade teachers at the school use other assessments for independent reading, and many teachers restrict the range of materials students can use for "free choice" reading.

**Factors jeopardizing internal validity:** Topping and Sanders (2000) suggest that when records are aggregated over many students, positive and negative short-term fluctuations tend to cancel each other out, and "any effects, if discernable, are likely to be small" (p. 308).

The school is in an area with some transience, and unmatched data can belie sample attrition, another threat to internal validity. Also, because the amount of time students and schools had been using AR was not known, there is the potential for students new to AR to demonstrate, as Topping and Sanders (2000) caution, a sort of Hawthorne effect at the point of new implementation.

**Factors jeopardizing external validity:** The proposed study uses seven years of data from a single school, and any findings are not necessarily generalizable outside of this particular locale. The school in question is an adolescent literacy demonstration site, and students receive more explicit reading instruction than the norm. While the single-school population is not ideal, Topping and Sanders (2000) discuss the difficulties in amalgamating reading gains from schools where hardware, software, teacher training and support and other human factors can vary wildly.

The STAR reading assessment, used in this case to determine reading gains, also interacts with the AR program, which uses a similar interface, giving AR students an advantage in terms of test familiarity. Unlike AR, STAR is an adaptive test, and subsequent questions are determined by a student's initial performance. Vollands, Topping, and Evans (1999) describe AR as "specifically intended to have strong formative effects on subsequent learning" (p. 198), through a point system designed to "raise metacognitive awareness" (p. 199). Participation in the AR program, with its frequent multiple-choice assessments, could have some relationship to the test-taking used as an outcome measure, particularly given that the STAR tests are delivered via a similar computerized mechanism.

**Adequacy of statistical procedures used:** Data analysis will explore covariance through regression, with the caveat that here, as in any multivariable system, one variable can mask another. These sorts of system can result in Type II errors as the opportunity to reject the null hypothesis is lost.

**Anticipated Results:** This study begins with a null hypothesis that there is no relationship between use of AR and reading gains. Calculating effect sizes, contrasting gains on the part of the students in the treatment AR classroom with the ninth-grade population, will determine if participation in the reading program held significance as an effective instructional mechanism.
Design improvement: The use of a single school, grade level, and treatment classroom as a sample of convenience provide the largest limitations to the proposed study. Expanding the study to a variety of school environments would also increase overall reliability. Experimental design improvement at a single school might involve using alternative measures for reading comprehension assessment in some sections of the current AR teacher’s classroom, and having other ninth-grade English teachers to implement AR in some classes as a point of comparison.

Pre-test, post-test comparison proposed as an outcome measure would result in data that could be used to contrast the control and treatment groups, though even that design presents many interactions you cannot quantify. Instruction may differ between ninth-grade classroom teachers in ways other than the incorporation of AR, what Topping and Sanders (2000) "significant intra-school variation" (p. 306) or "noise" (p. 308) at the classroom level.

Another limitation is the reality that students enrolled in the treatment course might not have fulfilled the AR point requirement, and exclusion of students who failed to earn a minimum number of points may prove appropriate.

AR is often presented as a part of "whole-school" revitalization, and the Reading Renaissance training materials provided with the Accelerated Reader software imply that, without the social motivation outside a particular classroom, students will not be appropriately incentivized to participate. The "Pavlovian" sensitization of students associating reading with material reward has been criticized (Chenoweth, 2001, p. 51); Vollands et al. (1999) found that the use of incentives was not critical to the success of the program in improving reading scores.

Significance of the Study

Most studies of Accelerated Reader focus on younger children, though Accelerated Reader is marketed to schools at all levels. In an outcome-driven educational environment, examining whether high school students using the computerized self-assessment can be said to demonstrate improved reading and test-taking skills can help determine whether to continue this ongoing expenditure or redirect funds for methods with better demonstrated instructional impact.

Chenoweth (2001) notes that all existing research failed to meet standards of scientific rigor, an implication that corresponds with the recent emphasis on establishing efficacy through experimental design and independent research to ensure schools can continue to pay for software components with federal funding such as Title I.
Works Cited


