Review of Educational Research
Spring 1989, Vol. 59, No. 1, pp. 87–116

Whole Language and Language Experience Approaches for Beginning Reading: A Quantitative Research Synthesis

Steven A. Stahl and **Patricia D. Miller**
Western Illinois University

To examine the effects of whole language and language experience approaches on beginning reading achievement, a quantitative synthesis was performed on two data bases: the five projects conducted as part of the United States Office of Education (USOE) first grade studies and 46 additional studies comparing basal reading approaches to whole language or language experience approaches. The results of both analyses suggest that, overall, whole language/language experience approaches and basal reader approaches are approximately equal in their effects, with several exceptions. First, whole language/language experience approaches may be more effective in kindergarten than in first grade. Second, they may produce stronger effects on measures of word recognition than on measures of reading comprehension. Third, more recent studies show a trend toward stronger effects for the basal reading program relative to whole language/language experience methods. Fourth, whole language/language experience approaches produce weaker effects with populations labeled specifically as disadvantaged than they do with those not specifically labeled. Finally, studies with higher rated quality tend to produce lower effect sizes and the lowest effect sizes were found in studies that evaluated existing programs, as opposed to newly implemented experimental programs. These results are discussed within a stage model of reading that suggests that whole language/language experience approaches might be most effective for teaching functional aspects of reading, such as print concepts and expectations about reading, whereas more direct approaches might be better at helping students master word recognition skills prerequisite to effective comprehension.

For at least the better part of this century, there have been voices advocating that reading instruction begin in a natural manner, using the child's own language as a bridge to beginning reading instruction (see Hildreth, 1965). These approaches have been termed *activity approaches, informal approaches, language experience approaches*, or, most recently, *whole language approaches*. These terms represent an evolution of an idea, so that a whole language approach discussed today might be very different from an activity approach discussed by Wrightstone (1951), for example. They are also manifestations of a core approach to children's learning to read, namely that the child's attention should be focused on the communicative function of written language rather than on its form (Goodman, 1986; Goodman & Goodman, 1979; Harste, 1985; Newman, 1985).

An earlier version of this paper was presented at the annual meeting of the National Reading Conference. We would like to thank Patricia Hart, Richard C. Anderson, Peg Richek, Kathy Dulaney Barclay, and all of the anonymous reviewers for their comments and assistance.

Whole Language and Language Experience

Proponents of whole language approaches emphasize it as a philosophy rather than a specific method (e.g., Altwerger, Edelsky, & Flores, 1987; Goodman, 1986; Newman, 1985), making it difficult to define for review purposes. These authors and others (e.g., Grundin, 1986; Weaver, 1988) have stressed the continuity between the whole language movement and the language experience approach used in the 1960s and 1970s.

There are several commonalities between the two approaches. First, both approaches stress the importance of children's own language productions as a bridge from oral to written language. Second, both approaches decry the use of skill sequences to organize instruction, as is done in most basal reading programs. Third, both approaches use children's literature, rather than basal readers, for instruction. It should be noted that basal readers also use children's literature in their texts, often the same stories. The basal reading programs "adapt" the stories, with more or less fidelity to the original language. Whole language theorists, such as Goodman (1986), for example, suggest that such adaptation distorts the predictability of the language, making the stories harder, rather than easier, to understand (see also Simons & Ammon, 1988). Fourth, both language experience and whole language approaches stress the importance of focusing on the meaningfulness of language and of not changing the focus to parts of nonmeaningful segments of language, such as individual sound-symbol relationships, unless that instruction is done within the context of a whole text. In other words, both language experience and whole language advocates do not recommend teaching words or individual sound-symbol relationships in isolation, but may teach them as needed to help students understand particular texts.

Although there is much in common between the earlier language experience approaches and current whole language approaches, there are some important differences. First, in language experience approaches, experience charts generated by children's dictation were used as the major part of instruction. In whole language approaches, such charts are recommended (see Goodman, 1986; Weaver, 1988), but more emphasis is placed on the reading of tradebooks, especially those with predictable patterns (see Bridge, Winograd, & Haley, 1983). In language experience approaches, trade books were also used extensively. They were read to groups of beginning readers who were encouraged to read them independently. It was the charts, however, rather than the trade books that received the most emphasis. Second, in whole language programs, greater emphasis is placed on children's own writing using invented spelling, rather than their dictated charts. Although language experience approaches stressed the interrelation of all four language processes (reading, writing, speaking, and listening), they recommended delaying writing until children had mastered a corpus of sight words (R. V. Allen, 1976; R. G. Stauffer, 1969); both Allen and Stauffer do, though, mention invented spelling. In whole language approaches, children are encouraged to write even before they can read words because of the belief that writing develops from scribbling to invented spelling to mature writing (e.g., Harste, 1985; Weaver, 1988).

The goal of both approaches is to bring children into literacy in a "natural" way, by bridging the gap between children's own language competencies and written language. Thus, written language should be seen as functional from the very

beginning (R. V. Allen, 1976). Because it is functional, it is argued that children learn written language the way they do oral language, through exposure to a literate environment (Goodman & Goodman, 1979). In such an environment, children are led to realize that their language, and the ideas of others, can be written and thus read. R. V. Allen (1976) expressed this as follows: "What I can think about, I can talk about. What I can talk about, I can write. What I can write, I can read. I can read what I write and what other people can write for me to read" (p. 51).

These approaches are also based on the premise that speaking, listening, writing, and reading are interrelated and interdependent. Instruction in reading begins where children are, in terms of their ability to think with words, and it stimulates language development in all media of expression and reception, with the ultimate goal as reading the writings of others.

We see these approaches as being both part of a continuous evolution, yet having some possibly important differences. Therefore, we will refer to the entire range of approaches as *whole language/language experience approaches*. If we are specifically referring to a study that uses one name or the other, we will use that label only.

Stage Models of Reading Development

Downing (1979) suggests three phases in the acquisition of reading skill: a *cognitive* phase in which the child becomes aware of the tasks needed to become a skilled performer, a *mastering* phase in which the skill is practiced until mastery is achieved, and an *automaticity* phase in which the learner practices until the skill can be performed without conscious attention.

Chall (1983b) has a similar stage model. In her initial stage, she suggests that prior to formal reading instruction, children need to develop skills prerequisite to learning to read. These skills and concepts have been investigated extensively under the framework of *emergent literacy* (see Teale, 1987; Teale & Sulzby, 1986). These include knowledge of their language, concepts about print, expectations about the nature of reading, and so forth. Researchers in this area have found similarities as well as differences in the functions that written language serves in different home environments and how children use written language across a variety of socioeconomic status (SES) groups. For example, some development of print awareness seems to be common across different cultures. They have also found vast differences in the quantity of exposure children get to written language, especially story book reading. Adams (in press) estimates that some children may receive more than 1000 hours of storybook reading by the time they begin formal reading instruction, whereas others may receive none. The informal interaction between parent and child during storybook reading may serve to familiarize the child with many of the conventions of print, such as where text begins on a page, directionality, punctuation, as well as the register of written language. Without such basic concepts, students may experience what Downing (1979) calls *cognitive confusion* when presented with text in formal reading instruction.

Whole language approaches may demonstrate the relations between written and spoken language in several ways. The process of directly translating oral language to experience charts may demonstrate the directionality of print, various print conventions, the concept of what a written word is, and some sound-symbol correspondences. The use of enlarged books may also approximate at school the

types of interactions around print that takes place in storybook reading. In these ways, whole language approaches might serve to clear up initial confusions about the functions of reading, of how written words relate to spoken words, and so forth (see McCormick & Mason, 1986).

Following this cognitive stage, according to the models of Downing (1979) and Chall (1983b), are stages concerned with mastery and automaticity. During the mastery stage, children learn to accurately decode print. Recent reviews have suggested that this is best accomplished through direct instruction of sound-symbol correspondences, rather than more indirect approaches (see Adams, in press; Anderson, Hiebert, Scott, & Wilkinson, 1985; Chall, 1983a). In the automaticity stage, children learn to apply their decoding skills fluently and automatically. In this stage, children may benefit from reading large amounts of relatively easy, connected text, and through repeated readings of the same text (see Samuels, 1985).

It might be that whole language/language experience approaches may work best in the cognitive phase because they simulate the environment in which literate behaviors begin to emerge. They might not work as well during the mastering phase, because sound-symbol correspondence instruction in these approaches is unsystematic and indirect. In discussing whether whole language/language experience approaches are effective, it is important to examine what role they are intended to play, whether they are expected to provide this initial introduction to literacy or whether they are intended to provide systematic reading instruction. (The automaticity phase is beyond the scope of this review.)

