Learning and Instruction, 17 Does Practice Make Perfect? Independent Reading Quantity, Quality and Student Achievement. 2007

The article you are about to read was reprinted with permission from the prestigious peer-reviewed journal entitled *Learning and Instruction*.

Does practice make perfect? Independent reading quantity, quality and student achievement

K.J. Topping a,*, J. Samuels b, T. Paul c,1

a School of Education, University of Dundee, Gardyne Road, Dundee DD5 1NY, Scotland, UK
b Departments of Educational Psychology and Curriculum and Instruction, University of Minnesota, MN, USA
c Renaissance Learning, Inc., Madison, WI, USA

Abstract

Does reading practice make perfect? Or is reading achievement related to the quality of practice as well as the quantity? To answer these questions, data on 45,670 students in grades 1–12 who read over 3 million books were analyzed. Measures largely of quantity (engaged reading volume) and purely of quality (success in reading comprehension) showed a positive relationship with achievement gain at all levels of achievement. However, both high quantity and high quality in combination were necessary for high achievement gains, especially for older students. Both were weakly associated with student initial reading achievement, but more strongly associated with the classroom in which the student was enrolled, possibly suggesting the properties of teacher intervention in guiding independent reading were important. Implications for theory-building, research and practice are explored.

"Practice makes perfect". This well known saying has some theoretical and empirical support in the research literature. However, it might be a considerable simplification of reality, and potentially misleading. The current study explores this issue in the important area of reading skills. Many studies have found a positive correlation between the amount of reading practice and reading achievement. However, more able readers might tend to read more, or more reading practice might enhance reading achievement, or both might happen. Nevertheless, some studies indicate that one causal direction is from practice to achievement (e.g., Cunningham & Stanovich, 1997). However, simply increasing reading practice time might not consistently improve student reading achievement. Other mediating variables might include actual student engagement during practice time (Berliner, 1990) and pure quality (rather than largely quantity) of reading practice (Topping & Sanders, 2000).

Quality of reading practice can be assessed by indicators of successful comprehension at an appropriate level of challenge. It has been argued that better curriculum-based formative assessment data of this type are needed to inform
student and teacher judgments (e.g., Visscher & Coe, 2002). Computerized progress-monitoring systems can provide such data, informing and potentially enhancing student self-management or teacher intervention or both. Within the context of such a system, the current study investigated the relationships between largely quantity of reading practice, pure quality of comprehension, classroom placement, and reading achievement. Theories assuming that reading practice is homogeneous and proposing that reading practice necessarily leads to automaticity and consequently reading achievement gain are challenged by this study.

1. Previous research

1.1. Reading practice

Stanovich, West, and Harrison (1995) suggested that reading practice was a major determinant of both crystallized and fluid intelligence in later life. Many studies have found a high positive correlation between volume of reading practice (at school or home) and reading achievement (e.g., Anderson, Wilson, & Fielding, 1988; Organisation for Economic Co-operation and Development, 2002; Snow, Burns, & Griffin, 1998). A smaller number of studies evidence a causal direction from practice to achievement (e.g., Cunningham & Stanovich, 1991, 1997; Shany & Biemiller, 1995; Taylor, Frye, & Maruyama, 1990). In a review of the literature, Lewis and Samuels (2005) found that 45 of 49 studies suggested that increased volume of reading caused improvement in reading. However, these studies used very various measures of reading practice, of unknown inter-correlation, rendering synthesis problematic. Despite this evident variety, the assumption that “reading practice” is homogeneous pervades the literature, and consequently relevant theoretical models can appear somewhat simplistic.

1.2. Theoretical issues

Much of the literature on the effects of practice is concerned with retention of facts or relatively simple skills. Practice is seen as facilitating the transfer from working memory to long term memory, so consciously controlled processing becomes automatic processing (associated with higher speed and lower effort). Distributed practice is likely to be more effective than massed practice (Schneider, Dumais, & Shiffrin, 1984). However, reading is a complex skill. In the specific context of reading, information-processing models (cf. Rumelhart, 2004) propose that student practice in decoding leads to automaticity in decoding which frees processing capacity to attend to higher order skills such as comprehension. This proposal appears to neglect two issues: (1) the effect of the challenge or difficulty (relative to the reader’s current ability) of the text currently being read, and (2) whether available processing capacity is actually deployed effectively.

