READ MORE – READ BETTER?

A META-ANALYSIS OF THE LITERATURE ON THE RELATIONSHIP BETWEEN EXPOSURE TO READING AND READING ACHIEVEMENT

M. Lewis
S. J. Samuels
University of Minnesota

Abstract

While most educators believe that students who read more are better readers; the National Reading Panel report questioned this assumption. Prior reviews on the relationship between reading exposure and reading achievement have either been too selective, omitting informative studies; or limited in the depth of analysis, using only narrative review. This meta-analysis attempts to address these deficiencies. Both a directional analysis and a meta-analysis of effect sizes were done on those studies included in the statistical analysis. Potential moderating factors were also examined. This review provided support for a moderately strong, positive, relationship between reading exposure and reading outcomes. Separate analysis of d-index effect sizes from experimental studies provided clear causal evidence that students who have in-school independent reading time in addition to regular reading instruction, do significantly better on measures of reading achievement than peers who have not had reading time. Reading time was especially beneficial for students at earlier stages of reading development: students in lower grades, those experiencing difficulties learning to read, and students learning English as a second language. The value of reading was also greater for students in rural and urban schools. While other factors with a potential moderating effect on the relationship between reading exposure and reading achievement need to be examined, the results of this meta-analysis support the belief that students who read more are better readers.

According to Gibson and Levin, "Reading is overtaught and underpracticed" (1975, p. 219). In support of this position, much has been written to suggest that if students had more reading practice time they would be better readers (Allington, R. L. 1977, 1980; Anderson, Wilson & Fielding, 1988; Cunningham & Stanovich, 1998; Guthrie, Wigfield, Metsala & Cox, 1999; Krashen, S. 1997-1998; Watkins & Edwards, 1992). However, the reviewers on the National Reading Panel (NRP, 2000) were "unable to find a positive relationship between programs and instruction that encourage large amounts of independent reading and improvement in reading achievement, including fluency. In other words, even though encouraging students to read more is intuitively appealing, there is still not sufficient evidence obtained from studies of high methodological quality to support the idea that such efforts reliably increase how much students read or that such programs result in improved reading skills" (p.13). Although the NRP report did not find significant statistical support for a relationship between independent reading practice and increases in reading achievement, it did not entirely close the door on the possibility that such a relationship might exist. Panel members did conclude that it was not possible to sustain a negative conclusion about the value of reading practice; they simply felt that current research had not clearly demonstrated that a positive relationship existed.

Justification for this Meta-Analysis

Time for reading has been a foundational part of many school reading programs such as Sustained Silent Reading, Reading is Fundamental, Reading Recovery, Success For All, the Accelerated Reader and the Electronic Bookshelf. These programs share a belief that time spent actively reading is a crucial component of learning to read; you need to practice reading to become a good reader. There are accepted beliefs and knowledge, grounded in theory and research, to suggest that independent reading time could be a beneficial instructional tool. First is the intuitive belief in the value of practice. Second, there are substantial differences in the amount of reading good and poor readers do and in the type of reading instruction they receive. Third, we have models of learning that support a positive relationship between time, learning, and achievement. And last, there is experimental validation of the benefits of reading practice from studies of assisted reading procedures.

No one would deny that practice is important. Practice is the mantra of musicians, athletes, and children learning the alphabet. Parent involvement with reading, summer and Saturday reading clubs, and reading programs for preschool children evolved from an intuitive, common sense notion that more reading is beneficial (Sullivan H. J., & Labeaune, C., 1970; Curry, J., & Zwskowsky G., 1999). A meta-analysis on joint book reading (Bus, van IJzendoorn & Pellegrini, 1995) supported the hypothesis that parent-child reading is related to outcome measures such as language growth, literacy development, and reading achievement.

Time is a precious commodity in schools; many topics and subjects have to be taught to a variety of students with widely differing abilities. Much time, attention, and instruction has been devoted to teaching students to read, but many students still do not learn to read beyond minimal literacy requirements. Reading is considered a vital and necessary skill, but the reality is that teachers still spend more time on skills instruction and isolated practice with phonics and word identification than on reading connected text. In a study of reading instruction time in first, third and fifth grades Kurth & Kurth (1987) found an average of 13.9% of reading instruction time given to actual reading. More time was spent on transitional, non-instructional activities (15%). A more recent survey revealed that third, fourth and fifth grade teachers in three southeastern suburban schools, were allocating only 35% of a 51-minute reading period to actual reading (McNinch, G. H., Shaffer, G. L., Campbell, P., & Rakes, S., 1998).