Reviews of Language Experience Approaches Studies

Hall (1972, 1977) summarized the results of 42 studies regarding the effects of language experience approaches. Her methodology was to report the significant and nonsignificant effects from each study as they were found. She concludes that there is evidence "to support that the overall reading achievement of students who receive language experience instruction is satisfactory, and, in some cases, it is superior to the achievement of children instructed by other approaches" (1977, p. 24). Although this claim is appropriately modest, it may overstate the effectiveness of language experience approaches. If two methods are approximately equal in their effects, both methods will produce some significantly higher scores than the other, by chance alone.

A partial synthesis of whole language/language experience approaches research was done by Grundin (1985) as part of his review of *Becoming a Nation of Readers* (Anderson et al., 1985). As part of their overall review of our knowledge of reading instruction, Anderson et al. suggested that the effects of language experience approaches have been indifferent, at least compared with conventional basal reading approaches. Grundin took issue with this conclusion. He argued that Anderson et al.'s conclusion was based only on one study, Bond and Dykstra's (1967) review of the Cooperative Research Program in First Grade, that this study was dated, and that the early language experience approaches were very different from modern whole language approaches.

Grundin (1985) did not cite any more recent research comparing whole language/language experience approaches and basal reader approaches that should have been considered by Anderson et al. Instead, Grundin reanalyzed the Bond and Dykstra (1967) data, totaling the average adjusted Stanford Achievement Test scores for

each group, "ranking each approach according to how it compares to the overall mean of the study it is in" (p. 265), and comparing the average ranks for language experience/structure approaches, linguistic approaches, basal plus additional phonics, i.t.a. approaches, and basal-only approaches. Grundin interprets his reanalysis to show that whole language/language experience approaches were in fact the best performing of all five approaches.

We feel that Grundin's (1985) analysis was not the best approach to examine the Bond and Dykstra (1967) data. Bond and Dykstra's achievement test results were interval data, that is, they conveyed information both about the order of the effects and the magnitude of differences between the effects. By ranking, Grundin disregarded the information about magnitude and retained only the information about order, leading to possibly misleading conclusions. For example, it is possible to find sets of numbers whose average rank is identical but whose means are vastly different and to find sets in which the means are similar but the average ranks are markedly different. The average ranking procedure can both overestimate the importance of small differences and underestimate that of larger differences (the differences between 29 and 30 and between 30 and 70 both count as differences of one rank point). In addition, Grundin did not provide any information about variations between studies. Because the different studies varied from the means for each method, small differences between means may well be due to chance and may not represent a real difference in methods. Under the ranking method used by Grundin, small differences would be indistinguishable from larger, reliable differences.

Various methods of quantitative synthesis have been developed that might be more appropriate for analyzing the Bond and Dykstra (1967) and other comparisons of whole language/language experience approaches with other forms of beginning reading instruction. This paper will attempt such a quantitative review of the effectiveness of these approaches to beginning reading, using two methods of quantitative synthesis, vote counting and meta-analysis, to evaluate our current knowledge of the effectiveness of whole language and language experience approaches to beginning reading and to suggest future research directions.

The synthesis will be directed toward four questions: (a) What are the overall effects of whole language/language experience approaches compared with the basal reading approach predominant at the time of the study? (b) Are these effects different at kindergarten (or at the child's first exposure to formal schooling) than in first grade? (c) Do whole language/language experience approaches have a differential effect on different aspects of reading (word recognition, decoding, reading comprehension, etc.)? and (d) Have whole language/language experience approaches grown more effective over time?

Method

Study Selection

We used several data sources to locate relevant studies. A study was deemed relevant if it compared an approach using a language experience or whole language approach as the majority or entirety of a beginning reading program to a basal or traditional approach. For the purposes of this review, we defined a whole language/language experience method as having the following characteristics, as best as we could determine:

1. The emphasis in the program was on using children's own language, either through experience charts or through their own writing using invented spelling, as a medium of instruction.
2. The lessons were child centered rather than teacher centered. In a whole language/language experience classroom, it is assumed that children's competence will develop out of a need to use language to communicate better. In a basal classroom, skill sequences were used to sequence instruction.
3. An emphasis on trade books, rather than basals. This is not to say that basal reading programs were not used in any of the programs reviewed, but that they were used sparingly and were not the major emphasis.
4. Phonics lessons were not directly taught in isolation. Lessons in decoding were given as the need arose in the context of reading whole text.

Although this definition will probably not satisfy some (see Altwerger et al., 1987), it seemed to capture methods using the whole language philosophy as well as the earlier language experience studies. The studies reviewed as a whole represent a clear contrast with traditional basal reading programs. However, Slaughter (1988), who observed whole language and skills-oriented classes, found that whole language classes did include some direct skill instruction and skills classes contained some activities typical of whole language approaches. Her results suggest that the differences between whole language and other approaches may be a matter of emphasis, rather than mutually exclusive approaches.

We conducted computer searches of the ERIC and Dissertation Abstracts data bases using the descriptors "Language Experience" and "Whole Language." In addition, we checked references of obtained papers and bibliographies such as Crismore (1985), Hall (1977), and R. G. Stauffer (1976), we checked conference programs, and wrote to major figures in the field. We found 46 studies in addition to the USOE first grade studies. Only 15 of these provided enough information for the meta-analysis, resulting in 50 effect sizes for the non-USOE studies and an additional 71 effect sizes derived from the USOE first grade studies, leaving a total of 121 separate effect sizes.

Several studies, mostly in the USOE group, were long-term follow-ups of children whose initial reading instruction used whole language/language experience approach procedures. These were included in both analyses. Although it could be argued that children observed in third grade should not be included with those observed in first grade, our review both with and without these studies found that studies might have exaggerated the effects of whole language/language experience approaches slightly because of the inclusion of more of R. G. Stauffer's (1976) results. As discussed below, because his program produced the strongest effects of the various language experience methods, it may have made whole language/language experience approaches appear slightly more effective than they would otherwise.

Because the purpose of the review was to get the broadest picture of the specific approach. There were clear differences between the language experience approaches espoused by R. G. Stauffer (1969) and R. V. Allen (1976), and, as noted above, differences between these and whole language programs, such as those described by Goodman (1986), Newman (1985), and Weaver (1988). Because few studies used observations to verify fidelity to the intended method, such limitations

would be artificial. Some of the authors have used only language experience charts, others have used language experience in combination with other materials. Thus, included in our review are "pure" language experience approaches, language experience plus predictable books (Bridge et al., 1983; Ribowsky, 1985), language experience on computers (Educational Testing Service (ETS), 1984; Pickering & Pope, 1986), and so forth. Also, several of the studies used special populations including disadvantaged children (e.g., Brazziel & Terrell, 1962; Harris & Serwer, 1966), mildly retarded children (Woodcock, 1967), and so forth. When in doubt, we chose to include a study in order to avoid bias.

The review was limited, however, to instructional studies that compared the effects of a predominantly whole language or language experience program to a basal reading program. We did not include studies that compared students' reading of self-authored versus other authored materials (e.g., J. Allen, 1985) or studies that compared the number of words generated in experience stories to those used in basal texts (e.g., Gunderson & Shapiro, 1987), although these types of studies have also been used to support the efficacy of whole language/language experience approaches (see Hall, 1977). We also did not include studies that only examined the effects of tradebooks or predictable books (such as Bridge & Burton, 1982) or the effects of increased story reading on later reading achievement (Feitelson, Kita, & Goldstein, 1986; McCormick & Mason, 1986), although these also have been used to support whole language programs. Although Hall (1977) cites studies dating back to 1933, we limited our review to studies published after 1960 because of our uncertainty of the descriptions of the early whole language/language experience-prototype approaches.

Procedures

Two procedures were used to evaluate studies. First, a vote counting procedure was used (Light & Pillemer, 1984). Each result was classified as either significantly favoring a whole language/language experience approach, significantly favoring the basal reader approach, or nonsignificant. Measures included not only standardized achievement tests but also attitude measures, miscue analysis, concepts about print, and so forth. Second, where available, results were translated into effect sizes (Glass, McGaw, & Smith, 1981; Light & Pillemer), using the following formula:

$$ES = \frac{\text{Mean}_{\text{whole language/language experience}} - \text{Mean}_{\text{basal}}}{SD_{\text{basal}}}$$

Where the standard deviation of the basal group was not available, the pooled standard deviation for the entire group (or the best approximation available) was used instead. Calculations were made using effect sizes in order to answer the four questions previously posed. In this analysis, a positive effect size means that the whole language/language experience group outperformed the basal reader group; a negative effect size suggests the opposite.

Because the majority of the effect sizes were obtained from the USOE Cooperative First Grade Studies, we have decided to present these separately from the non-USOE studies. For all analyses where this is appropriate, effect sizes will be presented for the group of USOE studies, the non-USOE studies, and the total. This will be

done to avoid the problems of excessive reliance on one body of literature and to examine possible changes because the USOE studies were carried out more than 20 years ago.

