iejee | INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

March 2019, Volume **11**, Issue **4**, **309-318**

The Inefficiency of Vocabulary Instruction

Jeffrey Lawrence McQuillan*

Received: **13 December 2018**
Revised: **12 February 2019**
Accepted: **27 February 2019**
ISSN: 1307-9298
Copyright © IEJEE
www.iejee.com

DOI: **10.26822/iejee.2019450789**

Abstract

Several researchers have advocated explicit instruction of vocabulary in order to help students improve their reading comprehension, especially low-achieving readers who need to "catch-up" to their age peers. Very few studies, however, have attempted to compare the time efficiency of direct instruction to its alternatives. In this review, I calculate the efficiency of vocabulary instruction in 14 studies taken from a recent research review (Wright & Cervetti, 2017). I then compare those results with estimates of vocabulary acquisition via a likely alternative source of vocabulary growth, free reading. Free reading was found to be 1.7 times more efficient than direct instruction in building vocabulary in short-term treatments, and 12 times as efficient for long-term treatments.

Keywords: Vocabulary Acquisition, Vocabulary Instruction, Pleasure Reading, Efficiency

Introduction

Success in school rests in significant measure on the ability to understand what one reads. Reading comprehension is in turn strongly influenced by one's vocabulary knowledge (Anderson & Freebody, 1981). Some researchers have concluded that the best way to help students improve both reading comprehension and academic achievement is through some form of direct, systematic vocabulary instruction (Beck, Perfetti, & McKeown, 1982; Biemiller & Boote, 2006; National Reading Panel, 2000; Stahl & Nagy, 2007; Stahl & Fairbanks, 1986). More recently, those emphasizing the importance of acquiring "academic" vocabulary have recommended teaching these words directly to students (Nagy & Townsend, 2012; Snow, Lawrence, & White, 2009).

While vocabulary instruction typically leads to some gains in word knowledge, not all instruction improves reading comprehension. In particular, vocabulary instruction that is limited to giving students the definitions of words – "shallow" instruction – often has little impact on comprehension of texts that contain those words (e.g. Pany & Jenkins, 1978; Pany, Jenkins, & Schreck, 1982). In place of shallow instruction, some researchers have proposed a more time-intensive "rich" instruction that, they claim, will lead not only to greater word knowledge but also increased comprehension. Beck, McKeown, and Omanson (1987), for example, identified several elements of what they considered effective rich vocabulary instruction, including:

> • Providing clear definitions;
>
> • Having students "manipulate" words in "rich and varied ways," describing how words relate to each other;
>
> • Requiring students to discuss words and give "justifications for the relationships" among words they discover;
>
> • Encountering the words frequently and in different contexts;
>
> • Encouraging the use of words outside of the vocabulary lessons (p. 149)

Some words are considered better candidates for this more extensive form of vocabulary instruction than others. Beck et al. (1987) categorized words into three "tiers" in order to determine appropriate targets for instruction. Basic vocabulary ("Tier 1") consists of words that most students will acquire by the early grades (e.g. cat, mother, talk, chocolate), and are therefore not good candidates for instruction. "Tier 3" words are those that are either used rarely or limited to a specific domain, the latter often referred to as technical vocabulary (e.g. photosynthesis, tidal pool, cosine). These again would not be good targets for instruction, since they can be learned "when the specific need arises, such as presenting nebula during a lesson or discussion of the solar system" (Beck et al., 1987, p. 155). Beck et al. recommend that teachers instead focus their vocabulary instruction on "Tier 2" words, those that are "of general utility not limited to a specific domain" (p 155). These are also sometimes called sub-technical words (Cowan, 1974), and can be found in a wide variety of genres and subject matter texts (e.g. influence, ponder, retort, thread).

Most vocabulary interventions have identified these words based on teacher or researcher judgment. A few more recent interventions have used words from the Academic Word List (AWL) (Coxhead, 2000). The list consists of 570 word families thought to be especially important in academic reading. Ming-Tzu and Nation (2004) found that the AWL word meanings were roughly similar across disciplines, meaning that acquiring an AWL word in one domain will be beneficial in other academic disciplines as well.

Studies of the effects of teaching words on reading comprehension have produced mixed results. Stahl and Fairbanks (1986) found a modest effect of instruction on standardized reading tests (d= .30) but a much stronger one for researcher-created passages that contained the words taught in the intervention (d= .97). Elleman, Lindo, Morphy, and Compton (2009), reanalyzing several of the same studies included in Stahl and Fairbank's review, found the effects of vocabulary instruction on comprehension were far lower, with no significant impact on standardized measures (d= .10) and modest but significant effects on researcher-created tests (d= .50).

*Correspondence Details: Jeffrey Lawrence McQuillan, Center for Educational Development in Los Angeles, California. P.O. Box 66577, Los Angeles, CA, U.S.A 90066. E-mail: jeff@eslpod.com

© 2018 Published by T& K Academic. This is an open access article under the CC BY- NC- ND license. (https://creativecommons.org/licenses/by/4.0/)

Wright and Cervetti's (2017) narrative review of vocabulary instruction studies came to a conclusion similar to Elleman et al.'s (2009). They found that words taught in certain interventions were effective in helping students improve their reading comprehension of a text containing those words, but this effect did not generalize to other texts, such as those found on standardized tests.

Incidental Vocabulary Acquisition via Free Reading

Even if vocabulary instruction can improve reading comprehension, it does not appear to be the main source of word growth for school age children. Nagy and Anderson (1984) observed that "even the most ruthlessly systematic direct vocabulary instruction could [not] account for a significant proportion of all the words the children actually learn" (p. 304). Evidence for the impact of reading on vocabulary acquisition comes from both experimental and correlational studies. In "read-and-test" experiments (discussed further below), subjects are given a text with unknown words in it and asked to read it for comprehension. They are then given a (usually surprise) test on the meanings of the new words. These studies have found that a small but reliable amount of knowledge is gained from even a single exposure to an unknown word (Swanborn & de Glopper, 1999). Nagy, Herman, and Anderson (1985) estimated that with a sufficient amount of reading, a seemingly low "pick-up" rate could account for most of the observed growth in vocabulary knowledge among school-age readers.