Students receive differing amounts and quality of instruction that contributes to disparities in reading ability (Allington, 1977, 1980, 1984; Applebee, Langer & Mullis, 1988; McGill-Franzen & Allington, 1991). A study by Allington (1983) indicated that good readers were actively reading and engaged in meaningful discussion 57% of allocated instruction time, poor readers only 27%. Good readers read silently, where the focus is on understanding, 70% of the time, while poor readers read orally, where the focus is on accuracy, 70% of the time. The trend has thus been for better readers to complete more reading in every session and do more silent reading than poor readers. With similar time allocated to reading instruction, poor readers, who are covering less reading material and spending less time reading those materials, will have increasing difficulty narrowing the gap with good readers. This

disparity has been observed not only for poor readers but also for other at-risk groups, such as deaf students. (Limbrick, McNaughton & Clay, 1992).

Education is premised on the belief that time spent in study and practice leads to learning. Models of learning stress time-on-task as an important factor in what and how much is learned, and support a linear relationship between study, practice time and learning, at least up to some asymptotic amount of time. In his models of school learning and time, Carroll says, " a learner will succeed in learning a given task to the extent that he spends the amount of time that he needs to learn the task" (1963, p. 7). According to Bloom (1974), with high quality instruction and sufficient time to learn, virtually all students can learn. Bloom concluded that time for learning, in conjunction with feedback, and at least moderate success, may well be the critical formula for "real" learning to occur. Harnischfeger & Wiley's model of instructional time (1985) also lists active learning time, or time-on-task, as its most important concept. There is also evidence that time may be even more critical factor for students at lower levels of aptitude and achievement than for those at higher levels (Brown & Saks, 1986; Paul, 1992, 1996).

Reading practice has been successfully used and experimentally validated in repeated reading and other assisted reading practices (Algozzine & Lockavitch, 1998; Dowhower, 1987; Rasinski, 1990; Samuels, 1979; Shany & Biemiller, 1995; Sindelar, Monda, & O'Shea, 1990). Research has demonstrated that students who engage in assisted reading practices improve on measures of reading achievement significantly more than students who do not get this practice time. Rossman (1986) found a strong positive correlation for first, third and fifth graders between reading automaticity and reading time in school. Students who had reached automaticity read with greater accuracy and comprehension than non-automatic readers. This implies an indirect, potentially causal relationship between time spent reading and reading achievement, at least up to the point where reading automaticity is attained.

Although prior reviews on the relationship between print exposure and reading outcomes have been done, they have largely ignored measures of reading practice outside the umbrella of sustained silent reading programs, and have been limited to narrative review and summary. The NRP had very narrow

criteria for inclusion of studies, examining only those experimental and quasi-experimental studies published in peer-reviewed journals after 1990. The report justified this restriction by citing a lack of rigorous, gold standard studies, stating that most evidence was correlational (2000). While correlation is not causal evidence, it does illuminate important relationships and provides meaningful evidence when combined with other sources of information. It is therefore, extremely important that the literature on reading practice be examined, systematically, critically, thoroughly and objectively to provide a clear and supported summary of this body of research that examines the role of time and practice and its relationship to reading outcomes.

To our knowledge, there has not been a meta-analysis of the literature on reading exposure that has looked at all available studies from all available sources. We hope this meta-analysis will provide a thorough, organized, clearly defined, broadly investigated analysis of the literature on reading exposure. Schwandt (1998), states that a review sums up or synthesizes knowledge so that we can enhance our capacity to solve educational problems, to consolidate what is known into a succinct and useful form. The hope is that this analysis will clarify our pool of knowledge about reading practice, help define the parameters of research on reading practice, and lay a foundation for more clearly focused future research. The questions addressed by this meta-analysis are:

1. Does the research literature show that students who read more do better on measures of reading achievement including comprehension, vocabulary and word knowledge, and measures of reading fluency such as reading rate and accuracy?

2. If time spent reading has a positive effect on measures of reading performance, does the research literature suggest factors that contribute to positive reading outcomes?

Prior Reviews of Reading Time

We found ten reviews on reading practice programs; most were reviews of sustained silent reading programs only. Of these 10, six found independent reading time to be a positive practice. Moore, Jones and Miller (1980) concluded that, overall, sustained reading practice has a positive effect on

reading comprehension, when combined with a program of reading instruction. Schaudt's (1983) review of independent reading time also led her to conclude that sustained silent reading is an effective addition to reading instruction. Tunnell and Jacobs (1989) concluded that literature based programs are successful in helping a wide range of students learn to read and to enjoy reading. Krashen (1994) concluded that what he calls free voluntary reading is not only the "major course of our reading competence" (p.113), but is the only way to literacy. Meyers and Lass (1998) also concluded that reading practice improves reading achievement. Chow and Chou's (2000) review of reading practice reported a positive effect on reading ability.