Results

Overall Effects

In the vote counting, overall, whole language/language experience approaches appear approximately equal to basal reader approaches in their effectiveness. Of the non-USOE studies, 26 comparisons favored whole language/language experience approaches, 16 favored basal reader approaches, and 58 did not find significant differences. For the USOE studies, 17 comparisons favored the whole language/language experience method, 6 favored the basal reading method, and 57 were nonsignificant. These include observations after first, second, third, and sixth grade. Thus, out of 180 comparisons, 22% favored whole language/language experience programs, 12% favored basal reading programs, and 66% were nonsignificant. (Tables 1 and 2 report the studies used in the vote-counting analysis.)

If the two approaches were identical in their effects, conservatively one would expect about 5% of the differences to significantly favor one method and another 5% to favor the other by chance alone (Light & Pillemer, 1984). A chi-square test comparing the observed findings with this distribution found more significant differences than would have been expected by chance ($\chi^2 = 112.18$, $p < .0001$). Although the majority of the comparisons showed no differences between the two approaches, the number of significant differences suggests that this agglutination of findings may not represent one population, or that both whole language/language experience and basal reading programs may have different effects for different subsets of the studies.

The results of the meta-analysis confirms the general results of the vote counting. The mean of all 117 effect sizes was 0.09 ($sd = .61$). Effect sizes ranged from 1.91 to -1.46. A one-sample t test found that this was not significantly different from zero, suggesting that whole language/language experience approaches were not reliably different from basal reader approaches in their effects. (Tables 3 and 4 report the studies used in the meta-analysis.)

The results of both analyses suggest that the overall effects of whole language/language experience and basal reading programs are similar, but that the effects are not homogenous. Because of the heterogeneity of the findings, the overall lack of differences between whole language/language experience and basal reading programs may obscure strong differences in certain subsets of the data.

Readiness Versus Beginning Reading

When the vote-counting data were broken down by whether the whole language/language experience approach was used as a readiness program, preparatory for another beginning reading approach, or whether it was the beginning reading approach, an interesting pattern was found. For these readiness studies, 17 comparisons favor whole language/language experience approaches, 2 favor the basal, and 14 are nonsignificant, suggesting that whole language/language experience approaches are more effective in kindergarten. For first grade, the approaches are more equal. Of the non-USOE comparisons, 13 favor basal reader approaches, 43

TABLE 1
Vote counting, non-USOE studies

Study	Grade	Word recognition	Decoding	Oral reading	Comprehension	Readiness	Letter names	Print concepts	Vocabulary
				Readiness					
Brazziel & Terrell (1962)	[a]					LEA			
Bergemann (1969)	[a]			n.s.	n.s.	n.s.			
Bruckner, Morsillo, & Sample (1978)	K				LEA				LEA
ETS (1984)	K				LEA				
Guillemette (1979)	K	n.s. n.s.			n.s.				
Hall (1965)	[a]	LEA n.s. n.s.			n.s.	n.s.			
O'Donnell & Raymond (1972)	K	LEA				LEA	n.s.		
Phinney (cited in Weaver, 1988)	K					WL			
Pickering & Pope (1986)	K					LEA n.s.			
Ribowsky (1985)	K		WL				WL	WL	
Stewart (1986)	K							WL	
Taylor, Blum, & Logsdon (1986)	K					n.s.		WL	
Trachtenburg & Ferruggia (1988)	Transition class			WL		WL			WL
Walraven (1981)	K							BR BR LEA n.s.	

TABLE 1 (Continued)

Study	Grade	Word recognition	Decoding	Oral reading	Comprehension	Readiness	Letter names	Print concepts	Vocabulary
				Beginning reading					
Abt Associates (1977)									
Bank Street	3rd				n.s.				BR
TEEM	3rd				BR				BR
Asphlund & Sunal (1976)	2nd	n.s.							
Blachowitz et al. (1979)	1st		n.s.						
Bridge et al. (1983)	1st	LEA							
Carrigan (1986)	1st				BR				
Crandall (1973)	1st				n.s.				
Dittus (1983)	1st		n.s.		n.s.				
Duquette (1972)	1st	n.s.	n.s.	n.s.					n.s.
	2nd		LEA	n.s.					n.s.
ETS (1985)	1st				n.s.				
Evans & Carr (1985)	1st				BR				
					BR				
					n.s.				
					n.s.				
Ewoldt (1976)	3rd			n.s.	n.s.				
Farber & Putnam (1983)	1st				n.s.				
Fishman (1977)									
Experimental Program	1st				LEA				
Allen's Program	1st				n.s.				
Freeman & Freeman (1987)	1st	n.s.		n.s.					
Fryburg (1972)	1st		BR		BR				
Gallagher (1975)	1st	n.s.							
		n.s.							

Harris et al. (1967)	1st	n.s.	n.s.		n.s.		
Hoffman (1977)	1st				n.s.		
Lamb (1971)	1st				n.s.		
Phinney (cited in Weaver, 1988)	1st				n.s.		n.s.
Pollack & Brown (1980)	3rd			n.s.			
Powell et al. (1987)	1st				n.s.		
Ramig & Hall (1980)	1st			n.s.			
Sinatra (1984)	1st				WL		
					n.s.		
Stallings (1975)							
Bank Street	1st				BR		
TEEM	1st				BR		
M. A. Stauffer (1976)	1st	LEA	LEA	LEA	LEA		
Stice & Bertrand (1987)	1st			n.s.			
Stubbs (1983)	1st				n.s.		
	2nd				n.s.		
Swanson (1981)	1st				n.s.		
Tumner & Nesdale (1985)	1st		BR		BR		
			BR				
Woodcock (1967)	EMR	n.s.			n.s.	n.s.	n.s.
		n.s.			n.s.		
					n.s.		

Note. LEA is used for programs specifically labeled as language experience, WL for programs specifically labeled whole language, and BR is used for programs using the basal reading method.

[a] Subjects were first graders, but LEA was compared to a basal reading readiness program rather than an initial reading program. See text.

TABLE 2
Vote counting, USOE studies

Study	Stanford Achievement Tests				Gilmore Oral Reading		Word lists			Other	
	WR	Word study	Vocabulary	Paragraph meaning	Accuracy	Rate	Fry	Gates	Karlsen	NYS WR	NYS Comp
					Observed after first grade						
Cleland[a]	LEA	LEA	LEA	LEA	n.s.	n.s.	LEA	LEA			
Hahn	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.		
Kendrick	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s.	n.s.		
Stauffer	LEA	n.s.	n.s.	n.s.	LEA	n.s.	LEA	LEA	LEA		
Harris	BR	BR	n.s.	BR	n.s.			n.s.			
					Observed after second grade						
Pooled[b]	n.s.	n.s.	n.s.	n.s.	n.s.[c]	n.s.[c]	LEA[c]				
Pooled	n.s.	n.s.	n.s.	n.s.							
Pooled	n.s.	n.s.	n.s.	n.s.							
Harris et al. (1967)	n.s.			n.s.	BR	BR	BR	n.s.			
					Observed after third grade						
Harris et al. (1967)			n.s.	n.s.	n.s.	n.s.				n.s.	n.s.
Stauffer	n.s.	n.s.	n.s.	n.s.	LEA	LEA	LEA	LEA			
					Observed after sixth grade						
Stauffer (1976)			n.s.	n.s.	n.s.	LEA					

Notes. LEA = learning experience approach; BR = basal reading approach.

[a] The data in this table are derived from the following reports: Bond & Dykstra (1967), Cleland & Vilscek (1964), Dykstra (1968), Hahn (1966, 1967), Harris, & Serwer (1966), Harris, Serwer, & Gold (1967), Harris, Serwer, Gold, & Morrison (1967), Kendrick & Bennett (1966, 1967), R. G. Stauffer (1966, 1976), Stauffer & Hammond (1967, 1969), Stauffer, Hammond, Oehlkers, & Houseman (1976), Vilscek & Cleland (1968), and Vilscek, Morgan, & Cleland (1966). We have used the convention adopted by Bond and Dykstra to refer to each project by a last name of one of its directors.

[b] Dykstra (1968) did not report individual effect sizes for the Cleland, Kendrick, and Stauffer projects continued into second grade. Because we could not obtain reports of all three projects, we reported the pooled means reported by Dykstra. Each pooled mean was counted three times in the analysis because it represented three separate studies.

[c] From R. G. Stauffer (1966).

are nonsignificant, and 9 favor whole language/language experience approaches. The effects of whole language/language experience programs on readiness and first grade reading were significantly different ($\chi^2 = 14.45$, $p < .001$). Cramer's Phi, a measure of the strength of the association, was 0.39, suggesting a moderately strong relationship between the function of whole language/language experience instruction and its effectiveness. (Although the Bergemann, 1969, Brazziel & Terrell, 1962, and Hall, 1965, studies were conducted with first graders, this was the first exposure of these children to school. The use of language experience served the purpose of a readiness program for these children and their progress was partially measured using readiness tests. Therefore, we have chosen to include these studies with the studies using kindergarten children. We also classified Trachtenburg & Ferruggia's, 1988, study with prefirst transition class children in the readiness group because their program was intended to prepare children for a first-grade reading program. The overall results would be similar no matter which way these are classified.)