Additional experimental evidence comes from sustained silent reading (SSR) and extensive reading studies, in which students are encouraged to read books they select for themselves. These studies have been conducted with children and adults, for both first and second language readers. Krashen's (2004a) narrative review of 54 studies concluded that SSR and extensive reading treatments were as good as or better than traditional language arts and reading instruction in promoting vocabulary and reading comprehension gains. Two meta-analyses of SSR and extensive reading studies have found significant, medium-to-large effects for free reading. Krashen (2007) examined studies for teens and young adults and found a large effect (d= .87) on comprehension tests. Jeon and Day (2016) found medium effect sizes for both vocabulary (d= .47) and comprehension (d= .54) for studies of both adults and children.[1]

Time Efficiency in Instruction

Many vocabulary interventions are aimed at helping low-performing students "catch up" to their age peers in reading, presumably in the most time-efficient way possible. Carlo, August, McLaughlin, Snow, Dressler, Lippman, Lively, and White (2004), for example, argued that "gaining access to the information taught in middle and secondary content area classes requires all students exit elementary school with good reading comprehension," and therefore "closing this gap has a high priority if the U.S. education system is to fulfill its goad of reducing inequalities" (p. 188, 190, emphasis added). Lawrence, Rolland, Branum-Martin, and Snow (2014) claimed that "without proper intervention, lower-skilled students are likely to fall further behind their more skilled peers in academic domains" (p. 77, emphasis added). Faw and Waller (1976) noted that despite the presumed goal of efficiency, most educational intervention studies lack any study or instructional time variable in their analyses. They argued:

> *It is absurd to think that psychologists and educators can be content with improving subjects' learning and retention of textual materials if the altered performance is simply a function of augmented study time. This would be analogous to attributing the increased length of a skier's jump to superior coaching when, in fact, the coach had simply provided a steeper and longer hill from which the jump could be made. (p. 703)*

Faw and Waller proposed that researchers distinguish between absolute performance levels and measures of efficiency. Absolute performance measures look only at the amount of learning that has taken place during the intervention, such as gain scores from pre-test to post-test. An efficiency measure takes these absolute gains and divides them by the study time of the intervention, to yield a gains-per-time estimate. It is then possible to calculate an efficiency index to compare the two approaches by dividing the efficiency score of the experimental group by the efficiency score of the control group. Faw and Waller point out that methods that produce greater absolute gains may in fact be less efficient than the alternatives.

Only handful of studies on vocabulary acquisition have applied the principles laid out by Faw and Waller. Krashen (1989) re-analyzed several vocabulary-teaching studies for word learning efficiency. Several studies by Mason and colleagues also reported the number of words per minute acquired in a second language classrooms in order to compare the efficiency of traditional instruction with comprehension-based language teaching methods (Mason, 2007; Mason & Krashen, 2004; Mason, Vanata, Jander, Borsch, & Krashen, 2008).

Research Questions

My analysis of the relative efficiency of direct instruction and free reading on word knowledge growth is organized around three questions:

1. What is the average efficiency of explicit vocabulary teaching for school-age students as measured in words learned per minute of instruction?

2. What is the average efficiency of free reading in words acquired per minute of reading, based on previous studies of incidental word acquisition rates, the percentage of unknown words in text, and average reading rates for students?

3. What is the relative efficiency of direct instruction compared to free reading, as measured by an efficiency index?

The Efficiency of Direct Instruction

The most recent comprehensive review on vocabulary instruction is Wright and Cervetti (2017), who reviewed the results of 36 studies on the effects of vocabulary instruction on word learning and reading comprehension. They began with the studies selected by two previous meta-analyses of vocabulary teaching interventions (Elleman et al., 2009; Stahl & Fairbanks, 1986), supplementing their pool of studies with those published after the Ellman et al. review. Their inclusion criteria differed somewhat from previous reviews. To be selected for the review, studies had to include PreK-12 students, a treatment involving the direct instruction of words, the teaching of "word-solving" strategies, or both, and a passage comprehension dependent measure.

Their analysis reported estimates of the time spent on vocabulary instruction in each study. These instructional times, however, are for the amount of time devoted to each word taught. For an efficiency measure, we need the time spent per word learned (Faw & Waller, 1976). In this analysis, I took the number of words learned in the intervention divided by the total instructional time (in minutes) of the vocabulary instruction, similar to Krashen (1989) and Mason (2007). For example, if 100 minutes were spent on instruction, and the gain score for vocabulary was 5 words, the efficiency estimate would be .05 words per minute (wpm) (5 words/100 minutes). An efficiency score is calculated for each study.

Gain scores for studies without comparison or control groups were calculated by subtracting the pre-test vocabulary scores from the post-test scores. Studies that had a control group but no information on whether the controls also received vocabulary instruction ("business-as-usual" or "typical practice" groups) were treated the same as those with no control groups. If a study had a reading-only comparison group, the gain scores of the comparison group were subtracted from gain scores of the treatment group, resulting in a "net" gain estimate.

Study Selection

Of the 36 studies included in the Wright and Cervetti review, nine studies included interventions aimed at teaching word learning strategies only, not a specific set of target words. Of the remaining group of 27 studies, I excluded 13 studies that failed to meet one of the following four criteria: (a) included a measure of target word knowledge due to the intervention; (b) included a pre-test or a no-treatment comparison group to control for pre-treatment knowledge of the target words; (c) provided sufficient data on instructional time to calculate efficiency; and/or (d) included subjects who would likely be able to read independently (grade 2 or older). This left 14 studies with sufficient data to calculate an efficiency estimate, listed in Tables 1 and 2.[2]

In cases where there was more than one experimental group and significant differences were found among them, the explicit instruction group that had the highest vocabulary gain scores was used in the calculations in order to provide the "best-case scenario" for explicit instruction. When there were multiple treatments and no significant differences found in their gain scores, I took the average gain for all the experimental conditions. Wright and Cervetti categorized the studies by the length of the intervention, with a "brief" intervention lasting four weeks or less, and a "long-term" intervention lasting more than four weeks (p. 209). Table 1 summarizes the data used to calculate time efficiency from the nine brief interventions. The data from the five long-term studies are found in Table 2.

Time Estimates

I used the "per word" instructional times provided by Wright and Cervetti (2017) in six of the 14 studies, taken from their Tables 2 and 4 (pp. 9, 15). For the other eight experiments, I used a different estimate, justified below. In all cases, my revised estimates were lower than those given by Wright and Cervetti. Total words learned was calculated by multiplying the raw score increase, pre- to post-test, by the quotient of total words divided by total words tested. For example, in Leseaux et al. (2010), there were 72 words taught but only 36 words tested. The raw score was multiplied by two (72/36) to yield the total words gained.
Bos & Anders (1990). Bos and Anders report that each intervention consisted of six 50-minute sessions over a period of seven weeks: three "practice" sessions and three experimental sections. Although Wright and Cervetti (2017) count both the practice and the experimental sessions for their time estimates, I have used only the experimental ones, for a total of 110 minutes of instruction across the three days.

Greene Brabham and Lynch-Brown (2002). Wright and Cervetti (2017) provided an estimate of 180 minutes for the experimental treatments (4.5 minutes per word taught for the 40 words). Greene Brabham and Lynch-Brown reported that the highest scoring group, the interactive group, spent 25 minutes on each of two stories. Since the stories were read three times, I used an estimate of 150 minutes of instruction. I excluded their first-grade subjects since it wasn't clear they were able to read independently.