If the text is slightly above the reader’s current independent reading level, much processing capacity may continue to be devoted to decoding, limiting successful comprehension. If it is well within that threshold, spare processing capacity might be available — but might not inevitably be devoted to higher order comprehension. Thus, McIntyre (1992) found that young children did not transfer to independent reading many of the skills they were able to deploy during direct instruction. Carver (1994) found that independent reading tended to include few unknown words and consequently did not result in vocabulary growth for either school or university students. Similar considerations might apply to texts of higher or lower ambiguity, the former requiring more spare processing capacity and processing efficiency than the latter (Rawson, 2004), particularly to separate task-relevant from task-irrelevant information (Haider & Frensch, 1999).

Consequently, the current study aimed to test a more complex model of the effects of reading practice on reading achievement, in which quality of independent reading practice (involving effective and accurate comprehension) was seen as equally important as quantity of reading practice or reading volume (which implies both having access to opportunities to practice and actually taking those opportunities).

1.3. Practical implications

If this model is valid, simply increasing time allocated to “reading practice” might not be effective in raising achievement (cf. Berliner, 1990), and indeed might have deleterious effects in other curricular areas owing to the opportunity costs involved. Elementary schools in the United States devote more hours per week on average to reading
instruction than any other country (Mullis, Martin, Gonzalez, & Kennedy, 2003), but appear some way down the international list in reading achievement. Studies of schemes for increasing time allocated to independent reading such as Sustained Silent Reading (SSR) show mixed results (Manning-Dowd, 1985; Yopp & Yopp, 2003). By contrast, Vollands, Topping, and Evans (1999) found achievement gains from improving the quality of engagement with literature by students, rather than merely increasing the quantity of time allocated to reading. This implies a need for more detailed monitoring of students’ independent reading practice.

1.4. Monitoring reading practice

Actually monitoring the behavior and progress of individual students in a consistent, reliable, valid, detailed and timely way presents large practical problems for a teacher. Computerized formative assessment systems for reading provide teachers with an activity- and progress-monitoring tool to enable them to achieve this otherwise daunting task. The study reported below utilized such a system for data capture.

Formative assessment involves frequent feedback on performance, and computerized systems enable this to be more frequent than otherwise. Research confirms the general importance of feedback (Butler & Winne, 1995; Lhyle & Kulhavy, 1987). Meta-analytic studies show substantial effect sizes (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Kluger & DeNisi, 1996; Walberg, 1999). Feedback through curriculum-based measurement or curriculum-based assessment facilitates frequent formative evaluation in an assess—teach—evaluate iterative cycle (Deno, 1985; Fuchs & Fuchs, 1986). Other reviews have confirmed the utility of formative assessment (e.g. Black & Wiliam, 1998; Crooks, 1988). Samuels and Wu (2005), among others, found immediate feedback had more impact than delayed feedback. After reviewing almost 8000 feedback studies, Hattie (1992) wrote: “The most powerful single modification that enhances achievement is feedback” (p. 11). This is corroborated by studies showing that actively monitoring reading practice leads to better gains for students (e.g., Anderson et al., 1988).

Feedback should enable greater student control over reading activity, in terms of management of appropriate challenge and other parameters of the reading process (Shapiro & Cole, 1994). Through this, meta-cognitive awareness might be heightened and feelings of self-efficacy as a learner enhanced (Schunk, 1994). However, information on learning is unlikely of itself to have effects if it is not acted upon. Bangert-Drowns et al. (1991) concluded from their meta-analysis that feedback improved learning effectiveness only where its “mindful reception” was assured. A computer system should free teacher “processing capacity” to intervene with students in a more timely and effective way — but that does not mean that such opportunities will be taken. However, the system itself indicates if teachers are indeed intervening effectively, and such data were captured in the current study.

2. Aims of the current study

In order to test a more complex theoretical model of the link between reading practice and achievement, the study reported below aimed to explore whether largely reading practice quantity (engaged reading volume) had an effect on reading achievement, or whether pure reading quality (in terms of successful comprehension) and classroom placement were also significant factors, analyzing data captured in a computer assessment system.