For the meta-analysis, we were able to derive effect sizes only from five studies in the readiness group. Because of the small number of studies involved and the use of different measures (print concept measures as opposed to reading performance measures), no similar comparison was made using the meta-analysis data.

Two of the studies we reviewed spanned both grade levels. Both ETS (1984) and Phinney (cited in Weaver, 1988, pp. 213–215) found whole language/language experience approaches to significantly improve children's reading skills in kindergarten but not first grade. Phinney's study was an informal evaluation of the implementation of a whole language approach, whereas the ETS study used computers equipped with speech synthesis to implement a language experience program. These implementations may not be typical. They do mirror the trends found in the overall analysis.

Differential Effects

The second question addressed in this analysis was whether whole language/language experience approaches had different effects on different measures of reading achievement. For the USOE studies, the mean effect size for word recognition measures was 0.17, whereas that for comprehension measures was 0.09. These are both small effects and essentially similar to each other. For the non-USOE studies, the mean effect size for word recognition was 0.33 and for comprehension measures was −.42. These effects are both moderate, favoring whole language/language experience programs on the word recognition measures and favoring basal readers on comprehension. The difference between them is large, and statistically significant [$t(28) = 4.69$, $p < .005$].

Older Versus Newer Studies

Another analysis examined whether whole language/language experience approaches were becoming more effective over time, as implied by Grundin (1985). To answer this, we calculated a Pearson correlation between the year of publication of the study and the obtained effect size. To avoid the biasing effects of having many effect sizes at the same year, we excluded the USOE studies from this analysis and used only the non-USOE effect sizes. The obtained correlation was −.14 ($N = 50$), which was not statistically significant ($p > .05$).

TABLE 3
Effect sizes derived from non-USOE studies

Study	Grade	Word recognition	Decoding	Oral reading	Comprehension	Print concepts	Vocabulary	Letter recognition	Readiness
				Readiness					
Bruckner et al. (1978)	K				0.70		0.62		
Guillemette (1979)	K	0.03			0.06				
		0.33							
Ribowsky (1985)	K		.55			1.91		.51	
Taylor, Blum, & Logsdon (1986)	K					.38			.33
Walraven (1981)	K					−0.27			
				Beginning reading					
Abt Associates (1977)									
Bank Street	3rd				−0.06		−0.28		
TEEM	3rd				−0.34		−0.32		
Asphlund & Sunal (1976)	2nd	0.78							
Bridge et al. (1983)	1st	1.09							

Carrigan (1986)	1st				−1.46	
Evans & Carr (1985)	1st				−1.09	
					−0.85	
					−1.26	
					−0.29	
Ewoldt (1976)	3rd			0.11	0.08	
Harris et al. (1967)	1st	0.13	−0.01		0.15	−0.11
	2nd	0.48			0.19	
		0.08			−1.12	
Gallagher (1975)	1st	0.02				
		0.03				
Lamb (1971)	1st				−0.49	
					−0.18	
Stallings (1975)						
Bank Street	1st				−0.45	
TEEM	1st				−0.86	
Tumner & Nesdale (1985)	1st		−1.17		−1.11	
			−0.91			
Woodcock (1967)	EMR	0.28			0.05	0.20
		0.38			0.33	
					0.33	

TABLE 4
Effect sizes derived from USOE studies

Study	Stanford Achievement Tests				Gilmore Oral Reading		Word lists			Other	
	WR	Word study	Vocabulary	Paragraph meaning	Accuracy	Rate	Fry	Gates	Karlsen	NYS WR	NYS Comp
					Observed after first grade						
Cleland[a]	0.50	0.36	0.45	0.41	0.07	0.13	0.24	0.36			
Hahn	0.25	0.13	0.16	0.13	−0.16	−0.08	0.12	−0.05	−0.03		
Kendrick	0.03	−0.10	−0.10	−0.26	0.39	−0.05	0.04	0.10	0.06		
Stauffer	0.33	0.03	0.14	0.21	0.64	−0.07		1.12	2.31		
Harris	−0.33	−0.34	−0.16	−0.43							
					Observed after second grade						
Pooled[b]	0.20	0.14	0.09	0.16	0.26[c]	0.03[c]					
Pooled	0.20	0.14	0.09	0.16							
Pooled	0.20	0.14	0.09	0.16							
Harris et al. (1967)	0.18			0.45	−0.82	−2.50	−0.64	−0.29			
					Observed after third grade						
Harris et al. (1967)			−0.09	−0.10	0.18	−0.05				−0.10	0.06
Stauffer		0.07	0.18	0.06	1.39	1.30	1.04	0.71			

[a] The data in this table are derived from the following reports: Bond & Dykstra (1967), Cleland & Vilscek (1964), Dykstra (1968), Hahn (1966, 1967), Harris & Serwer (1966), Harris, Serwer, & Gold (1967), Harris, Serwer, Gold, & Morrison (1967), Kendrick & Bennett (1966, 1967), R. G. Stauffer (1966, 1976), Stauffer & Hammond (1967, 1969), Stauffer, Hammond, Oehlkers, & Houseman (1976), Vilscek & Cleland (1968), and Vilscek, Morgan, & Cleland (1966). We have used the convention adopted by Bond and Dykstra to refer to each project by a last name of one of its directors.

[b] Dykstra (1968) did not report individual effect sizes for the Cleland, Kendrick, and Stauffer projects continued into second grade. Because we could not obtain reports of all three projects, we reported the pooled means reported by Dykstra. Each pooled mean was counted three times in the analysis because it represented three separate studies.

[c] From R. G. Stauffer (1966).

Because the earlier analysis suggested that whole language/language experience programs function differently as a readiness program than they do as a beginning reading program, separate correlations were calculated for each data set. For the readiness studies, the relationship between year of publication and effect size was moderate and positive (0.22) but was not significantly different from zero probably because of the small number of effect sizes ($N = 14$) used in the analysis. For the beginning reading studies, the effect was also moderate, $-.58$ ($N = 36$), which was statistically significant ($p < .001$). Because this correlation is negative, it suggests that between 1967 and 1986 there was a tendency for the whole language/language experience approaches used as beginning reading methods to be less effective in relation to basal reading approaches with later years of publication. This could be interpreted to mean that, as language experience approaches have evolved into whole language approaches, they have been associated with higher relative achievement in kindergarten and lower relative achievement in first grade.

Related were differences found between the USOE and non-USOE studies. The mean effect size derived from the USOE studies was small and positive (0.14). The mean effect size for the non-USOE studies, was near zero (.01). The average effect sizes for the USOE and non-USOE groups were significantly different from each other [$t(115) = 4.68$, $p < .001$], although both were small and neither was significantly different from zero. Therefore, the language experience treatments used in the USOE studies were more effective relative to the basals used at the time than that used in the non-USOE studies, confirming the trend for later studies to show lower effects for these approaches.

Disadvantaged and Lower SES Populations

It has been suggested by some (see Hall, 1972) that whole language/language experience approaches are especially suitable for disadvantaged or lower SES populations. To examine this, we looked separately at studies specifically examining these populations. We found 12 studies that specifically included those terms in the subject descriptions or were part of programs specifically set up for such populations, such as Project Follow Through. (We could not determine the SES of about a third of the studies we examined from the descriptions we had.) Of the three readiness studies who examined a lower SES population (Brazziel & Terrell, 1962; Hall, 1972; O'Donnell & Raymond, 1972) five comparisons favored whole language/language experience approaches, while five were nonsignificant. This is similar to that found in the overall analysis.

In the beginning reading group, of the nine studies that specifically examined lower SES groups (Abt Associates, 1977; Ewoldt, 1976; Fryburg, 1972; Gallagher, 1975; Harris et al., 1967; Hoffman, 1977; Lamb, 1971; Powell, Needham, & Cochran, 1987; Stallings, 1975), 9 comparisons favored the basal reader approach and 15 were nonsignificant. None of the comparisons made with specifically lower SES populations at this level favored whole language/language experience approaches. In the overall analysis, the results were more equal. A contingency table analysis found significant differences between those beginning reading studies using specifically lower SES populations and the remainder of the studies ($\chi^2(2) = 17.78$, $p < .01$). Any conclusion should be tempered by the number of studies for which SES information was not available. It does appear, however, that whole language/language experience approaches do not have a particular advantage with lower SES

populations. In fact, they may have less of an effect with this population than with more homogenous or middle or upper class populations.

Standardized Versus Naturalistic Measures

Some authors have suggested that children in basal reading programs are exposed to more testlike events in their instruction, biasing these measures toward programs with basic skills orientations (e.g., Harste, 1985; House, Glass, McLean, & Walker, 1978). It has been suggested, then, that more naturalistic measures should be used to measure more naturalistic programs. Such measures, such as oral reading miscue analysis and attitude measures, were included mostly in the vote-counting analysis because only one of these studies provided numerical data necessary to derive effect sizes.