Hawkins, Musti-Rao, Hale, McGuire, and Hailley (2010). Hawkins and colleagues studied vocabulary acquisition using a within-subjects, post-test-only design with a group of fourth-grade students. The reading-only condition read three 400-word stories. Since no treatment time was reported for this condition, I began with the average 4th grade silent reading rate, which is estimated to be around 150 wpm (Carver, 1989; Spichtig, Hiebert, Vorstius, C., Pascoe, J., Pearson, P. D., & Radach, R., 2016). However, because students knew they were going to be quizzed on the content of the passages, I lowered the reading rate to 100 wpm, as students who know they are to be tested tend to read more slowly (Carver, 1990). I estimated four minutes was spent by the controls on reading the story (400 words/100 wpm).

The listening + vocabulary instruction condition produced the greatest absolute number of words gained. For the listening part of the treatment, I doubled the estimate of the reading-only condition (8 minutes), since students were asked to repeat each sentence after it was read by the teacher. For the added pre-reading vocabulary instruction, I estimated one minute per target word, which is what Coyne et al. (2009) used as a time estimate for giving word definitions in storybook reading treatments. Thus the total listening + vocabulary practice treatment time was 18 minutes (10 minutes instruction plus 8 minutes reading and repeating). Wright and Cervetti's estimate was "less than one minute" on each of the 30 target words.

Pany, Jenkins, and Schreck (1982) (Study 1). Only a range of per-word instructional times (two to ten minutes) was provided by Wright and Cervetti. Students saw two sets of target words in each condition. The highest scoring condition was "meanings practiced," which spent 6.5 minutes per set of words, for a total of 13 minutes.

Seifert and Espin (2012). Wright and Cervetti estimated that the researchers spent 12 minutes on each target word taught, for a total of 120 minutes. But it would appear from the procedures section of the study (p. 241) that students spent 30 minutes in each condition for each set of 10 target words, so the total time by condition was 30 minutes, not 120.

Tuinman and Brady (1974). Wright and Cervetti estimated 10 minutes per word taught (for 660 minutes), although the time reported for the treatment by the researchers was 585 minutes (14 instructional sessions of 45 minutes each, p. 179), so this lower estimate was used.

Beck, Perfetti, & McKeown (1982); McKeown, Beck, Omanson, and Perfetti (1983). An estimate of 15 minutes per word was given by Wright and Cervetti for both of these studies, for a total of 2,288 minutes for 104 target words. I used a slightly lower estimate of 2,250 minutes based upon Beck et al. description's of the intervention as consisting of 75 30-minute lessons.

Summary of Direct Instruction Studies

Tables 1 and 2 summarize the efficiency scores in words learned per minute for the 14 direct instruction studies taken from Wright and Cervetti's (2017) review. Also listed are the grade level, sample size, treatment duration (in minutes), number of words learned, and type of vocabulary test (meaning recall or meaning recognition) for each experiment. For short interventions, the average number of words learned per minute was .07. For long-term interventions, the average number of words gained per minute was much smaller, at .01. There was a large standard deviation for the short-term studies, indicating considerable variability in efficiency scores.

Incidental Word Acquisition from Reading

Several studies of K-12 and adult readers have measured the amount of vocabulary gained incidentally from reading. As noted above, in these "read-and-test" experimental studies, subjects are given texts containing unknown words and told to read the texts for comprehension. They are then given a surprise vocabulary test on the unknown words, either immediately or after some delay. We can use these data to estimate the number of unknown words a typical reader might acquire through reading for pleasure, given a certain percentage of unknown words in the text. Combined with data on reading rate, we can then estimate the number of words per minute that could have been gained through free reading for each of the 14 direct instruction studies.

Table 1. *Efficiency of Direct Instruction of Vocabulary in Long-Term Interventions*

Study	Grade Level (Sample Size)	Duration	Words Learned (Test Type)	Efficiency Score
Bos & Anders (1990)	7, 8 (*N*= 61)	110 minutes	7.86 (Recog.)	.07 wpm
Greene Brabham & Lynch-Brown (2002) (Grade 3)	3 (*N*= 129)	150 minutes	11.9 (Recog.)	.08 wpm
Hawkins et al. (2010)	4 (*N*= 21)	18 minutes	2.67 (Recall)	.15 wpm
McKeown et al. (1985)	4 (*N*= ?)	360 minutes	13.17 (Recog.)	.04 wpm
Nash & Snowling (2006)	2, 3 (*N*= 24)	360 minutes	5.14 (Recog.)	.014 wpm
Pany et al. (1982) (Study 1)	4 (*N*= 12	13 minutes	0.92 (Recall)	.07 wpm
Seifert & Espin (2012)	10 (*N*= 20)	30 minutes	3.1 (Recog.)	.10 wpm
Stahl (1983)	5 (*N*= 28)	75 minutes	8.32 (Recall)	.05 wpm
Tuinman & Brady (1974)	4, 5, 6 (*N*= 210)	585 minutes	12.18 (Recog.)	.02 wpm
Average				.07 wpm
(Std Dev)				(.04)

Recog.= Meaning recognition test.

Table 2. *Efficiency of Direct Instruction of Vocabulary in Long-Term Interventions*

Study	Grade Level (Sample Size)	Duration	Words Learned (Test Type)	Efficiency Score
Beck et al. (1982)	4 (*N*= 27)	2250 minutes	52.75 (Recog.)	.02 wpm
Lesaux et al. (2010)	6 (*N*= 476)	3240 minutes	12.08 (Recog.)	.004 wpm
Lesaux et al. (2014)	6 (*N*= 2082)	4095 minutes	9.62 (Recog.)	.002 wpm
McKeown et al. (1983)	4 (*N*= 82)	2250 minutes	48.59 (Recog.)	.02 wpm
Simmons et al. (2010)	4 (*N*= 903)	1620 minutes	21.98 (Recog.)	.01 wpm
Average				.01 wpm
(Std Dev)				(.009)

Recog.= Meaning recognition test.

I have included in Table 3 the studies from Swanborn and de Glopper's (1999) meta-analysis on incidental word acquisition among school-age readers. I excluded four unpublished dissertations used by Swanborn and de Glopper, but added one published study not in their review (Herman, Anderson, Pearson, & Nagy, 1997). I also excluded studies in which the researchers deliberately choose or manipulated the contexts around the target words in order to make them all either "informative" or "uninformative," since neither extreme is representative of natural texts. Only studies in which words were chosen solely on the basis of whether they were likely to be unknown to the reader regardless of context were used.