Three reviews concluded that allowing students time for independent reading, in addition to regular reading instruction, is as beneficial as reading instruction alone. Sadoski (1980) stated that reading is a skill that develops with practice but he found sustained silent reading to be neither more nor less effective than other approaches to reading. Both the Weisendanger and Birlem (1984) and the Manning-Dowd (1985) reviews found the effects of sustained silent reading inconclusive, but expressed the belief that practice time was probably effective when combined with systematic skills instruction. The NRP (2000) found no evidence to support or refute the benefits of reading time in improving reading skills and achievement.

Overall, these reviews concluded that time spent reading, largely in the form of sustained silent reading, to be a positive and effective practice, though not always significantly so. Most believed the practice to be at least as effective as regular instruction alone, and best used in combination with more traditional reading instruction. Reviewers emphasized that reading practice is just that, practice of reading skills that first need to be taught and learned.

Method

Inclusion Criteria

Studies included in this meta-analysis had to meet several criteria:

— First, the focus had to be on reading practice effects, or some measure of exposure to reading and its relationship to reading achievement.

— Second, reading practice had to be either a stand alone program or be measured and analyzed independently of an overall reading program.

— Third, studies had to include a quantitative measure of both reading exposure and achievement.

— Fourth, studies had to identify the student sample, the number of students and their age or grade.

— Fifth, the report had to specify how much daily or weekly time was spent reading, for how many weeks or months.

— Sixth, the study had to specify how reading time or exposure and achievement were measured.

— Seventh, the study had to report correlations or other data from which effect sizes could be computed.

— Eighth, outcome measures had to focus on students.

— Ninth, reports had to be in English and related to the effects of reading on English reading ability.

No study that examined the relationship between reading experience and reading achievement was totally excluded from this review. Studies that did not report enough data to compute an effect size, but did indicate the direction of results, were included in a directional analysis.

Meta-analytic Procedures

In a meta-analysis, three assumptions are crucial to validity (Cooper, 1998). First, the individual research findings in a meta-analysis should all test the same construct or relationship. Exposure to reading was been measured in a variety of ways: time spent reading, amount of reading, and indirect measures of reading experience. These different measures can be consolidated as measures of reading exposure. The second assumption is the independence of the separate dependent variables. Each contributing piece of information should be unique. Reading achievement has been assessed with a variety of measures: vocabulary, comprehension, accuracy in context, and reading rate. They are not completely independent, but results did reveal differing effects of time on these measures. Analysis was therefore done on overall reading achievement and on the individual reading assessment measures. The third assumption concerns the validity of the primary studies included. One must assume that the primary

researchers made valid assumptions of independence and normal distribution when they reported test results, to justify doing a meta-analysis on those studies.

Effect size computation

The d-index effect size, was calculated for each study that reported means and standard deviations, t-tests, or F tests with one degree of freedom in the numerator. Multiple effect sizes on the same construct are a potential problem in meta-analysis. They are statistically dependent and violate the assumption of independence of data points. A single effect size can be selected, or an average of the effect sizes can be calculated to produce one single effect size for each construct measured (Cooper, 1998). There were no consistent or especially suitable criteria to justify choosing one measure over another, therefore, multiple effect sizes on the same construct were averaged to contribute one effect size for each sample for each reading achievement construct measured.

Another complicating factor in estimating effect size occurs when studies compare more than two groups. A sample may compare two experimental conditions to a single control. The three d-indexes from these comparisons are not statistically independent. According to Cooper (1998), using multiple effect sizes is preferable to using multiple inference tests. When several comparisons were made between sample groups, we included all comparisons of interest.

The r-index effect size is the Pearson product-moment correlation coefficient that describes the relationship between two continuous variables in terms of degree and direction. In combining correlations in a meta-analysis, sampling distributions become non-normal as values approach the limits of +1.00 and –1.00. For this reason, several statisticians (Cooper, 1998; Lipsey & Wilson, 2000; Rosenthal, 1995) recommend converting r-indexes to their corresponding z-scores, which have no limiting value and a normal distribution. Once the average z-score has been calculated, it can be converted back to r for interpretation.

For overall analysis of all studies, the r-index was chosen as the effect size metric. Although different effect size metrics can be easily converted, it is important to choose one to describe results. The

confidence interval around the effect size statistic was calculated to provide a test of the null hypothesis that there is no relationship in the population between reading exposure and reading achievement. Homogeneity analysis was used to compare observed effect size variance to expected variance, testing whether they come from the same population. Homogeneity is calculated using the Q statistic. Q has a chi-square distribution with k-1 degrees of freedom, where k = the number of effect sizes. If the observed variance is significantly different from expected variance, we reject the null hypothesis that the variation among effect sizes is due to sampling error alone. We then need to examine whether other characteristics might be associated with variations in effect size. These potential moderator variables were also examined using homogeneity analysis. Studies were grouped according to various characteristics and the average effect size tested. A significant Q value indicates that particular characteristics are possible explanations for observed variations in effect size.