The results from more naturalistic measures mirror those from the other measures. Four studies (Ewoldt, 1976; Pollack & Brown, 1980; Ramig & Hall, 1980; Stice & Bertrand, 1987) found no significant differences in the number of oral reading miscues produced by whole language/language experience trained and basal trained first graders (see also Blachowitz, McCarthy, & Ogle, 1979; Wilkinson & Brown, 1983). Ewoldt also found no differences between the groups on a qualitatively scored measure of retelling. On another naturalistic measure of comprehension—the number of predictions made after reading an open ended story—Farber and Putnam (1983) also failed to find significant differences between language experience trained first graders and basal trained first graders. Harris et al. (1967) found that language experience trained students read more books than basal students when number of different books was the unit of analysis, but basal students read more pages in free reading.

The results of measures of children's attitudes toward reading are equally ambivalent. Some proponents of whole language/language experience approaches suggest that these approaches improve students' attitudes toward reading, whereas basal approaches deaden enthusiasm (e.g., Harste, 1985). The results of the vote counting found that whole language/language experience approaches had significant effects on attitude measures in 3 studies, whereas no significant difference was found in 11 studies. One difference favored the basal reader group.

The results of the naturalistic measures, therefore, mirror those derived from standardized tests. On both types of instruments, whole language/language experience approaches produce approximately the same levels of achievement and attitude toward reading.

Study Quality

Meta-analysis has been criticized by some (e.g., Slavin, 1986) for including all available studies, irrespective of quality. Slavin suggests a *best evidence* synthesis in excluding effect sizes from studies that did not meet rigorous criteria. Such a synthesis has its own problems because the criteria are necessarily post hoc, in that they are applied after the data is collected, no matter how logical they might seem. To examine study quality, we have chosen to present how each exclusionary criterion would affect the overall results.

Through an examination of the corpus of studies, we determined six criteria that characterized the best qualities of research in this area. They are as follows:

Treatment and control groups both are taught by more than one teacher. If only one teacher is assigned per group, then treatment and teacher factors are confounded.

Both programs should be at least in their second year of implementation. This is to insure both that teachers are comfortable using the programs and to dilute the high expectations that accompany any new program.

Treatment should last at least 6 months. We felt a better test of the effectiveness of a program was in how it fared over a relatively long term. Most of the studies were a full year in length.

Study should include observations in order to insure fidelity of treatment. Studies such as those of Chall and Feldmann (1966) suggest that how teachers identify their program may have little relation to what they are doing in the classroom. Studies in which a researcher observed instruction in the classroom were likely to have greater fidelity to the intended treatment.

Initial differences should be accounted for. This would usually involve covarying initial differences in readiness test scores or intelligence, but might also include initial differences on criterion measures in a pretest-posttest gain score design.

Study should use traditional first graders. We excluded here studies that used as subjects Sikhs for whom English was a second language (Carrigan, 1986), educable mentally retarded children (Woodcock, 1966), and learning disabled children (Guillemette, 1979), but included those using disadvantaged populations (Ewoldt, 1976; Gallagher, 1975; Harris et al., 1967; Lamb, 1971).

In addition, we included scores only from measures of reading, word recognition, oral reading, decoding, comprehension, or silent vocabulary given in kindergarten or first grade, and excluded measures of prereading skills or measures given as long-term follow-ups so that the studies would be more homogenous.

Table 5 lists the exclusionary criteria, the studies each criterion includes, and the resulting effect size. Only two studies of the non-USOE group met all six criteria: Harris et al. (1967) and Stallings (1975). The mean effect size from these studies was $-.19$, a small effect favoring the basal reading program. This would be the estimate from a best evidence synthesis. (The USOE studies, taken as a whole, met all of the above criteria except the second.) When the effect of applying each criterion was examined singly, for five of the six criteria, the resulting effect sizes were markedly lower than the overall effect size of 0.09 found for all studies. In addition, the correlation between the number of criteria each study met and its resulting effect size was $-.41$ ($p < .05$), also suggesting that as the study quality increased, the effect size tended to decrease or favor basal reading programs. This analysis thus suggests that the results of the overall meta-analysis might be an overestimation of the whole language/language experience methods' effects.

Outliers

Because massing of effect sizes can mask individual effective programs, separate descriptions will be made of outlier studies. An outlier study is one producing an effect size of ± 1 standard unit from the mean of all studies. Examination of outliers allows one to test the trends found in the overall analysis, and to look at the characteristics of both very effective and very ineffective approaches.

The four outlier studies that showed strong effects for whole language/language experience showed similar trends. These studies found at least one effect size greater

TABLE 5
Effects of exclusionary criteria on meta-analysis

Criterion	Studies meeting criterion	Effect size
More than one teacher	Carrigan (1986), Evans & Carr (1985), Harris et al. (1967), Lamb (1971), Stallings (1975), Tumner & Nesdale (1985), Woodcock (1967).	−.36
Second year or later	Carrigan, Evans & Carr, Stallings, Tumner & Nesdale.	−.44
At least 6 months	Ribowsky (1985), Carrigan, Evans & Carr, Lamb, Stallings, Tumner & Nesdale, Woodcock.	−.30
Observations for fidelity	Ribowsky, Bridge et al. (1983), Carrigan, Evans & Carr, Harris et al., Stallings.	−.34
Initial differences accounted for	Ribowsky, Asphlund & Sunal (1976), Bridge et al., Carrigan, Evans & Carr, Gallagher (1975), Lamb, Stallings, Woodcock.	.08
Conventional students	Ribowsky, Asphlund & Sunal, Bridge et al., Evans & Carr, Harris et al., Gallagher, Lamb, Stallings, Tumner & Nesdale.	−.30

than +1 standard unit, or that the language experience or whole language group outperformed the basal control by a full standard deviation. Of these, three (Bridge et al., 1983; Ribowsky, 1985; Walraven, 1981) were concerned with the initial stages of reading, either in the beginning of first grade (Bridge et al.) or kindergarten (Ribowsky; Walraven). These support the hypothesis that whole language programs are effective as an initial introduction to literacy.

The only other outlier showing a strong advantage for a whole language/language experience approach was the third grade follow-up of Stauffer's (Stauffer & Hammond, 1969) USOE study. Unlike many of the USOE studies, Stauffer's second and third grade instruction was also especially adapted rather than conventional instruction. This large effect in third grade may be the culmination of Stauffer's overall curriculum, which included specially developed teaching techniques at the second and third grade levels as well rather than just the effects of a language experience approach. Stauffer's version of language experience, however, also produced consistently larger effects on first grade measures than the versions used in other USOE projects. One possible reason for this will be discussed later.

Four studies found average effect sizes for whole language/language experience methods of −1.00 or less, or that an average student in the whole language/language experience group scored a full standard deviation below the average student in the basal group. All of these studies, Carrigan (1986), Evans and Carr (1985), Harris et al. (1967), and Tumner and Nesdale (1985), were evaluations of existing programs rather than deliberate manipulations in which an experimenter trained teachers to use a whole language/language experience approach and then evaluated the results of that instruction. The two other evaluation studies, the Follow Through evaluations by Abt Associates (1977) and Stallings (1975), also found basal reader

programs to be superior to language experience approaches. (They also both found code emphasis approaches to be more effective than either a language experience approach or a basal approach.)

The results of a deliberate manipulation may be subject to an experimenter effect (Campbell & Stanley, 1966). In other words, the training or experimenter expectations may have biased the results in favor of the manipulated treatment when compared with a status quo, basal treatment. In some studies, the experimenter taught some of the experimental classes but not the basal control classes (e.g., Fishman, 1977; Ribowsky, 1985). Even when others do the teaching, they must not only be specially trained to be an effective whole language/language experience teacher, but also must invest time, effort, and ingenuity preparing materials for their classes. Thus, the superiority of basal reader instruction on these evaluation studies suggests that the long-term effects of whole language/language experience might be lower than this meta-analysis found.

Discussion

The results of the meta-analysis and vote-counting procedures appear to suggest that, overall, whole language/language experience approaches were approximately equal to basal reading approaches in their effects. This was found on both standardized and nonstandardized measures and on measures of both attitude and achievement. These results, however, were not consistent across all studies. Whole language/language experience approaches appear to be more effective when used in kindergarten or when used instead of a reading readiness program and seem to have had greater effects on measures of word recognition than on measures of comprehension in the more recent non-USOE studies. In addition, the studies that met more of our rigorous criteria for inclusion tended to favor basal reading programs over whole language/language experience programs. An analysis of outliers also confirms these general trends.