The studies include readers at every reading ability level, including less-able readers.3 I report the results by reading achievement level for those studies that provided a breakdown, taken in part from Swanborn and de Glopper's meta-analysis (Table 3, p. 273).4 Table 3 lists grade levels tested, the number of subjects, and the subjects' reading levels for each study. The final column of Table 3 reports the probability of acquiring an unknown word from a single exposure. Nagy et al. (1985) defines this probability as "the increase in the number of words divided by the number of words originally not known" (p. 248). In some studies, students were tested on both the target words that appeared in their assigned text and on words that appeared in a text they did not read. This was done instead of a pretest to control for prior knowledge of the target words. For these studies, I used Nagy et al.'s (1987) formula to calculate probability:

(Proportion of Target Words Correct – Proportion of Control Words Correct) / (1 - Proportion of Control Words Correct)

In those studies where a pretest was used instead of control words, I followed a similar formula, subtracting the proportion of correct pretest words from the proportion of correct post-test words, and then dividing that result by the proportion of incorrect pretest words.

Table 3. *Incidental Acquisition Pickup Rates in 12 Studies*

Study	Grade Level (Sample Size)	Subjects' Reading Level[a]	Acquisition Probability
Herman, Anderson, Pearson, & Nagy (1987) (81st to 99th Percentile)	8 (*N*= 413)	H	0.26
Herman, Anderson, Pearson, & Nagy (1987) (31st to 80th Percentile)	8	A	0.12
Herman, Anderson, Pearson, & Nagy (1987) (3rd to 30th Percentile)	8	L	0.05
Nagy, Anderson, & Herman (1987)	3, 5, 7 (*N*= 352)	H, A, L	0.05
Nagy, Herman, & Anderson (1985)	8 (*N*= 57)	H, A	0.11
Schwanenflugel, Stahl, & McFalls (1997)	4 (*N*= 43)	H, A, L	0.08
Stahl (1989)	6 (*N*= 182)	H, A, L	0.13
Shu, Anderson, & Zhang (1995) (English experiment)	3, 5 (*N*= 170)	H, A, L	0.10

[a] Subjects' Reading Level: H = High, A = Average, L = Low, taken in part from Swanborn & de Glopper (1999). Scores from both meaning recall and meaning recognition measures were averaged to calculate probability if data on both were provided.

The probability of acquiring an unknown word incidentally through reading ranged from .05 to .26. In their meta-analysis, Swanborn and de Glopper (1999) calculated the average probability of acquisition to be .15 for the 15 experiments

they included. The average probability of acquisition for the eight experiments in Table 3 is slightly lower, at .11. What is the appropriate acquisition rate for comparing free reading to direct instruction? Since the goal of direct instruction is often to help low-achieving students, the most conservative approach is to use one of the lower estimates. In Table 3, we find that the lowest probability estimate is .05. This is the figure used for the efficiency scores calculations below

Percentage of Unknown Words in Text

The number of words that a reader can acquire incidentally from reading depends in part on how many unknown words are present in the text. Anderson and Freebody (unpublished, reported in Nagy et al., 1985) estimated the number of unknown words likely to be encountered by a "50th percentile fifth-grader" in text is between three and six percent, depending on the criteria used for "knowing" a word (p. 250). The researchers did not specify the source of the texts analyzed. Stahl (1990) stated that "a reader typically encounters between one and a half and three unknown words per hundred running words" (p. 6), but he gave no source for his estimate.

Carver's (1994) attempted to determine the percentage of unknown words by asking third through sixth grade students to circle words they did not know in a set of passages. Students first were tested to determine their current reading level, and then given passages to read that were below, at, or above their grade level. Passages were taken from both textbooks and library or trade books (p. 416). Carver noted that all of the percentages are likely to be underestimates, however. A large number (40%) of his initial sample failed to underline three embedded low-frequency words and had to be excluded from the study, suggesting that students had a tendency to under-identify unknown words in the texts.

Since Carver's (1994) estimates are the best documented, I have used for my calculations the average number of unknown words he found for library books, from two grade levels above grade level (3.35%), at grade level (1.30%), and two grade levels below reading level (1.35%), giving us an estimate of 2%. This number falls at the lower end of the range given by Stahl (1990), and slightly below the low end of Anderson and Freebody's (cited in Nagy et al., 1985) results.

Reading Rates

Efficiency calculations for incidental vocabulary acquisition depend in part on the reader's reading rate. The most recent large-scale study of reading rates was a partial replication of Taylor (1965) by Spichtig et al. (2016). Like Taylor, Spichtig and colleagues measured reading rate along with comprehension and eye movements for a large sample (*N*= 2,203) of students, but limited their study to grades 2, 4, 5, 8, 10, and 12. Unlike Taylor's study, the researchers attempted to stratify their sample to reflect the current demographics of U.S. schools, but it was not a random sample.

Table 4 (column 2) lists the mean reading rates, controlling for comprehension, reported in Spichtig et al. The researchers also reported rates by quartile, so I have taken the average of the bottom two quartiles in order to provide an approximate "below average" or "slow" reading rate at each level (column 3). Since again vocabulary instruction is often advocated especially for less-able readers, I will use these lower rates in making efficiency estimates when comparing acquisition rates to direct instruction studies in the following section, even when it appears the actual subjects in the study were average or above average readers.

Free Reading Efficiency Scores

Having established a rate of acquisition (.05), a percentage of unknown words typically found in text (2%), and average and low-achieving reading rates for school-age children, we can now estimate the number of unknown words that would likely be acquired from free reading. Table 4 (columns 4 and 5) reports estimates, at various reading rates, of the number of words per minute likely to be acquired from reading under these assumptions

Table 4. *Estimated Incidental Word Acquisition from Free Reading at Average and Slow Reading Rates*

Grade Level	Average Reading Rate (*SD*)	Slow Reading Rate	Efficiency: Average Rate	Efficiency: Slow Rate
2	116.1 (41.7)	87	.12 wpm	.09 wpm
4	147.8 (45.4)	115	.15 wpm	.12 wpm
6	164.2 (54.4)	128	.16 wpm	.13 wpm
8	169.9 (51.8)	130	.17 wpm	.13 wpm
10	186.6 (53.4)	147	.19 wpm	.15 wpm
12	187.5 (55.5)	181	.19 wpm	.18 wpm

Measured in words per minute of reading, students become more efficient in word acquisition as they age, although the trend is not perfectly linear due to the plateauing of reading rates between grades 4 and 8. Students reading at an average rate for their grade level will acquire around .12 words per minute in grade 2, rising to .19 words per minute by grade 10. For low-achieving students, the efficiency of incidental acquisition goes from .09 words per minute at grade 2, up to .18 words per minute in grade 12.