3. Method

3.1. Sampling and participants

Data for one school year were collected from 139 schools using the computer assessment system for reading in 24 of the United States (TX, AL, MS, ID, GA, IL, AR, AZ, DE, LA, NC, PA, CA, NM, OK, KY, SC, VA, WI, CO, MI, MO, NE, and UT). The geographical distribution of responses reflected distribution of program use at the time of the study, and a degree of bias to southern states is evident. Analyses reported in the Results section below were also conducted comparing Southern states and non-Southern states (as defined by the Census Bureau — www.census.gov/popest/geographic) and schools banded by socio-economic status, and no significant differences from the aggregate analysis found. Schools (n = 515) were contacted via a recruitment letter, with follow-up telephone calls. Responding schools were given instructions for uploading or emailing data directly to an independent third-party data collection
and storage company (this allows other independent researchers to verify the data and/or analyses reported here). The overall response rate was 27%. Schools in their first year or two of program implementation yielded more than half the data.

Overall, in the 139 schools there were 2365 classrooms containing between 12 and 30 students, slightly more males than females. All grade levels included students with a wide range of pretest NCE scores and the distribution of scores was similar in all grades (i.e. there was not a disproportionate number of remedial or learning disabled students in the sample — see Fig. 1). Racial origin was 63.8% White, 19.6% Black, 14.3% Hispanic/Latino, 1.1% Native American, and 1.2% Asian. According to the US census, the general population was 75.1% White, 12.3% Black, 12.5% Hispanic/Latino, .9% Native American, 3.1% Asian or Pacific Islander, and 5.5% Other (www.census.gov/prod/2001pubs/c2kbr01-1.pdf).

This yielded a database of the reading practice records of 45,670 students who had read over 3 million books in total within the system. The majority of data were for grades 1–6, but some data were available up to grade 12. First grade data were noisy — unsurprising as these students were variably emergent readers. In the results, grade 1 is included only when data are broken down by grade; in combined analyses across grades, grade 1 data are excluded to prevent skewing. Similarly, because of the probable differences in small-group implementation as compared to full-class implementation, a principled decision was taken to exclude classes/groups with fewer than 12 students from the classroom level analyses in this study.

3.2. Measures

3.2.1. STAR Reading

STAR Reading (Advantage Learning Systems, 1997) is a computerized adaptive standardized (norm-referenced) reading test for readers above first grade (100 words sight vocabulary). It continuously tailors each student’s test to his or her achievement level based on his/her responses to previous test items. It contains an item-bank, and can be used repeatedly during the year to measure student, class, or school progress. It yields grade equivalent, percentile, and normal curve equivalent (NCE) scores. Of these, NCE is based on an equal interval scale, consequently the most appropriate for aggregated statistical analysis, and used extensively below (although percentile rank is occasionally given). STAR Reading test—retest reliabilities range by grade from .85 to .95 (within 7 days), with overall reliability of .94. Split-half reliabilities range from .89 to .93. STAR Reading correlates with other standardized tests, including the Iowa Test of Basic Skills (average \( r = .74 \) across Forms K, L, and M for grades 1–6) and the Stanford Achievement Test, Ninth Edition (average \( r = .69 \) for grades 1–6) (Renaissance Learning, 2003). The majority of schools participating in the current study administered the STAR Reading test at the beginning of the school year and again in late April and May.
3.2.2. Accelerated Reader

Accelerated Reader (AR) (Advantage Learning Systems, 1993) is a system for free-standing computer assisted individualized assessment and progress monitoring of comprehension of “real books”. Students individually select books from the more than 100,000 titles for which AR reading practice comprehension quizzes are available, and read independently at their own pace, at school and/or at home. On completion, they take the multiple-choice AR reading practice quiz for the book at the computer. Each book has a maximum point value according to its length, adjusted for text difficulty. When the student completes a quiz the program awards points up to this maximum, according to the proportion of correct responses. The default condition is that students can quiz on a book only once.

The AR reading practice quizzes are brief. They primarily assess literal comprehension rather than idiosyncratic reader inferences or other more complex responses (which might be culturally specific), and do not measure all relevant reading behaviors in school or elsewhere. At a simplistic level, the AR reading practice quiz helps ensure the student has actually read the book, since reading choices are individualized and any attempt to copy quiz responses from another reader is both difficult and pointless. AR reading practice quizzes comprise 5, 10 or 20 questions (depending on the length of the book) on important facts or events relevant to the story grammar, each offering four response options. (For example, one question and response options for the book “Born Free”: Why do elephants sometimes kill lions? A. They regard lions as the only enemies of their young; B. Elephants often kill animals that encroach on their territory; C. Lions deprive elephants of their food supply; D. When an elephant is cornered, it becomes very aggressive.) Systematic procedures for quiz development assure question consistency and reliability to an extent beyond the capacity of a class teacher. Consequently, AR points gained are a relatively stable and accurate measure of words read and comprehended (across a variable number of reading occasions and number of books of varying length), and therefore a useful measure of volume of engaged and effective reading practice. (There are other kinds of AR quizzes available, but these are less psychometrically stable.)