The finding that whole language/language experience programs appear more effective when used prior to a formal reading program fits well into the stage models of reading acquisition discussed earlier. Whole language approaches may approximate the kind of incidental learning about books that takes place in middle-class households (see Snow, 1983; Snow & Ninio, 1986) in which children are initiated to print through dialogue about its features. This learning may not take place in some households, especially where books are not present or that literacy has a lower status (Chall & Snow, 1982). Although these print concepts are important, once they are mastered children may be ready for more systematic learning about the code of written language (see Chall, 1983b).

In the mastering phase, a more systematic approach to decoding than whole language approaches provide may be needed, at least for some children. It is possible to compare more systematic programs, such as those suggested by Adams (in press), Anderson et al. (1985) and Chall (1983a) as most effective in first grade reading programs to language experience programs, at least in a limited manner. Bond and Dykstra (1967) included programs labeled "Phonic/Linguistic," which correspond most closely to Anderson et al.'s and Chall's (1983a) suggestions, in their analysis. Looking only at the results at the end of first grade, the three phonic/linguistic projects (Hayes, Tanyzer, and Wyatt) produced a mean effect size of 0.91

on the Stanford Word Reading subtest and 0.36 on the Stanford Paragraph Meaning subtest. The corresponding effects of the five whole language/language experience projects are 0.16 and 0.01. Thus, when compared with similar basal reading programs, phonic/linguistic programs used in first grade produced strikingly larger effects than language experience programs, suggesting that, in this mastering phase, the more systematic code-emphasis approaches were the more successful.

Many authors have argued that using a whole language or language experience approach precludes (or is an alternative to) teaching decoding systematically. This does not have to be the case. R. G. Stauffer's (1969) approach places the largest emphasis on word recognition in isolation, making extensive use of word banks and explicit teaching to bring word forms to the child's attention during reading of experience charts. Stauffer's program also produced larger effects in first grade than the other language experience programs in the USOE studies. This greater attention to word forms may have contributed to that success. Fishman (1977), similarly, integrated systematic attention to decoding into a language experience approach. This approach produced significantly higher reading comprehension achievement than either a language experience method based on R. V. Allen's (1976) work or a basal reader control.

Some authors have argued that students learn about the code through whole language approaches, through exposure to print and through invented spelling (e.g., Goodman, 1986; Weaver, 1988). They argue that such an approach is preferable, that because it is integrated into the context of reading, students remain focused on comprehending whole texts. There is some evidence that children do learn decoding skills through such approaches (Gunderson & Shapiro, 1987), but such instruction is unavoidably less efficient for the specific purpose of learning sound-symbol relations than direct instruction. Less efficient approaches require more time for learning. Because lower SES and middle and upper SES students vary considerably in the amount of listening to storybooks and other conventional literacy directed behaviors in the home (see Adams, in press, for review), this might explain why whole language/language experience approaches might be more effective as beginning reading programs with middle and upper SES populations. Such populations probably have learned more about the code through long-term exposure to print and the interactions around storybooks (see Snow, 1983). In other words, they have had more time to learn. Children who have not had as much exposure and the same types of interactions need direct instruction to catch up (see Chall, 1983a). Delpit (1988) suggests that more progressive approaches may simply give children the opportunity to show what they have learned, but systematic approaches teach letter sounds and other basic literacy concepts as new information children have not learned.

At least for some of the approaches studied, the effects found for whole language/language experience programs may have been related to the amount of reading done by the students. Harris and Serwer's (1966) formal observations of language experience and basal programs found that the children in their project's basal reading program spent more time on direct reading activities, such as reading connected text. Language experience pupils spent more time in indirect reading activities, such as talking about what they were going to write, talking about what they read, and so forth. They further found that the amount of time spent on direct reading activities was positively correlated with reading achievement while time

spent on indirect activities tended to be negatively correlated. This second finding has been replicated by other researchers (e.g., Berliner, 1981).

It certainly would not be the intent of the proponents of whole language/language experience approaches that children spend less time reading. However, the philosophy that reading and speaking are equivalent aspects of a same general process seems to suggest that time spent on language, oral or written, would be equally profitable. If the goal is to improve reading skill, it does not appear to be the case. Evans and Carr (1985) found that gains in spoken language skills were negatively correlated to gains in reading skill, that is, the greater gains in oral language were associated with lower reading skills. Their interpretation was that children need *print-specific skills* for success in reading, and that children cannot benefit from their language skills unless they have a requisite amount of these print-specific skills. Therefore, spending more time on oral language development might not show an effect until children are able to identify a certain number of written words (see also Chall, 1983b). Curtis (1980), for example, found word recognition skill but not language comprehension skill to be related to reading comprehension in second grade, whereas both word recognition and language comprehension related strongly to reading comprehension in fifth grade.

Integrating Perspectives

Some authors (DeFord, 1985; Harste, 1985) have attempted to trichotomize the practice of reading, separating practices into either phonics approaches, skills approaches, or whole language approaches. In this scheme, whole language approaches involve increased use of quality children's literature, writing, and insuring that all skills are applied in the context of reading, rather than treated as isolated exercises. These are virtues that should be a part of any reading program. It would be wrong to interpret the indifferent effects found for whole language/language experience programs at the first grade level as supporting the banality of some basal reading stories, the excessive use of worksheets, or other aspects of the basal reading approach.

One goal of any reading instruction should be that children become efficient at constructing meaning from text. However, approaches that always emphasize the construction of meaning may not be the most effective in achieving their purpose in the end. It appears clear from the reviews of Adams (in press), Anderson et al. (1985), Chall (1983a, 1983b), and others, that children need to go through intermediate stages of mastering word recognition abilities to better develop the reading abilities necessary to read good quality literature with enjoyment and understanding. These intermediate stages appear to be better served with direct and systematic phonics instruction. On the other hand, it appears that even the most systematic phonics approaches work better as children get increased opportunities to read children's literature (Adams).

The research base seems to point to an amalgam of the three approaches. The trichomization has implied, however, that an acceptance of the virtues of one approach requires rejection of the other approaches. Quality phonics instruction need not be synonymous with excessive worksheets, nor must it exclude the use of quality literature. We have observed first grade teachers who integrate direct instruction of phonics with a broad program using children's literature and individual writing. We have also observed teachers whose classes largely do round robin

reading of the reader and worksheets on both phonics and comprehension skills. The first type of teacher seems ideal, and all too rare. The second type may create the impression that reading is drudgery to be avoided, and is all too common.

Limitations and Suggestions for Further Research

It is important to note some limitations of this review. First, we did not examine the effects of whole language/language experience approaches on writing. Individualized writing, possibly using invented spelling, could have positive effects on children's development of writing skills, irrespective of its effect on reading. Second, we limited our review to programs in the kindergarten or first grade year. We do not intend for these results to apply to later grades, where the effects of whole language approaches might be quite different. The stage model of Chall (1983b) might suggest that increased amounts of more challenging reading and writing in reaction to that reading might have salutary effects in the later grades, after decoding becomes automatic and reading requires more information and reasoning skills. There is not enough research available, however, to evaluate the effects of whole language programs at these levels.

Third, because we limited our review to programs that predominantly used whole language/language experience approaches, these results may not apply when aspects of whole language/language experience approaches are used as supplements to other reading programs. Teachers report that language experience is often used in this way, but we have not found enough research examining the effectiveness of these approaches used in this role to draw conclusions. Swanson (1981) found nonsignificant correlations between the amount of language experience activities used in the first grade classes she observed and reading achievement. The applicability of these results are limited, however, by the generally small amounts of language experience instruction she observed.

Last, many researchers interested in whole language approaches have used ethnographic methods. Typically, studies using these methods have painted an appealing picture of the whole language classroom, but, of course, studies of this type do not license conclusions about whether the approach is more or less effective than other approaches to teaching beginning reading.

The use of ethnographic research, however, may be premature. Such research is useful to get a participant's view of an effective instructional setting, or to contrast an effective setting with an ineffective one. Given the results reported here, one cannot establish that whole language classes are more effective than the status quo. Rather than more "Method A versus Method B" research comparing multicomponent packages, we suggest that future research be directed toward isolating effective components of beginning reading programs, regardless of philosophy. Evans, and Carr (1985), Harris and Serwer (1966), and Stallings (1975) combined observations of the activities in whole language and basal oriented classrooms with overall evaluations of the effectiveness of the different approaches. They, however, used only general categories to characterize the activities they observed. Evans and Carr, for example, found that the amount of silent reading in both settings correlated significantly with overall achievement, whereas the amount of oral reading and word analysis activities did not. Without more information about the nature of these activities, these results are difficult to interpret. Does, for example, silent reading in trade books with "authentic" language have different effects than reading

"basalese"? What are the effects of invented spelling specifically on decoding skill? Are they more effective than practice on worksheets? Such questions can best be answered through more fine-grained comparisons between classes. Ultimately, such comparisons are the only way to examine the contributions of the whole language movement to beginning reading instruction.