Efficiency Indexes for Direct Instruction and Free Reading

Having calculated the efficiency scores for both direct instruction and free reading, we can now provide efficiency indexes to compare the two approaches for each of the direct instruction studies. Table 5 shows the efficiency scores estimates of direct instruction conditions and our hypothetical reading-only conditions. I took the estimated number of words gained and words per minute for each study in Cervetti and Wright (2017) as reported in Tables 1 and 2 above. For the incidental acquisition estimates, I used the "slow" reading speed for that grade level as found in Table 5. For studies that included odd-numbered grades, I used the estimate from the even-numbered grade below it (e.g. for grade 5, grade 4 reading speeds were used.). Shown also in Table 5 is the estimated number of words one could acquire from reading (an absolute measure), to compare to those gained in the direct instruction experiment (Table 5, column 4, "Reading Words Gained").

As was done in Faw and Waller (1996), the efficiency index for a study was calculated by dividing the efficiency score of experimental condition (direct instruction) by the efficiency score of the control condition (reading-only). An efficiency index score of 1.0 means the two approaches were equally efficient, a number smaller than 1.0 indicates incidental acquisition was more efficient, and a number

Table 5. *Efficiency of Direct Instruction versus Incidental Acquisition in 14 Studies*

Study	Duration	DI Words Gained[a]	Reading Words Gained	DI Efficiency Score	Reading Efficiency Score	Eff. Index[b]
Short-Term Studies:						
Bos & Anders (1990)	110	7.86	14.3	.07 wpm	.13 wpm	0.54
Greene Brabham & Lynch-Brown (2002)	150	11.9	13.5	.08 wpm	.09 wpm	0.89
Hawkins et al. (2010)	18	2.67	2.16	.15 wpm	.12 wpm	1.25
McKeown et al. (1985)	360	13.17	43.2	.04 wpm	.12 wpm	0.33
Nash & Snowling (2006)	360	5.14	32.4	.01 wpm	.09 wpm	0.16
Pany et al. (1982) (Study 1)	13	1.42	1.56	.07 wpm	.12 wpm	0.58
Seifert & Espin (2012)	30	3.1	4.5	.10 wpm	.15 wpm	0.67
Stahl (1983)	75	8.32	9	.05 wpm	.12 wpm	0.42
Tuinman & Brady (1974)	585	12.18	70.2	.02 wpm	.12 wpm	0.17
Average (SD)				.07 wpm (.04)	.12 wpm (.019)	0.56 (.35)
Long-Term Studies:						
Beck et al. (1982)	2250	52.75	270	.02 wpm	.12 wpm	0.17
Lesaux et al. (2010)	3240	12.08	421.2	.004 wpm	.13 wpm	0.03
Lesaux et al. (2014)	4095	9.62	573.3	.002 wpm	.14 wpm	0.01
McKeown et al. (1983)	2250	48.59	270	.02 wpm	.12 wpm	0.17
Simmons et al. (2010)	1620	21.98	145.8	.01 wpm	.09 wpm	0.11
Average (Standard Deviation)				.01 wpm (.009)	.12 wpm (.019)	0.10 (.08)

[a] DI = direct instruction. [b] Eff. Index = efficiency index.

greater than 1.0 means direct instruction was more efficient. Results are reported in Table 5, as in Tables 1 and 2 above, by Wright and Cervetti's classification of "short-term" and "long-term" treatments.

Short-term direct instruction studies had an average rate of word learning of .07 wpm. The average rate of word acquisition for free reading was .12 wpm, resulting in an average efficiency index of .56, favorable to free reading. This difference is statistically significant ($t(16)= 3.36$, $p< .01$), yielding a large effect size ($d= 1.58$). Put another way, reading would be on average 1.7 times (.12 wpm/.07 wpm) more efficient than direct instruction in word acquisition. Direct instruction was found more efficient than reading in only one of the nine short-term comparisons.

For long-term studies, the results favor incidental acquisition to an even greater degree. The average efficiency of long-term direct instruction treatments was .01 wpm, compared to .12 wpm for incidental acquisition, and the average efficiency index was .10. As with the short-terms studies, this difference is statistically significant ($t(8)= 11.83$, $p< .0001$), with a very large effect size ($d= 7.48$). This means free reading would be about 12 times (.12 wpm/.01 wpm) more efficient than direct instruction in helping children acquire new vocabulary in long-term treatments.

There was a moderate negative correlation between the time devoted to instruction and the efficiency estimate ($r= -.67$, $p<.01$), meaning that the more time teachers spent on vocabulary instruction, the fewer number of words per minute their students learned.

Discussion

Our results indicate that neither short-term nor long-term instruction is efficient in teaching new words compared to just reading. Students in short-term direct instruction treatments learned about four words per hour (.07 wpm), compared to our estimate of around seven words per hour (.12 wpm) via free reading. For long-term treatments, students learn only about a .5 words per hour of instruction (.01 wpm).

Greater investments of time into vocabulary instruction appear to have diminishing returns in terms of the number of words students learn. It is clear from the estimates in Table 5 that long-term interventions were considerably less efficient than short-term ones. The least efficient instruction was found in those studies (Leseaux et al., 2010; 2014) that focused on teaching words from the Academic Word List. Students learned only one new AWL word every five and a half hours or so of instruction. At this rate, students would need to spend roughly 1,600 hours of instruction to learn just half of the 570 AWL terms, a feat that would take a decade or more of language arts classes devoted to nothing but vocabulary teaching.

Some Objections

Hypothetical Comparison Groups. Since nearly all of the direct instruction studies we examined lacked a reading-only condition, we do not have a set of "head-to-head" comparisons of incidental acquisition versus direct instruction. But there is no reason to think the subjects in the studies reviewed here, had they been given the opportunity to read, would not also have acquired vocabulary at rates found in previous incidental acquisition studies. Confidence in our findings is bolstered by the results of two reviews of studies that did include direct comparisons between a reading-only and a reading plus explicit instruction condition. McQuillan (2016a) reviewed eight such experiments with adult second-language acquirers, and found that reading-only conditions had the same or greater efficiency than direct instruction in six experiments, and lower efficiency in two treatments. McQuillan (2019) looked at five studies that compared direct instruction and reading-only conditions in storybook reading for young children. The reading-only conditions were found on average to be 66% more efficient than direct instruction conditions for acquiring new words.

Overly Optimistic Estimates. While I have used conservative estimates consistent with previous studies of incidental acquisition rates, percentage of unknown words, and reading

rate, some might argue they should be even lower. To provide an even stricter test for my assumptions, I ran a separate analysis of the efficiency data in Table 5 in which I halved the incidental reading acquisition efficiency estimates in all the comparisons, the equivalent of lowering the estimate of unknown words in a text to 1% or the probability of acquisition to .025. Even under these very pessimistic assumptions for incidental acquisition, free reading was still significantly more efficient than direct instruction for the long-term studies ($t(8)= 8.61$, $p< .0001$). For the short-term studies, the revised average efficiency estimate for free reading was .06 wpm, which was not significantly different from the direct instruction estimate of .07 wpm ($t(16)= 0.46$, $p= .665$). Free reading, then, was as good as or better than direct instruction in promoting word growth regardless of the study length.