In addition to tracking reading practice on different kinds of material, AR tracks, analyzes and automatically updates percent correct on each quiz taken, points, average percent correct (APC), words read, book titles, book readability levels, and fiction or non-fiction. The system identifies students who are not reading successfully, via his/her accumulating points (largely quantity metric) and APC (pure quality metric). The software designers recommend that teachers target an average quiz success rate of 85% correct as optimal, with either independent or supported reading. Automatically computer generated “At Risk” reports flag a need for the teacher to intervene (through informal reading counseling or other guidance) with any student whose reading activities appear currently ineffective, subsequently using the AR system to track the effect of their intervention. This might include students reading at high as well as low levels. AR also allows teachers to establish individual student point, percent correct, and book level goals. The software tracks how well each student is doing with respect to his/her own goals. Personalized student goals establish clear expectations and provide checkpoints to measure progress and trigger intervention. A number of meta-analytic studies report moderate effect sizes for goal setting, ranging from .46 to .55 (Lipsey & Wilson, 1993; Walberg, 1999; Wise & Okey, 1983).

Teachers may also choose to allow students to quiz on books read to them and with them, as well as those read independently and silently, especially in the case of emergent or delayed readers (the software tracks these three different types of reading practice). Where the program is used with class-wide, selective or elective peer tutoring (Topping, 2001), both assisting and assisted participants may subsequently self-assess their comprehension of the book. For emergently literate tutees, AR provides digitized speech output quizzes.

3.3. Variables

Included at the student level of analysis were variables such as points gained and Average Percent Correct (APC) on AR reading practice quizzes, with STAR Reading test NCE scores. The AR point value for each book is based on the number of words in the book and its readability level derived from the ATOS formula (which correlates highly with Flesch-Kincaid and Lexile levels) applied to a scan of the entire book (Renaissance Learning, 2001), thus

\[
\text{AR points} = \frac{\text{Words in book}}{100,000} \times (10 + \text{Readability level})
\]
The total number of points earned by a student is a proxy for Engaged Reading Volume (ERV) — but not for mere exposure to books, nor for time allocated to reading but not used for reading. ERV is an indicator not just of words read, but of difficulty of words read successfully. It includes two elements of quality. The maximum point value focuses upon length of the book, but this is adjusted for text difficulty. On quiz completion, points are awarded up to this maximum according to the proportion of correct responses. Both of these adjustments are necessary, as words of unknown difficulty read unsuccessfully would be unlikely to correlate with any outcomes and are not of interest. This measure might be considered superior to alternative measures, which assay reading volume but do not consider how well or even whether time has been used. Thus, ERV is a measure of quantity with elements of quality — a measure of volume of engaged and effective reading practice. By contrast, APC is a pure indicator of quality, but based upon a different metric, encompassing percent correct averaged over all books. There is no suggestion that these two measures are orthogonally related.

The portion of variance in individual student reading time that remained unexplained was the classroom placement variable. This might be partly attributable to the student’s teacher, and his/her inclination or ability to pursue a lower APC. However, it might also have included contextual variables of relevance (such as availability of a wide range of reading material and extent to which peers set good models of independent reading).

3.4. Data analysis

Most analyses explored absolute levels or changes in NCE within or between groups. The general assumption was that “normal” growth from the beginning to end of a year in NCE (or percentile rank) is zero. This might have been a somewhat risky assumption in first grade, and possibly for other grades (e.g. sixth grade, when students are likely to be changing schools). One concern with using change scores is controlling for regression to the mean. The corrections for regression to the mean proposed by Campbell and Kenny (1999) were utilized here, and change scores given are residualized change scores, which allow for differences in pretest scores and give more conservative estimates of effects. Interactions between the variables of ERV (reading quantity) and APC (reading quality) have recently been shown to be salient indicators of implementation fidelity and influences on reading achievement (Borman & Dowling, 2004; Topping, Samuels, & Paul, 2004, 2007), and these variables feature extensively in the analyses reported below.

4. Results

4.1. Quantity of reading

The relationship between ERV and NCE change at different levels of reading achievement was explored (Table 1). A consistent positive relationship between ERV and reading achievement gain was evident at all levels of reading achievement.