From the data reviewed, it appears that whole language approaches may have an important function early in the process of learning to read, but that as the child's needs shift, they become less effective. It could also be that the philosophy behind whole language/language experience approaches, that the function of reading is to communicate, needs to be learned by children early, but, once learned, children need to be able to decode written language fluently and automatically in order to be able to use reading for that purpose.

References

Abt Associates. (1977). *Education as experimentation: A planned variation model: Vol. IV-B. Effects of Follow Through Models.* Cambridge, MA: Author.

Adams, M. J. (in press). *Phonics and beginning reading instruction.* Champaign, IL: Reading Research and Education Center, University of Illinois.

Allen, J. (1985) Inferential comprehension: The effects of text source, decoding ability, and mode. *Reading Research Quarterly, 20,* 603–615.

Allen, R. V. (1976) *Language experiences in communication.* Boston: Houghton Mifflin.

Altwerger, B., Edelsky, C., & Flores, B. M. (1987). Whole language: What's new? *The Reading Teacher, 41,* 144–154.

Anderson, R. C., Hiebert, E., Scott, J., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers.* Champaign, IL: Center for the Study of Reading and National Academy of Education.

Asphlund, B. B., & Sunal, C. S. (1976). *The effectiveness of the language experience approach as a supplement to a basal reader program.* (ERIC Document Reproduction Service No. ED 182 705)

Bergemann, V. E. (1969). *A modified language experience approach versus a basal approach to reading with first grade children.* Unpublished doctoral dissertation, University of Maryland, College Park.

Berliner, D. C. (1981). Academic learning time. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 203–226). Newark, DE: International Reading Association.

Blachowitz, C. L. F., McCarthy, L. B., & Ogle, D. M. (1979). Testing phonics: A look at children's response biases. *Illinois School Research and Development, 16,* 1–6.

Bond, G. L., & Dykstra, R. (1967). The cooperative research program in first grade reading instruction. *Reading Research Quarterly, 2*(4), 5–142.

Brazziel, W. F., & Terrell, M. (1962). For first graders: A good start in school. *The Elementary School Journal, 62,* 352–355.

Bridge, C. A., & Burton, B. (1982) Teaching sight vocabulary through patterned language materials. In J. A. Niles & L. A. Harris (Eds.), *New inquiries in reading research and instruction: Thirty-first yearbook of the National Reading Conference* (pp. 119–123). Washington, DC: National Reading Conference.

Bridge, C. A., Winograd, P. N., & Haley, D. (1983). Using predictable materials vs. preprimers to teach beginning sight words. *The Reading Teacher, 36,* 884–891.

Bruckner, J. H., Morsillo, C. M., & Sample, E. E. (1978). Supportive evidence for the language experience approach at the kindergarten level. *Graduate Student Association Journal, 1,* 15–29. (ERIC Document Reproduction Service No. ED 157 635)

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Carrigan, A. (1986, December). *Reading achievement of grade one students involved in language experience programs vs. basal reader programs.* Paper presented at the annual meeting, National Reading Conference, Austin, TX.

Chall, J. S. (1983a). *Learning to read: The great debate* (2nd ed.). New York: McGraw-Hill.

Chall, J. S. (1983b). *Stages of reading development.* New York: McGraw-Hill.

Chall, J. S., & Feldmann, S. (1966). First grade reading: An analysis of the interactions of professed methods, teacher implementation, and child background. *The Reading Teacher, 19,* 569–575.

Chall, J. S., & Snow, C. E. (1982). *Families and literacy: A report to the National Institute of Education.* Cambridge, MA: Harvard Graduate School of Education.

Clay, M. (1985) *The early detection of reading disabilities* (3rd ed.). Portsmouth, NH: Heinemann.

Cleland, D. L., & Vilscek, E. (1964). *Comparison of the basal and the coordinated language experience approaches in first grade reading instruction.* (ERIC Document Reproduction Service No. ED 012 687)

Crandall, A. H. (1973). *A comparison of reading attitude and reading achievement among first graders in open concept and more formal classes.* Unpublished doctoral dissertation, University of Connecticut, Storrs.

Crismore, A. (1985). *Landscapes: A state-of-the-art assessment of reading comprehension research 1974–1984.* Bloomington, IN: Language Education Department, Indiana University.

Curtis, M. E. (1980). Development of components of reading skill. *Journal of Educational Psychology, 72,* 656–669.

DeFord, D. (1985). Validating the construct of theoretical orientation in reading instruction. *Reading Research Quarterly, 20,* 351–367.

Delpit, L. D. (1988). The silenced dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review, 58,* 280–298.

Dittus, C. H. (1983). *The influence of instructional approach, instruction/test overlap, and student behavior on reading achievement.* Unpublished doctoral dissertation, Rutgers, The State University of New Jersey.

Downing, J. (1979). *Reading and reasoning.* New York: Springer-Verlag.

Duquette, R. J. (1972). Barnette-Duquette study. *Childhood Education, 38,* 437–439.

Dykstra, R. (1968). Summary of the second-grade phase of the cooperative research program. *Reading Research Quarterly, 4,* 49–70.

Educational Testing Service. (1984). *The ETS evaluation of Writing to Read: Executive summary.* Princeton, NJ: Author.

Evans, M. A., & Carr, T. H. (1985). Cognitive abilities, conditions of learning, and the early development of reading skill. *Reading Research Quarterly, 20,* 327–349.

Ewoldt, C. (1976). *Miscue analysis of the reading of third grade Follow Through and non-Follow Through children in Wichita, Kansas.* Tuscon: Arizona Center for Educational Research and Development. (ERIC Document Reproduction Service No. ED 136 219)

Farber, F. D., & Putnam, L. P. (1983, October). *Convergent/divergent predictions of urban first graders.* Paper presented at the annual meeting of the College Reading Association, Atlanta, GA (ERIC Document Reproduction Service No. ED 236 564)

Feitelson, D., Kita, B., & Goldstein, Z. (1986). Effects of listening to series stories on first graders' comprehension and use of language. *Research in the Teaching of English, 20,* 339–356.

Fishman, T. G. (1977). *A comparison of two different programs for implementing the language experience approach to beginning reading with a basic reader program and phonic supplement.* Unpublished doctoral dissertation, Boston University.

Freeman, R. H., & Freeman, G. G. (1987). Reading acquisition: A comparison of four approaches to reading instruction. *Reading Psychology, 8,* 257–272.

Fryburg, E. L. (1972). *The relations among English syntax, methods of instruction, and*

reading achievement of first grade disadvantaged black children. Unpublished doctoral dissertation, New York University.

Gallagher, J. A. M. (1975). *Vocabulary retention of lower-class students in language experience and in basal text approaches to the teaching of reading.* Unpublished doctoral dissertation, University of South Carolina.

Glass, G. V, McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Goodman, K. S. (1986). *What's whole in whole language.* Portsmouth, NH; Heinemann.

Goodman, Y., & Goodman, K. S. (1979). Learning to read is natural. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading* (Vol. 1, pp. 137–154). Hillsdale, NJ: Erlbaum.

Grundin, H. (1985). A commission of selective readers: A critique of Becoming a Nation of Readers. *The Reading Teacher, 39,* 262–266.

Guillemette, M. (1979). *A study of the effectiveness of sight reading versus phonetic instruction for children with auditory learning disability.* (ERIC Document Reproduction Service No. ED 184 074)

Gunderson, L., & Shapiro, J. (1987). Some findings on whole language instruction. *Reading-Canada-Lecture, 5,* 22–26.

Hahn, H. T. (1966). Three approaches to beginning reading instruction—i.t.a., language arts, and basic readers. *The Reading Teacher, 19,* 590–594.

Hahn, H. T. (1967). Three approaches to beginning reading instruction—i.t.a., language arts, and basic readers—Extended into second grade. *The Reading Teacher, 20,* 711–715.

Hall, M. A. (1965). *The development and evaluation of a language experience approach to reading with first-grade culturally disadvantaged children.* Unpublished doctoral dissertation, University of Maryland.

Hall, M. A. (1972). *The language experience approach for the culturally disadvantaged.* Newark, DE: International Reading Association.

Hall, M. A. (1977). *The language experience approach for teaching reading: A research perspective.* Newark, DE: International Reading Association.

Harris, A. J., & Serwer, B. L. (1966). The CRAFT project: Instructional time in reading research. *Reading Research Quarterly, 2*(1), 27–56.

Harris, A. J., Serwer, B. L., & Gold, L. (1967). Comparing reading approaches in first grade teaching with disadvantaged children—extended into second grade. *The Reading Teacher, 20,* 698–703.

Harris, A. J., Serwer, B. L., Gold, L., & Morrison, C. (1967). *A third progress report on the CRAFT project: Teaching reading to disadvantaged primary grade urban Negro children.* (ERIC Document Reproduction Service No. ED 015 841)

Harste, J. (1985). Becoming a nation of readers: Beyond risk. In J. Harste (Ed.), *Toward practical theory: A state of practice assessment of reading comprehension instruction.* Bloomington, IN: Language Education Department, Indiana University.