In any such "tie" between free reading and direct instruction, the advantage clearly belongs to free reading, since reading is less work for the teacher and more enjoyable for the student. In addition, free reading has important benefits in addition to vocabulary, such as improving reading comprehension, writing, and grammar (Krashen, 2004a). Nagy et al. (1985) also make this point, noting that "[a]ny comparison of approaches ought to take account of the fact that time spent in reading has more benefits than just growth in vocabulary...no doubt the ancillary benefits of vocabulary instruction are less rich" (p. 251).

An argument can be made that the probability estimates such as those provided by the read-and-test studies should not be applied in the way done in this analysis. Our analysis assumes that acquisition is incremental, meaning that we pick up a small percentage of an unknown word's meaning each time we see it in a text. As Stephen Krashen (personal communication) has pointed out to me, read-and-test researchers such as Nagy et al. (1985) seem to assume that a .05 probability means that out of 100 unknown words in a text, a reader acquires the full meaning of five of them. This is not the same as saying that you pick up partial meanings (say, 5%) of 100 words. If the former interpretation accurately represents these researchers' reasoning, then multiplying a probability by the number of word occurrences may in fact be inappropriate and overestimate vocabulary gains.

Still, our finding that free reading is superior to direct instruction for word acquisition is consistent with analyses that used a very different approach to the problem, that of corpus analysis. Nation (2014), for example, looked at a large corpus of classic novels to determine how much one would have to read to have a reasonable chance of acquiring new words. Potentially unknown words that occurred 12 or more times in the text were considered "acquired" in Nation's study. For texts written with 98% vocabulary coverage in the 3,000- through 8,000-word-family levels (the levels at which popular, young adult fiction is written (McQuillan, 2016b)), Nation estimated that on average 6.9 words would be acquired per hour, or 0.12 wpm. This is identical to our overall incidental acquisition estimate for school-age readers. Similarly, McQuillan (in press) examined a corpus of 1.2 million words from young adult popular fiction and, using a similar "cut-off" method as used by Nation, found that reading was between two and six times more efficient than direct instruction in acquiring academic vocabulary.5

Short-Term Benefits of Instruction. Since Wright and Cervetti (2017) concluded that "pre-teaching" vocabulary appears in a text improves comprehension of that text, it could be argued that direct instruction interventions have real benefits when used in this short-term, "text-by-text" approach. The weakness of this argument is that such teaching is by definition only a temporary fix, akin to bailing out a sinking boat with a bucket instead of fixing the hole at the bottom. Stahl (1990) made a similar observation, noting that "[m]ore intensive instruction is going to take away time from other activities, including wide reading that will not only better allow them to solidify their vocabulary gains but also will itself lead to greater vocabulary growth" (p. 11). Allington, McCuiston, and Billen (2015) argued that students need to read texts they can understand independently for real progress in reading to occur, preferably on topics that are compelling and comprehensible to them (Krashen et al., 2018). One solution, proposed by Krashen (2004a) and others, is to give students the opportunity to read extensively by providing a large number of interesting and comprehensible books in the school and classroom library, time to read daily, and a comfortable environment. Such programs have been found to be as good as or better than traditional instruction in promoting both vocabulary growth and reading comprehension (Krashen, 2004a; Krashen & Mason, 2017).

Free Reading as a Bridge to Academic English

I do not claim that free reading alone can give students 100% of the vocabulary or academic language needed for success in school. Some explicit teaching of terms related to new concepts, for example, may be required. More importantly, there are characteristics of academic language that are only found in academic texts (Biber, 1985), and therefore can only be acquired through academic reading (Krashen, 2010). Pleasure reading can, however, provide an important "bridge" to more challenging school reading, including sub-technical vocabulary. Rolls and Rogers (2017) analyzed a large corpus of science fiction and fantasy literature for the presence of sub-technical vocabulary specific to the sciences, based on Coxhead and Hirsh's (2007) list of 318 word families (e.g. degrade, module, uptake). They found that nearly all of the words (92%) occurred at least once in a corpus of one million words, and majority of those words occurred six times of more, giving students a good chance to acquire them.

Krashen (2012a; 2012b) advocates a two-stage approach for helping students advance in both academic vocabulary and content knowledge via free reading. Stage 1 consists of "massive, but not necessarily wide, self-selected voluntary reading" (Krashen, 2012b, p. 9). Reading done at this stage builds general vocabulary and knowledge of the world that will make academic reading more comprehensible. Ideally, students read narrowly in order to take advantage of prior knowledge of a topic or book series (Krashen, 2004b; Kyungho & Nation, 1989; Schmitt & Carter, 2000).

Since general reading will not give students all of the academic language they need, Krashen proposes a second stage called narrow academic reading. This consists of students reading about an academic topic that they themselves are interested in. This sort of reading will give students knowledge of academic conventions and language (such as the words on the AWL) that in turn will help them across disciplines. Krashen (2012b) gives his own case history of narrow reading in linguistics and medicine that gave him sufficient knowledge of the academic language register to read scholarship in other fields. Indeed, it would seem that nearly all of us acquired academic language in this way, and not through direct instruction (Krashen, 2012a). Providing low-achieving students with an opportunity to follow that same path should at least be considered.

Refrences

Allington, R. (2014). How reading volume affects both reading fluency and reading achievement. *International Electronic Journal of Elementary Education, 7*(1), 13-26.

Allington, R., McCuiston, K., & Billen, M. (2015). What the research says about text complexity and learning to read. *The Reading Teacher, 68*(7), 491-501.

Anderson, R., & Freebody, P. (1981). Vocabulary knowledge. In J.T. Guthrie (Ed.), *Comprehension and teaching: Researcher Reviews* (pp. 77-117). Newark, DE: International Reading Association.

Apthorp, H. (2006). Effects of a supplemental vocabulary program on third-grade reading/language arts. *The Journal of Educational Research, 100*(2), 67-79.

Apthorp, H., Randel, B., Cherasaro, T., Clark, T., McKeown, M., & Beck, I. (2012). Effects of a supplemental vocabulary program on word knowledge and passage comprehension. *Journal of Research on Educational Effectiveness, 5*(2), 160-188.

Beck, I., McKeown, M., & Omanson, R. (1987). The effects and uses of diverse vocabulary instructional techniques. In M. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 147-163). Hillsdale, NJ: Erlbaum.

Beck, I., Perfetti, C., & McKeown, M. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*(4), 506-521.

Biber, D. (1985). Investigating macroscopic textual variation through multi-feature/multidimensional analyses. *Linguistics, 23*(2), 337-360.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*(1), 44-62.