A meta-analysis involves assumptions about the population represented by the included studies; these may be fixed or random. A fixed effects model assumes that all the comparisons come from the same population, so observed effect size differences are a result of sampling variability. This model is appropriate if an exhaustive search for and examination of moderator variables can be and is done. It is also considered appropriate in situations where the inferences made are conditional and concern only the observed studies (Hedges & Vevea, 1998). A random effects model assumes other sources of variability in addition to sampling error. If these factors are believed to be a potentially significant source of error, then a random effects model is appropriate. It is also appropriate when one wishes to make inferences that are generalizable beyond the observed studies, implying that the set of observed studies may reveal something about a larger presumed set of studies that are the real object of interest (Hedges & Vevea, 1998). Random effects models typically yield a more conservative analysis than fixed effects models. In examining reading performance, individual differences may be related to a variety of influences beyond reading exposure. Some of these factors were not reported and so could not be analyzed. It is also true that inferences would be made to the larger population of all students. Therefore, we chose to apply both fixed and random effects models, running all analyses twice.

Sample of studies

Two search strategies were used to locate studies. Primary searches were done through ERIC, PsychINFO, Education Abstracts, Dissertations Online and Dissertation Abstracts. The reference section of every article retrieved was also examined for additional articles missed through direct search. Search terms were initially entered using the keyword index. This turned up thousands of articles. Search terms were therefore indexed in the abstract first, and if no articles were found then searches were done through the subject and major descriptor indexes.

Table 1 reports the overall search results. After eliminating duplications, 1306 articles were found. The abstract of each of these was examined. Of the 251 studies retrieved, 97 were coded, 48 studies in the directional analysis and 49 in the meta-analysis. A second search was made shortly before the final analysis was done. This search turned up 33 additional reports; none were suitable for inclusion.

(insert table 1 here)

Coding Format and Database

All reported information related to reading was coded. Information was classified by design type; sample selection criteria and group equivalence; student and teacher characteristics; school environment; reading instruction methods and materials; pretests administered and results; reading practice measures; and outcome measures. A brief description of the researcher's findings was provided, statistical analyses done and results, and researcher conclusions. It was noted whether or not the study satisfied the hypothesis and if not, why not. Effect size magnitude and direction was computed for each outcome measure. The r-index, was coded as reported. Space was provided for any additional information readers felt was important, and to note limitations, design flaws, or other factors that might have affected results.

Both a directional summary and a meta-analysis were planned because of the disparity in type of data reported across studies. A sorting sheet was developed to identify studies for inclusion in the meta-

analysis database. This sheet included a reviewer developed rating for studies based on design type, sample size, group equivalence and analysis.

Inter-rater Reliability

Two graduate students were employed to read and code articles. Coding issues and problems were discussed in weekly meetings. The primary reviewer recoded any articles readers had difficulty with not resolved in the meetings. The primary reviewer also briefly examined each of the articles coded by the readers, and independently coded 20 percent of the studies they coded. There were no significant disagreement; inter-rater reliability was 95%. All data used in the calculation of effect sizes was rechecked before entering into the database.

Results

Directional Analysis Results

Forty-eight studies that were not useable in the meta-analysis were used in the directional analysis. Of these, 43 looked at time spent reading and were analyzed using vote count methods. The other five studies were individually reviewed. Table 2 shows the relationship between time spent reading and reading outcomes. For example, a study might show a positive relationship between time and outcomes, or no effect at all. In 79% of the student samples that had some form of independent reading experience, reading ability scores improved in relation to their pretest scores or in comparison to a non-reading control group. The reading ability scores of 8% of the student samples who participated in some form of independent reading experience decreased in relation to their pretest scores or in comparison to a control group that did not have any independent reading experience. All of these samples involved one group pre to posttest comparisons of students who were using the Accelerated Reader computer reading management program. Research conducted by Renaissance Learning reports that negative scores do not mean that students' reading ability has decreased, only that they have kept up with, but not surpassed national norm standards. Thus, in no instance did independent reading practice experience have a detrimental effect on reading growth and achievement. The scores of 13% of the student samples with reading exposure stayed the same in relation to pretest scores or in comparison to a control group. It is

important to note that in these studies, means were not reported so we don't know if there was any difference in scores. It is possible that students engaged in some form of reading exposure made substantial pre to posttest reading achievement gains or had higher scores in comparison to a control group, even though those differences were not statistically significant. If students are benefiting from independent reading time, then we need to evaluate the practical as well as statistical benefits.