<table>
<thead>
<tr>
<th>STAR percentiles</th>
<th>n</th>
<th>ERV r</th>
<th>APC r</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–10</td>
<td>6114</td>
<td>.140*</td>
<td>.171*</td>
</tr>
<tr>
<td>11–20</td>
<td>5082</td>
<td>.167*</td>
<td>.237*</td>
</tr>
<tr>
<td>21–30</td>
<td>4320</td>
<td>.155*</td>
<td>.236*</td>
</tr>
<tr>
<td>31–40</td>
<td>4537</td>
<td>.144*</td>
<td>.232*</td>
</tr>
<tr>
<td>41–50</td>
<td>5049</td>
<td>.167*</td>
<td>.263*</td>
</tr>
<tr>
<td>51–60</td>
<td>4135</td>
<td>.126*</td>
<td>.232*</td>
</tr>
<tr>
<td>61–70</td>
<td>3917</td>
<td>.098*</td>
<td>.172*</td>
</tr>
<tr>
<td>71–80</td>
<td>4062</td>
<td>.156*</td>
<td>.225*</td>
</tr>
<tr>
<td>81–90</td>
<td>4453</td>
<td>.131*</td>
<td>.215*</td>
</tr>
<tr>
<td>91–100</td>
<td>4001</td>
<td>.121*</td>
<td>.199*</td>
</tr>
<tr>
<td>All deciles</td>
<td>45,670</td>
<td>.137*</td>
<td>.197*</td>
</tr>
</tbody>
</table>

*p < .001.

# Grades 2–12, all classes.
achievement. This countered any proposal that high achieving readers tend to read more rather than vice versa, because students read about the same amount regardless of initial reading level. A grade-level analysis showed the same pattern for all grades.

The relationship between reading quantity and student initial achievement was further explored by examining how much of the variation in ERV was explained by differences in student reading achievement and the classroom in which they were placed. ERV was first regressed on initial NCE level to determine the variance in time attributable to student initial achievement (Table 2). Within each grade, the correlation between pretest score and ERV was small, ranging from .061 to .251. The variance in ERV due to pretest score ranged from .004 to .063, indicating that between .4% and 6.3% of the variance in reading time was explained by pretest score.

The relationship between the quantity of reading (ERV) and the classroom in which the student was placed was then investigated. The residuals from the previous regression were used to create a classroom average residual for each student, excluding the residual of that individual student. The individual student’s residual (unexplained variance) was then regressed on their classroom average residual (Table 2). The variance in ERV due to the teacher ranged from .101 in eleventh grade to .590 in first grade. The $R^2$s were substantially higher than those for initial student achievement, suggesting that the classroom placement accounted for more of the variance in reading achievement than initial student achievement.

The relationship between quantity of reading and student reading achievement gain across all grades was then investigated for linearity. Notwithstanding the positive relationship between ERV and reading achievement gain, this relationship was curvilinear. Beyond a certain point, there was evidence of diminishing marginal returns in achievement for time spent engaged in reading (see Fig. 2). This curvilinear relationship was similar at all grade levels and for both above- and below-average students.

4.2. Quality of reading

The main indicator of quality of reading (successful comprehension) in the RPD was APC. However, the sources of variance merited further exploration. The relationships between NCE gain, varying levels of APC, and initial student achievement are given in Table 1. APC was positively associated with reading growth at all achievement levels.

Further analyses explored the relationship between quality of reading (APC), classroom placement and student initial achievement (Table 3). As was the case with ERV, far more of the variance was explained by the student’s teacher than by initial NCE level. Again, this was an encouraging finding, as the former is presumably more manipulable than the latter. The relationship between APC and NCE change was also explored for curvilinearity, but was found to be largely linear, with no evidence of diminishing marginal gain up to 97% (and sample size above this margin was quite small).
4.3. Quantity and quality of reading

Subsequent analyses explored the interaction of quantity (ERV) and quality (APC), across the grades. A striking interaction between quantity and quality was evident. Gains in achievement were found only with the combination of higher ERV and higher APC. High ERV with low APC was rare. Students with the combination of low ERV and low APC did particularly poorly — actually seeming to regress relative to peers. Even moderate APC (75—79%) did not yield gains, even at moderate ERV. Not until the threshold of 80—84% APC was exceeded did higher ERV begin to yield higher achievement. This relationship was particularly pronounced in the higher grades (6—12) (see Tables 4 and 5).