Bos, C., & Anders, P. (1990). Effects of vocabulary instruction on the vocabulary learning and reading comprehension of junior-high learning disabled students. *Learning Disability Quarterly, 13*(1), 31-42.

Carlo, M., August, D., McLaughlin, B., Snow, C., Dressler, C., Lippman, D, Lively, T., & White, C. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*(2), 188-215.

Carver, R. (1989). Silent reading rates in grade equivalents. *Journal of Reading Behavior, 21*(2), 155-166.

Carver, R. (1990). *Reading rate: A review of research and theory.* New York: Academic Press.

Carver, R. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Reading Behavior, 26*(4), 413-437.

Cho, K.S., & Krashen, S. (1994). Acquisition of vocabulary from Sweet Valley Kids series: Adult ESL acquisition. *Journal of Reading, 37*(8), 662-667.

Cowan, J. (1974). Lexical and syntactic research for the design of EFL reading materials. *TESOL Quarterly, 8*(4), 389-400.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

Coxhead, A., & Hirsch, D. (2007). A pilot science-specific word list. *Revue Française de Linguistique Appliquée, 12*(2), 65-78.

Coyne, M., McCoach, D., Loftus, S., Zipoli Jr, R., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *The Elementary School Journal, 110*(1), 1-18.

Elleman, A., Lindo, E., Morphy, P., & Compton, D. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness, 2*(1), 1-44.

Faw, H., & Waller, T., (1976). Mathemagenic behaviors and efficiency in learning from prose. *Review of Educational Research, 46*, 691-720.

Greene Babham, E., & Lynch-Brown, C. (2002). Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. *Journal of Educational Psychology, 94*(3), 465-473.

Hawkins, R., Musti-Rao, S., Hale, A., McGuire, S., & Hailley, J. (2010). Examining listening previewing as a class-wide strategy to promote reading comprehension and vocabulary. *Psychology in the Schools, 47*(9), 903-916.

Herman, P., Anderson, R., Pearson, P., & Nagy, W. (1987). Incidental acquisition of word meaning from expositions with varied text features. *Reading Research Quarterly, 22*(3), 263-284.

Jeon, E., & Day, R. (2016). The effectiveness of ER on reading proficiency: A meta-analysis. *Reading in a Foreign Language, 28*(2), 246-265.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal, 73*(4), 440-464.

Krashen, S. (2001). More smoke and mirrors: A critique of he National Reading Panel report on fluency. *Phi Delta Kappan, 83*, 119-123.

Krashen, S. (2004a). T*he power of reading.* 2nd edition. Portsmouth, NH: Heinemann.

Krashen, S. (2004b). The case for narrow reading. *Language Magazine, 3,* 17–19.

Krashen, S. (2005). Is in-school free reading good for children? Why the National Reading Panel Report is (still) wrong. *Phi Delta Kappan, 86*, 444-447.

Krashen, S. (2007). Extensive reading in English as a foreign language by adolescents and young adults: A meta-analysis. *International Journal of Foreign Language Teaching, 3*(2), 23-29.

Krashen, S. (2010). Academic language proficiency: Acquired or learned? In *Selected papers from the Nineteenth International Symposium on English Teaching* (pp. 34-43). Taipei: Crane.

Krashen, S. (2012a). Direct instruction of academic vocabulary: What about real reading? *Reading Research Quarterly, 47*(3), 223-224.

Krashen, S. (2012b). Developing academic language: Some hypotheses. *International Journal of Foreign Language Teaching, 7*(2), 8-15.

Krashen, S. & Mason, B. (2017). Sustained silent reading in foreign language education: An update. *Turkish Online Journal of English Language Teaching (TOJELT), 2*(2), 70-73.

Kyongho, H., & Nation, P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language, 6*(1), 323-335.

Lawence, J., White, C., & Snow, C. (2010). The words students need. *Educational Leadership, 68*(2), 22-26.

Lawrence, J., Rolland, R., Branum-Martin, L., & Snow, C. (2014). Generating vocabulary knowledge for at-risk middle school readers: Contrasting program effects and growth trajectories. *Journal of Education for Students Placed at Risk, 19*, 76-97.

Lesaux, N., Kieffer, M., Faller, S., & Kelley, J. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly, 45*(2), 196-228.

Lesaux, N., Kieffer, M., Kelley, J., & Harris, J. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal, 51*(6), 1159-1194.

Mason, B. (2007). The efficiency of self-selected reading and hearing stories on adult second language acquisition. In *Selected papers from the sixteenth International Symposium on English Teaching* (pp. 630-633). Taipei: Crane Publishing Company.

Mason, B., & Krashen, S. (2004). Is form-focused vocabulary instruction worthwhile? *RELC Journal, 35*(2), 179-185.

Mason, B., Vanata, M., Jander, K., Borsch, R., & Krashen, S. (2008). The effects and efficiency of hearing stories on vocabulary acquisition by students of German as a second foreign language in Japan. *Indonesian Journal of English Language Teaching, 5*(1), 1-14.

McKeown, M., Beck, I., Omanson, R., & Perfetti, C. (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Reading Behavior, 15*(1), 3-18.

McKeown, M., Beck, I., Omanson, R., & Pople, M. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly, 20*(5), 522-535.

McQuillan, J. (2016a). Time, texts, and teaching in vocabulary acquisition: A rebuttal to Cobb (2016). *Reading in a Foreign Language, 28*(2), 307-318.

McQuillan, J. (2016b). What can readers read after graded readers? *Reading in a Foreign Language, 28*(1), 63-78.

McQuillan, J. (2019). We don't need no stinkin' exercises: The impact of extended instruction and storybook reading on vocabulary acquisition. *Language and Language Teaching, 8*(1), 25-37.

McQuillan, J. (in press). Where do we get our academic vocabulary? Comparing the efficiency of direct instruction and free voluntary reading. T*he Reading Matrix.*

Ming-Tzu, K. W., & Nation, P. (2004). Word meaning in academic English: Homography in the academic word list. *Applied Linguistics, 25*(3), 291-314.

Nagy, W., & Anderson, R. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*(3), 304-330.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), 91-108.

Nagy, W., Anderson, R., & Herman, P. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*(2), 237-270.

Nagy, W., Herman P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly, 20*(2), 233-253.

Nash, H., & Snowling, M. (2006). Teaching new words to children with poor existing vocabulary: A controlled evaluation of the definition and context methods. *International Journal of Language Communication Disorders, 41*(3), 335-354.

Nation, P. (2014). How much input do you need to learn the most frequent 9,000 words?. *Reading in a Foreign Language, 26*(2), 1-16.

National Reading Panel. (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Washington, D.C.: National Institute of Child Health and Human Development, National Institutes of Health.