Insert table 2 here

Results were analyzed with vote count methods that use the proportion of positive and negative results to calculate the underlying magnitude of a treatment effect. If there is no relationship in the population between reading exposure and reading achievement, we would expect the positive and negative results to cancel each other out. A sign test is performed to see if the cumulative findings suggest that one direction occurs more often than would be predicted by chance. The computational formula for the sign test is:

$$Z_{vc} = \frac{(N_p) - (\frac{1}{2} N)}{\frac{1}{2} \sqrt{N}}$$

where Z_{vc} equals the standard normal deviate, the z score for the overall series of findings; N_p equals the number of positive findings, and N equals the total number of findings. These 43 studies had a standard normal deviate score of $Z_{ve,} = 5.965$, $p < .0001$. They provide substantial evidence for a strong positive relationship between reading exposure and reading achievement. They also suggest that reading growth occurs as a result of participation in some form of independent reading, or in programs that promote and support independent reading practice time. We can conclude from this evidence, that, in general, as reading time increases so too does reading achievement.

The five miscellaneous studies included measures of specific comprehension skills, measures of literacy using different types of written documents, and studies that examined learning vocabulary from text exposure. Results showed that specific comprehension skills such as finding sequences, locating

main ideas and supporting details, and drawing conclusions, as well as vocabulary, improved with reading experience.

These positive results add to the evidence about the value of reading time, and can help inform the most effective use of reading practice. They suggest that programs that combine reading time with some form of skills instruction help students become better readers. The reading programs examined here included readers workshops, sustained silent reading periods that supplemented skills instruction and frequently included some responsive activities, and computerized reading management programs that assess student understanding of text. It is possible that reading programs that encourage depth of processing through activities such as taking quizzes, writing reports, or discussing books that the students have read contribute to positive reading outcomes.

Meta-analysis Results

Sample of Studies

Forty-nine additional studies described the effects of reading exposure on measures of reading achievement with quantitative data that could be used to compute effect sizes. There were 17 correlational studies. There were 24 quasi-experimental designs that randomly assigned classes to conditions, 13 of the quasi-experimental studies had control groups, and nine compared different treatments. There were eight true random experiments, six with control groups and two that compared treatments. These 49 studies involved 106 samples and 192 reading achievement comparisons. Sample size ranged from 5 to over 17,424. More than 112,000 students in all were involved in this examination of reading exposure.

Insert table 3 here

Table 3 lists the overall direction and significance level of results. Twenty-seven studies compared the reading achievement of students who were given time to read during school, to that of students who had reading instruction only. Of these, 17 had positive, significant results. Two studies that

compared varying amounts of reading time, reported additional time as valuable, at least up to some asymptotic amount of time, about 15 minutes a day. Six studies reported positive, non-significant results for readers of varying ability in intact classes. One study compared students in three different reading instruction programs that all involved reading and understanding text, and demonstrating that understanding. All three groups improved, but differences were not significant. Three studies reported negative, non-significant results. No study reported significant negative results; in no instance did allowing students time for independent reading result in a decrease in reading achievement. Some studies massed reading practice into one or two sessions a week, others provided daily reading time. Results did not fully support either practice. This is a variable that needs further study. Do students benefit more from massed practice or smaller amounts of distributed practice?

The five studies that assessed reading time with print exposure measures all revealed a strong positive relationship between reading experience and reading growth. Regression analysis revealed that reading experience was significantly predictive of reading ability even after other factors such as age, and general, phonological and orthographic ability were accounted for. Here also, there were no instances where reading time had a negative relationship with reading achievement.

Eleven studies examined the benefits of recreational reading time outside of school programs. Seven of these studies used activity logs to measure reading exposure; time spent reading ranged from none to 49 minutes a day. In all instances, students who read more scored higher on measures of reading achievement; only one study reported a positive relationship that was not significant. Most correlations were in the small to moderate range. The one experimental study that had gifted students, found that fourth graders who read for pleasure had significantly higher reading comprehension scores than those who did not read for pleasure. Four studies used data from surveys to assess reading exposure. In three studies the relationship between reading time and achievement was positive and significant; in the fourth study correlations were small but positive. There were no negative correlations, no instance where students who read more were at a disadvantage.

The remaining six studies looked at the relationship between reading time and reading achievement in several ways. One study found that students in schools with higher than average gains in reading achievement had more in-school reading time than students in schools with lower than average reading achievement test gains. Two studies used observations of time spent reading in a learning disabled classroom. The first found a significant relationship between time spent reading and reading achievement. The second analysis of the same data concluded that, although reading time is an important factor in reading achievement, it is not as significant as pre-existing reading ability. Three studies examined the impact of the Accelerated Reader, a computerized reading management and assessment program that monitors student reading with comprehension quizzes. All three revealed a positive relationship between how much students read and scores on quizzes. Correlations were in the moderate to high range, especially for developmental and remedial readers.