Hildreth, G. H. (1965). Experience related reading for school beginners. *Elementary English, 42,* 280–297.

Hoffman, J. R. (1977). *An assessment of the language experience approach to beginning reading and additional teacher characteristics in eight summer migrant schools.* Unpublished doctoral dissertation, University of Colorado, Boulder.

House, E. R., Glass, G. V, McLean, L. D., & Walker, D. F. (1978). No simple answer: Critique of the Follow Through evaluation. *Harvard Educational Review, 48,* 128–160.

Kendrick, W. M., & Bennett, C. L. (1966). A comparative study of two first grade language arts programs. *Reading Research Quarterly, 2*(1), 83–118.

Kendrick, W. M., & Bennett, C. L. (1967). A comparative study of two first grade language arts programs—extended into second grade. *The Reading Teacher, 20,* 747–755.

Knight, J. J. (1971). *A comparison of four different reading programs on children's expressed attitudes toward reading.* Unpublished doctoral dissertation, University of New Mexico.

Lamb, P. (1971). The language experience approach to teaching beginning reading to culturally disadvantaged pupils. (ERIC Document Reproduction Service No. ED 059 314)

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research.* Cambridge, MA: Harvard University Press.

McCormick, C. E., & Mason, J. M. (1986). Intervention procedures for increasing children's interest in and knowledge about reading. In E. Sulzby & W. Teale (Eds.), *Emergent Literacy* (pp. 90–114). Norwood, NJ: Ablex.

Newman, J. M. (1985). *Whole language: Theory and use.* Portsmouth, NH: Heinemann.

O'Donnell, C. M., & Raymond, D. (1972). Developing reading readiness in the kindergarten. *Elementary English, 49,* 768–771.

Pickering, T., & Pope, S. (1986, October). *A study of computer assisted instruction vs. a traditional basal reader approach in kindergarten.* Paper presented at the annual meeting of the Mid-Western Educational Research Association, Chicago, IL.

Pollack, J. F., & Brown, G. H. (1980, August). *Observing the effects of reading instruction.* Paper presented at the annual meeting of the World Congress on Reading, Manilla, The Phillipines. (ERIC Document Reproduction Service No. ED 195 500)

Powell, D., Needham, R., & Cochran, K. (1987, December). *A comparison of Hispanic/Anglo, middle/low income, first and second graders' reading achievement: A pilot study.* Paper presented at the annual meeting of the National Reading Conference, St. Petersburg Beach, FL.

Ramig, C. J., & Hall, M. A. (1980). Reading strategies of first grade children taught by a language experience and a basal approach. *Reading World, 19,* 280–289.

Ribowsky, H. (1985). The effects of a code emphasis and a whole language approach upon the emergent literacy of kindergarten children. (ERIC Document Reproduction Service No. ED 269 720)

Samuels, S. J. (1985). Automaticity and repeated reading. In J. Osborn, P. T. Wilson, & R. C. Anderson (Eds.), *Reading education: Foundations for a literate America* (pp. 215–231). Lexington, MA: Lexington Books.

Simons, H. D., & Ammon, P. (1988). Primerese miscues. In J. E. Readence & R. S. Baldwin (Eds.), *Dialogues in literacy, 37th yearbook of the National Reading Conference* (pp. 115–121). Chicago, IL: National Reading Conference.

Sinatra, R. (1984, October). *How a holistic language arts curriculum influenced the reading and writing proficiency of kindergarten and primary grade children.* Paper presented at the annual meeting of the Eastern Regional Conference of the International Reading Association. (ERIC Document Reproduction Service No. ED 264 583)

Slaughter, H. B. (1988). Indirect and direct teaching in a whole language program. *The Reading Teacher, 42,* 30–34.

Slavin, R. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher, 15*(9), 5–11.

Snow, C. E. (1983). Language and literacy: Relationships during the preschool years. *Harvard Educational Review, 53,* 165–189.

Snow, C. E., & Ninio, A. (1986). The contracts of literacy: What children learn from learning to read books. In W. H. Teale & E. Sulzby (Eds.), *Emergent literacy: Writing and reading* (pp. 116–138). Norwood, NJ: Ablex.

Stallings, J. (1975). Implementation and child effects of teaching practices in follow through classrooms. *Monographs of the Society for Research in Child Development, 40*(7–8, Serial No. 163).

Stauffer, M. A. (1976). Comparative effects of language arts approach and basal reader approach to first grade reading achievement. In R. G. Stauffer (Ed.), *Action research in L.E.A. instructional procedures* (pp. 205–226). Newark, DE: University of Delaware.

Stauffer, R. G. (1966). The effectiveness of language arts and basic reader approaches to first grade reading instruction. *The Reading Teacher, 20,* 18–24.

Stauffer, R. G. (1969). *Teaching reading as a thinking process.* New York: Harper & Row.

Stauffer, R. G. (1976). *Action research in L.E.A. instructional procedures.* Newark, DE: University of Delaware.

Stauffer, R. G., & Hammond, W. D. (1967). The effectiveness of language arts and basic reader approaches to first grade reading instruction—Extended into second grade. *The Reading Teacher, 20,* 740–746.

Stauffer, R. G., & Hammond, W. D. (1969). The effectiveness of language arts and basic reader approaches to first grade reading instruction—Extended into third grade. *Reading Research Quarterly, 4,* 468–499.

Stauffer, R. G., Hammond, W. D., Oehlkers, W. J., & Houseman, A. (1976). The effectiveness of language arts and basic reader approaches to first grade reading instruction—extended into sixth grade. In R. G. Stauffer (Ed.), *Action research in L.E.A. instructional procedures,* Newark, DE: University of Delaware.

Stewart, J. P. (1986). *A study of kindergarten children's awareness of how they are learning to read: Home and school perspectives.* Unpublished doctoral dissertation, University of Illinois, Urbana.

Stice, C. F., & Bertrand, N. P. (1987, December). *The effects of a whole language program on the literacy development of at-risk children.* Paper presented at the annual meeting of the National Reading Conference, St. Petersburg Beach, FL.

Stubbs, C. A. (1983). *An investigation of the effect of structure of intellect training on academic achievement with low achieving primary students.* Unpublished doctoral dissertation, University of Tennessee.

Swanson, B. B. (1981). *The beginner's concepts about reading, attitudes, and reading achievement in relation to language environment.* Paper presented at the annual meeting of the University of South Carolina Conference of Educational Research, Columbia, SC. (ERIC Document Reproduction Service No. ED 209 637)

Taylor, N. E., Blum, I. H., & Logsdon, D. M. (1986). The development of written language awareness: Environmental aspects and program characteristics. *Reading Research Quarterly, 21,* 132–149.

Teale, W. H. (1987). Emergent literacy: Reading and writing development in early childhood. In *Annual yearbook of the National Reading Conference,* Chicago, IL: National Reading Conference.

Teale, W. H., & Sulzby, E. (Eds.). (1986). *Emergent literacy: Writing and reading.* Norwood, NJ: Ablex.

Trachtenburg, P., & Ferruggia, A. (1988). Big books from little voices: Reaching high risk beginning readers. *The Reading Teacher, 42,* 284–289.

Tumner, W. E., & Nesdale, A. R. (1985). Phonemic segmentation skill and beginning reading. *Journal of Educational Psychology, 77,* 417–427.

Vilscek, E., & Cleland, D. L. (1968). Two approaches to reading instruction: Final report. (ERIC Document Reproduction Service No. ED 022 647)

Vilscek, E., Morgan, L., & Cleland, D. (1966). Coordinating and integrating language arts instruction in first grade. *The Reading Teacher, 20,* 31–37.

Walraven, S. T. (1981). *The effect of linguistic awareness instruction on linguistic awareness concept development and reading attitudes of kindergarten children.* Unpublished doctoral dissertation, University of Georgia.

Weaver, C. (1988). *Reading process and practice.* Portsmouth, NH: Heinemann.

Wilkinson, I. A. G., & Brown, C. A. (1983). Oral reading strategies of year one children as a function of level of ability and methods of instruction. *Reading Psychology, 4,* 1–9.

Woodcock, R. W. (1967). The Peabody-Chicago-Detroit Reading Project—A report of the second year results. (ERIC Document Reproduction Service No. ED 017 413)

Wrightstone, J. W. (1951). Research related to experience records and basal readers. *The Reading Teacher, 5,* 5–6.

Authors

STEVEN A. STAHL, Associate Professor, Western Illinois University, Elementary Education/Reading Department, Horrabin Hall 58, Macomb, IL 61455. *Specializations:* vocabulary development and reading comprehension, research synthesis.

PATRICIA D. MILLER, Chapter I Teacher, Southern Elementary School, District 120, 37–39 South Street, Stronghurst, IL 61480. *Specialization:* reading.