Pany, D., & Jenkins, J. (1978). Learning word meanings: A comparison of instructional procedures. *Learning Disability Quarterly, 1*(2), 21-32.

Pany, D., Jenkins, J., & Schreck, J. (1982). Vocabulary instruction: Effects on word knowledge and reading comprehension. *Learning Disability Quarterly, 5*(3), 202-215.

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do "Things Fall Apart"? *Reading in a Foreign Language, 22*(1), 31-55.

Rolls, H., & Rodgers, M. P. (2017). Science-specific technical vocabulary in science fiction-fantasy texts: A case for 'language through literature'. *English for Specific Purposes, 48*, 44-56.

Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal, 9*(1), 4-9.

Schwanenflugel, P. J., Stahl, S. A., & Mcfalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research, 29*(4), 531-553.

Seifert, K., & Espin, C. (2012). Improving reading of science text for secondary students with learning disabilities: Reading, vocabulary learning, and combined approaches to instruction. *Learning Disability Quarterly, 35*(4), 236-247.

Shu, H., Anderson, R., & Zhang, H. (1995). Incidental learning of word meanings while reading: A Chinese and American cross-cultural study. *Reading Research Quarterly, 30*(1), 76-95.

Simmons, D., Hairrell, A., Edmonds, M., Vaughn, S., Larsen, R., Willson, V., Rupley, W., & Byrns, G. (2010). A comparison of multiple-strategy methods: Effects on fourth-grade students' general and content-specific reading comprehension and vocabulary development. *Journal of Research on Educational Effectiveness, 3*(2), 121-156.

Snow, C., Lawrence, J., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness, 2*(4), 325-344.

Spichtig, A., Hiebert, E., Vorstius, C., Pascoe, J., Pearson, P. D., & Radach, R. (2016). The decline of comprehension-based silent reading efficiency the United States: A comparison of current data with performance in 1960. *Reading Research Quarterly, 51*(2), 239-259.

Stahl, S. (1983). Differential word knowledge and reading comprehension. *Journal of Reading Behavior, 15*(4), 33-50.

Stahl, S. (1989). Task variations and prior knowledge in learning word meanings from context. In S. McCormick, & J. Zutell (Eds), *Cognitive and social perspectives for literacy research and instruction: Thirty-eighth yearbook of the National Reading Conference* (pp. 197-204). Chicago: National Reading Conference.

Stahl, S. (1990). *Beyond the instrumentalist hypothesis: Some relationships between word meanings and comprehension.* Urbana, IL: Center for the Study of Reading.

Stahl, S., & Fairbanks, M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*(1), 72-110.

Stahl, S., & Nagy, W. (2007). *Teaching word meanings.* New York: Routledge.

Taylor, S.E. (1965). Eye movements in reading: Facts and fallacies. *American Educational Research Journal, 2,* 187-202.

Tuinman, J., & Brady, M. (1974). How does vocabulary account for variance on reading comprehension tests? A preliminary instructional analysis. In P. Nacke (Ed.), *Twenty-third Yearbook of the National Reading Conference* (pp. 176-184). Clemson, SC: National Reading Conference.

Wright, T., & Cervetti. G. (2017). A systematic review of research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly, 52*(2), 203-226.

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review, 57*(4), 541-572.

Notes

1. The National Reading Panel (2000) concluded that SSR and extensive reading programs do not help students become better readers, but as Krashen (2001, 2005) pointed out, the Panel omitted several studies of silent reading programs in its review, and misreported the results of some of those they did examine. Lewis and Samuels (2005, reported in Allington, 2014) conducted a review of 49 SSR studies, and reached a similar conclusion to Krashen's (2004a): "No study reported significant negative results; in no instance did allowing students time for independent reading result in a decrease in reading achievement" (p. 17). The eight "true experiment" studies Lewis and Samuels included in a meta-analysis had an effect size favoring free reading (*d*= .42).

2. Apthorp et al. (2012) included a comparison group that received a different, less intensive form of vocabulary instruction. No raw post-test scores on the target words were reported, only the hierarchical linear modeling results (Table 7, p. 174). Apthorp et al. claimed that the significant effects of the intervention held even when taking into account instructional time (p. 173), but this merely indicates that the experimental form of direct instruction was more efficient than the comparison form of vocabulary teaching. No reading-only comparison group was used.

3. I was unable to locate a read-and-test study of incidental acquisition among second language or language minority K-12 students. However, studies of adult L2 vocabulary acquisition have reported similar probability of acquisition estimates as those found in Table 3, ranging from .05 (Zahar, Cobb, & Spada, 2001) to .17 (Pellicer-Sanchez & Schmitt, 2010).

4. Two studies (Nagy, Anderson, & Herman, 1987; Shu, Anderson, & Zhang, 1995) also reported acquisition rates by the "conceptual difficulty" of the target word. The highest difficulty rating ("Level 4") was given to words that required new factual information to be understood. Nagy et al. found that none of the Level 4 words in the passages read by the students in their study were acquired, while Shu et al. found with a similar group of subjects that the probability of acquisition for such words was .07, within the low-end of the range of probabilities reported in Table 3. Nagy et al.'s Level 4 words appear in part to be technical vocabulary, words specific to a given discipline (e.g. divide meaning "boundary between drainage systems" (Nagy et al., 1987, p. 250). These are not the type of words generally included in vocabulary teaching programs such as those reviewed by Wright and Cervetti (2017), most of which used "Tier 2" or sub-technical words.

In Herman et al.'s (1987) study, nearly half of the target words could be classified as conceptually difficult or requiring new factual information to understand, terms such as renal, floodplain, ventricle, oxbow lake, and aorta. Yet the probability of acquisition in Herman et al. (.10) was comparable to the results from Stahl (1989) (.13), which used only "difficult synonyms" for the target words. This is additional reason to suspect that Nagy et al.'s (1987) finding may be an outlier.

5. There is an obvious problem with determining the rates of incidental acquisition during self-selected reading conditions: how do you assess word gains when every subject is reading a different text? Cho and Krashen (1994) attempted in part to do this in their study of a group of adult second language subjects (N = 4). Each of their subjects read texts of her own choosing. Three of the subjects also underlined words they did not know. A fourth subject, Alma, was not part of the original reading study group, and did not underline any words while reading. She instead was given a pretest and post-test on 165 words that the other three subjects had consistently marked as unknown from the reading series all four subjects were using.

Alma was not aware she was going to participate in the reading study when she took the pretest, so there is little chance she would have attempted to study or memorize the words. Applying the formula from Nagy et al. (1987) used above to the data reported by Cho and Krashen (1994, Table 3), we find Alma's rate of acquisition was an impressive .56. Alma reported she did not use a dictionary at all during her reading. These results suggest that self-selected reading may yield much higher rates of acquisition than in the laboratory conditions typically used in other studies, where the texts are chosen by the researchers.