The studies in this meta-analysis measured exposure to reading in a variety of ways: time spent reading, amount read, frequency of reading, and print exposure measures. In virtually all studies, across a wide variety of reading measures, the conclusion was the same. Independent reading time, both in and out of school, is an effective means of helping students improve reading ability and achievement. There were no studies that reported overall significant negative results; in no instance did exposure to reading lead to a decrease in reading achievement. In the three studies that did report negative results, the authors concluded that giving students independent reading time in addition to regular reading instruction was not less effective than skills instruction only. It is hard to look at the accumulated data from these 49 studies and not conclude that students who read more are better readers.

It was interesting to note that many studies with significant results involved other factors interacting with reading. These activities included adult support; responsive activities that encourage depth of processing, such as written and oral reports that foster understanding; and ongoing assessment and feedback from programs such as the Accelerated Reader. When students are successful, it leads to additional growth and improvement. Children are motivated to do that which they can do well.

Table Four

Table 4 lists some of the characteristics of the studies included in the meta-analysis: first author, year of the report, number of student participants and their grade level, type of comparison made, reading level of students, what outcome measures were used to assess achievement, the effect size statistic, and whether or not the study overall supported reading practice. The variables examined as potential factors in effect size variation included: reading level and grade of student participants; community location of student population; how reading exposure was measured; how reading achievement was measured; and methodological rating of study design.

Results

All reading achievement measures

Table 5 lists the Q values as well as the mean z, and 95% confidence interval value across all reading achievement measures. If the effect size value is larger than the confidence interval value, then z is significant at $p < .05$. The average z, under a fixed effects model, for all reading achievement measures, on 192 comparisons, was 0.100. The 95% confidence interval had a lower value of 0.094 and an upper value of 0.105. Under a random effects model with a z value of 0.194, the confidence interval expanded to a lower value of 0.168 and an upper value of 0.261. These intervals do not contain zero; the mean z is significant at .05, and we can reject the null hypothesis that reading exposure has no effect on measures of reading achievement and performance. Thus the meta-analysis results support the hypothesis that exposure to reading is an important factor associated with reading achievement.

---

insert table 5 here

The fixed effects test for homogeneity of variance was significant, Q (191, k=192) = 2843.43, $p < .001$. This means we can reject the hypothesis that sampling error alone accounts for the variation among effect sizes and need to examine moderator variables that may be influencing the variation in the distribution of effect sizes. Table 6 lists k, Q value, mean z effect size, and 95% confidence interval value for the effects on reading achievement of reading level, grade and grade.

Insert table 6 here

The moderator variables examined included reading level and grade. Reading level was collapsed into 5 categories: below grade level, at grade level, above grade level, heterogeneous groups, and ESL and LD students. If intact classrooms were used and reading level was not mentioned, it was assumed that the students represented a range of reading levels. Student reading level was a strong predictor of effect size variation using both fixed and random assumptions ($p < .001$ and .025). Results suggest that below level readers, ESL and LD students, and average readers, benefited more from reading exposure than above average readers or mixed ability groups.

Grade level was grouped into three categories: grades one through three, four through six; and seven through nine. Ten studies were not included because they examined students in multiple grades. Grade level was a strong predictor of effect size variation under both fixed and random assumptions, ($p< .001$). Students in grades one through three benefited the most from reading exposure.

Individual Measures of Reading Achievement

Separate analysis was run on each of the individual measures of reading achievement: vocabulary, comprehension, total reading score, accuracy in context, word recognition and reading rate.

Table 7 lists k, Q value, mean z and 95% confidence interval value for the individual measures of reading achievement. The average effect size (z), for vocabulary as a function of reading exposure was significant under both fixed and random effects model, ($p < .05$). The average effect size (z), under a fixed effects model, for comprehension, was also significant at .05. It was not significant using random model assumptions. The average effect size (z), for total reading score, accuracy in context and reading rate was significant at .05 using both fixed and random assumptions. We can reject the null hypothesis that reading exposure has no effect on these individual reading outcomes. Homogeneity of variance analysis was significant on all individual measures of reading achievement. This indicates that factors, other than sampling error were moderating effect size variation.[2]

---

insert table 7 here

Discussion

This study examined the relationship between the amount of reading exposure and reading achievement, and possible evidence linking the amount of exposure to reading outcomes in a causal relationship. According to the results of this meta-analysis, the overall correlation between reading exposure and reading achievement was .10 under fixed assumptions and 0.192 under random assumptions. Cohen (1988) labeled effect sizes of $r = .10$ as small, $r = .30$ as medium and $r = .50$ as large. These correlations are small to moderately small, but significant because confidence intervals do not contain zero. We can conclude that there is a positive and significant relationship between exposure to reading and reading outcomes. Significant Q value indicates that sampling error alone does not account for the variation in effect sizes. Moderator variable analysis revealed that students who are at earlier levels of reading ability, and those having difficulty learning to read such as remedial and second language students, benefit more from reading exposure. These findings suggest that students who read more have higher reading achievement.