4.4. Summary

Regarding quantity of reading, there was a moderate positive relationship between ERV and gains in reading achievement at all levels of student reading achievement, suggesting that initial reading level does not affect reading volume. Variance in ERV was attributable much more to the student’s teacher than to student initial reading achievement. The relationship between ERV and reading achievement was curvilinear, with evidence of diminishing marginal returns in achievement beyond a certain point. Concerning quality of reading, there was a positive relationship
between successful reading comprehension and gains in reading achievement at all levels of student reading achievement suggesting that the relationship between successful reading comprehension and gains in reading achievement is independent of initial reading ability. Variance in successful comprehension was also attributable much more to classroom placement than to student reading achievement. As for the interaction between quantity and quality of reading, gains in reading achievement were evident only when both ERV and APC were high. An APC of \(80-84\%\) appeared to be an important threshold in this respect. This relationship was particularly evident in the higher grades.

5. Discussion

This study involved post hoc analysis at a national level of archival data initially gathered for local purposes, and as such is in a particular research tradition. It had the strength of the very large size and extensive regional scope of the research practice database from which it was developed. Additionally, integrating correction of possible regression artifacts led to a conservative interpretation of the data. While socio-economic status data at individual student level were not available, analysis of socio-economic status at school level showed minimal variation from the aggregate analysis. Thus, while the sample in this study showed slightly higher ethnic minority membership than the US as a whole, this did not appear to significantly socio-economically bias the sample. The computer-administered and scored format seems likely to be less susceptible to administration variation than traditional paper tests. The measure of ERV used here was a proxy indicator, primarily based upon number of words read but also taking into account the difficulty of the words and successful comprehension.

Albeit from correlational rather than experimental research, the present study shows that both ERV (largely quantity of reading) and APC (pure success in reading comprehension — quality of reading) are important for reading achievement gain. This held true at all initial reading levels (a significant departure from previous related report.
in the literature), was thus important for practitioners, and also suggested causal direction from ERV + APC to achievement rather than vice versa. High ERV and high APC together were particularly important for reading achievement in older readers, who might be more inclined to read independently carelessly and below their optimal readability level. However, there were signs that classroom placement countered this tendency and had a strong relationship with achievement gains. This tends to suggest that a large factor in classroom placement was indeed the behavior of the teacher.

From a broader theoretical perspective, the current study suggested that simple information-processing models of reading practice were inadequate. Volume of practice is only one relevant variable, and not all practice is the same. Pure quality of independent reading practice and classroom placement were as important as quantity of reading practice. Theoretical models need to take account of three variables not one, and distinguish between affordances and the extent to which they are actively utilized. This is likely to have relevance for theoretical models of effective practice for complex skills in areas other than reading.

6. Conclusion

Positive relationships between reading achievement gain and both largely quantity of reading and pure quality of reading were evident at all levels of student reading achievement, suggesting that the relationship between quantity and quality of reading and gains in reading achievement is independent of initial reading ability. Both largely high quantity (ERV) and pure high quality (APC) in combination were necessary for high reading achievement gains, especially for older students. ERV and APC were weakly associated with student initial achievement, but more strongly associated with classroom placement, suggesting that the properties of teacher action in guiding independent reading were important.

Time spent reading without guidance has only a modest influence on reading achievement. Thus, practice does not make perfect — but attuned, successful practice makes perfect. The mere allocation of time to independent reading might have little impact upon reading achievement. Similarly, the mere presence of a computer assessment system in the classroom is unlikely to have any impact on reading achievement. Appropriate, effective implementation involves not only the monitoring of reading practice, but also implies action to guide the student towards successful comprehension.

It is encouraging that far more of the variance was explained by classroom placement than by student initial reading achievement, since the former might be more easily changed than the latter. This has implications for educational practice and policy, particularly for older students, and across all student initial achievement levels. Without a computer assessment system of some sort, exploring and addressing such questions is difficult. Teachers need better data for formative assessment of student progress. However, such data are only useful where feedback consistently informs subsequent adaptive action and leads to better learning.

The harvesting of computerized formative and summative assessment data at individual student level (analyzed at classroom, school, school district or national level) holds considerable promise for the investigation of the effects of at least some student choices/behaviors in a way that leads to implications for practice and policy. This can be done at low cost and is becoming even easier through web technology. There are implications here for the cost-effectiveness of research.

References