As discussed earlier, r values are correlational and do not address cause and effect relationships. There were eight true experiments that used random assignment of students and do address cause and effect relationships. The overall d-index effect size for these eight studies was .422. Cohen (1988) labels d-indexes less than .20 as small, .50 as medium and greater than .80 as large. These d-index values fall unto the moderately high range, indicating that students who have some form of reading experience have reading achievement scores that are 2/5 to 3/5 of a standard deviation better than comparable students who do not have this exposure to reading. Researchers consider effect sizes greater than .33 to have practical significance; "the effect size is large enough to make a worthwhile difference in the outcome" (Borg and Gall, 1989, p. 7*).* We can therefore conclude that spending time reading has at least a moderate causal effect on growth in reading outcomes. The instructional implications are that curricula that allocate reading time up to some asymptote should foster increased achievement in reading.

Another way to interpret d-index effect sizes is with the U_3 measure (Cohen, 1988). This measure tells us what percent of scores in the lower-mean group (students who have less or no independent reading time) are below the average score of the higher-mean group (those who have more reading time). The average reading achievement score of students who had some form of independent reading was higher than 65.5% of the reading achievement scores of students who did not have independent reading time. One can conclude from this that the amount of time spent in reading practice had a positive effect on reading achievement.

According to the research reviewed here, students who have more time and opportunity to read perform better on measures of reading achievement. From the studies reviewed here, the optimum independent reading time appears to be 10 to 30 minutes a day in addition to some form of skills instruction. It also appears from this literature review that those students who are in the early stages of reading benefit the most from the opportunity to spend time reading. Not enough information was available to address the potential moderating effects of instructional methods, teacher characteristics, instructional materials, support and assessment activities, performance feedback, or gender and ethnic

differences. These factors need to be examined to determine those conditions that would optimize reading practice time.

Added to the strong correlational support of the relationship between reading exposure and reading achievement and the causal support from the randomized experiments is the evidence from the directional analysis. Vote count analysis indicated a 1 in 10,000 chance that the positive findings of the directional analysis occurred by chance, given that in the population there is equal reading growth both with and without independent reading exposure. We can conclude from an examination of all of the components of this review on the relationship between reading exposure and reading achievement that, in general, students who read more do better on a variety of measures of reading achievementl

All this evidence reveals not only a strong positive correlational relationship between reading time and reading achievement, but some probability of a causal relationship as well. This causal connection may well go in both directions at different points along the way to becoming an effective, independent reader. This idea has support in Stanovich's research on "Matthew Effects" in reading. Better readers continue to read more and do better . Those who experience early success in reading are likely to read more and so progress at a faster rate than those who have a less positive relationship with learning to read (Stanovich, 1976). But success at any age can set this process in motion. Furthermore, students who experience a slow start in reading often receive remedial instruction that emphasizes the mechanical aspects of word recognition, while those students who have early success in reading are introduced sooner to reading meaningful stories. This in itself tends to motivate the better students to read more (Allington, 1977, 1980, 1983, 1884). It may be more important to look at how to facilitate this relationship rather than try to figure out in which direction it is moving.

According to the literature reviewed here, students at earlier levels of reading development appeared to benefit the most from more time spent reading. This includes students in elementary grades, those having difficulty learning to read and those learning English as a second language. Box (1984) looked at the effects of in-school independent reading time on the achievement of third graders. Students who had 10 minutes daily independent reading time in addition to reading instruction had significantly

higher scores on the vocabulary subtest of the California Achievement Test than similar students who had reading instruction only. Comprehension scores were higher as well, but differences were not significant.

Box also examined the effects of reading time on students at varying levels of ability. Low ability readers benefited the most, performing significantly better on measures of vocabulary and comprehension than control group students of similar reading ability. Denslow (1985) also examined the relationship between reading time and achievement for below grade level readers. Students in grades one through six who read 15 to 20 minutes, in addition to reading instruction, had significantly higher scores on the ITBS vocabulary and comprehension subtest compared to similar students who had reading instruction only.

Several studies of students learning English as a second language showed that students in grades one, three, four and five who used book immersion programs had significantly higher scores on measures of word recognition and comprehension than students who used oral and taped language learning methods (Elley & Mangubhai, 1983, Elley, 1991). Hafiz and Tudor (1989) showed that after school reading programs can also benefit reading ability and achievement. Students, 10 and 11 years old, who participated in a voluntary after school reading program had significantly higher vocabulary, sentence and passage comprehension scores than students who did not participate in the program. The students in the experimental group also had lower pretest scores than control group students.

Learning disabled students also benefit from the opportunity to practice reading. Melton (1993) compared the achievement of LD students who had 10 minutes daily in-school reading time to that of similar students who had reading instruction only. Students who read for 10 minutes a day had significantly higher achievement on word recognition in context and reading comprehension.

Time was the primary variable of interest. The amount of time that students spent reading varied considerable across these studies, from two to three minutes daily to 40 to 50 minutes several days a week. Significant results were seen with as little as 10 minutes of daily reading time, especially with developmental readers. There may well be an upper limit of time as well. There were two studies that looked at the effects of variations in independent reading time on reading achievement. Both studies

indicated that simply increasing the amount of independent reading time did not necessarily lead to greater achievement. Lawson (1964) compared 45 minutes reading instruction, 45 minutes of independent reading with no instruction, 30 minutes instruction plus 15 minutes reading practice, and 15 minutes instruction plus 30 minutes reading practice. Students in sixth through eighth grade who had 15 minutes reading instruction and 30 minutes reading practice time had significantly higher scores on the MAT word knowledge subtest than the other three groups. Students who had 30 minutes instruction and 15 minutes reading practice had significantly higher scores than students who had instruction or reading practice only. On the reading subtest, students who had 30 minutes instruction and 15 minutes reading practice did as well as students who had reading instruction only, and both groups scored significantly higher than the other two groups.

McGroarty (1982) looked at the effects of 30, 50, 75, 100 or 125 minutes of weekly reading time in addition to regular reading instruction. Results on the Gates MacGinitie vocabulary subtests were significant, but results were not perfectly linear. While more time was generally better, students who had 50 minutes weekly reading time had significantly higher vocabulary scores. Scores on comprehension were also inconsistent. Here again students who read for 50 minutes a week had the highest scores, but the differences among groups were not significant. Daily reading times were not reported, so it is not known if reading time was distributed throughout the week or massed into one or two periods a week. This factor could have a significant effect on results. Information processing theories state that students who distribute practice, learn and remember more than students who mass practice for a single extended period (Mumford, Costanza, Baugham, Threlfall, & Fleishman, 1994). Students who read on a daily basis, even for a few minutes, may experience more growth in reading ability than students who read for an extended period once or twice a week, even though total reading time is the same. Providing daily in-school independent reading time takes advantage of this concept as opposed to relying on students reading on their own. Research on variations in independent reading time is needed to address this issue. Time-on-task research also stresses the need for an adequate amount of time for optimum learning. The optimum amount of reading time that may lead to increased achievement is probably influenced by

factors such as student ability, task difficulty, motivation, quality of materials and instruction, and effective use of available time.

There are other variables that likely impact the relationship between reading exposure and reading achievement that were not analyzed because of insufficient information reported in the primary studies. These variables included teacher education and experience; types of reading instruction, and assessment; issues of motivation and feedback; type and difficulty of reading materials; reading response and other related reading activities; how reading achievement was measured, with formal or informal measures; the home environments and other characteristics of good readers; and differences due to gender, ethnicity or socio-economic status. These are important variables and should be examined in future research of the relationship between time spent reading and reading achievement.

Recommendations for Future Research: Methodology

The methodology in current research has varied considerably. Reading research needs to find a common ground, protocols and procedures that are methodologically sound and powerful, and useful to educators. Educational research involves unique situations and contexts that need to be considered when conducting reading research. Students come to us in already formed classrooms. Research that uses true random assignment may be methodologically best but not always practical; it is not always possible to randomly assign students within a classroom to different treatment conditions. Even when possible, it may result in very small study groups that are likely to change through normal student attrition. This reduces statistical power and generalizability. Using intact classes is more realistic, creating study groups that mimic actual classroom conditions. It is essential, though, to closely match student groups on as many characteristics as possible to improve statistical power and generalizability.

Reading does not occur in a vacuum and cannot be examined as such. All aspects of teaching and learning affect reading activity and should be considered when doing reading research. Research needs to move beyond looking at isolated aspects and integrate all the variables and contexts that contribute to the

relationship between reading exposure and reading achievement. Finding out what works is not simply "do this or that", it is how "this" interacts with all the "that's".

Implications for Reading Instruction

The evidence from this review strongly supports the belief that students who read more are better readers; there is also evidence that reading contributes to higher achievement and does not merely correlate with it. What recommendations for reading practice can we find in the existing research base? One of the recommendations is that students need strong instruction in basic reading skills along with the opportunity to practice those skills reading books and stories. Secondly, this reading needs to be monitored to assess strengths to build on and weaknesses to remediate. Third, students need to read books at the most appropriate level to promote reading growth. Finally, students also benefit from interactions with others by discussing and sharing books and ideas.

The purpose of this review was to conduct a comprehensive examination of the literature on the relationship between reading exposure and reading outcomes, and we trust that this study will shed light on this issue.