

THREE

Research on Beginning Reading— Science or Ideology?

ONE OF THE MOST important things, if not *the* most important thing, I learned from studying the existing research on beginning reading is that it says nothing consistently. It says too much about some things, too little about others. And if you select judiciously and avoid interpretations, you can make the research “prove” almost anything you want it to.

What I have learned, too, or at least hope to have learned, is humility. As a result of reading the research of others, interviewing leading proponents of different approaches, and talking informally with administrators and teachers, I have been struck by how easy it is to misinterpret research findings. The best of us can be led into making too hasty conclusions and overgeneralizing from limited evidence. Therefore all of us concerned with reading—laymen, teachers, experts in reading, critics of various approaches—should question whether we have been rigorous enough in our search for evidence.

Since we look to the researchers for our facts, however, we are bound to turn our gaze on them first, to ask, How good a job have you done?

RESEARCH RATIONAL OR EMOTIONAL?

The answer is disturbing. Taken as a whole, the research on beginning reading is shockingly inconclusive.

Why should this be the case? Educational researchers have, after all, devoted more time and effort to the study of reading than to any other school subject.

Despite the volume of research, we could hardly have expected the reading specialists to have produced better results than they did, according to a study by David Wilder (1966). They were never given the institutional support that scientists require if they are to produce definitive research. The schools of education where the researchers studied and worked were cut off from essential contacts with research activity in the basic sciences; these schools did not even offer full-time research positions until quite recently. Furthermore, reading research has never received financial support on the scale and over the period of time required to yield reliable answers to questions that needed answering. Instead of conducting much-needed high-quality research, therefore, the reading specialists turned to teaching courses, speaking at conferences, and giving workshops on reading. Or they engaged in clinical practice and remedial reading or wrote instructional materials for children.

Since research was largely a part-time interest of a minority of experts—indeed, only 1.6 percent of reading experts queried in 1963 by Barton and Wilder (1964) received half or more of their income from research bureaus or projects—it is not surprising that it exhibits serious failings.

The research in beginning reading has generally been inadequate in both depth and scope. The questions it has attempted to answer have usually not been answered well. Other important questions have not been asked at all.

Most studies of beginning reading, it seems, were undertaken *not* because they grew logically out of previously published studies that needed further refinement in order to build a unified theory about the reading process. Instead, it appears as if researchers usually wanted to buttress a strongly held view about a method or practice that was already in use or that was a reaction to one in use in a particular school system at a particular time. Thus, they attempted to answer such practical questions as: Which method is better? When is the best time to start? When the researcher obtained an answer, he seldom went beyond to investigate *what* in the method or practice produced better results and *why*. We can identify a series of practical questions of this type that have been researched over and over again, the most recent studies going very little

beyond the earlier ones. In fact, some of the older classroom experiments, particularly those carried out by Gates, were more firmly grounded in a theory of the reading and learning process than the more recent ones.

The research in beginning reading has also been parochial. Like scientists in other fields, especially in the social sciences, these researchers have been influenced by the philosophical assumptions and social problems of their time, both in selecting problems and, more particularly, in drawing conclusions and making recommendations.¹ Such influence is inevitable, since no one can escape the time in which he lives. However, a researcher must be aware that the present will be tomorrow's past, just

¹ See, in this connection, Huey (1908), especially chap. XV, "The Views of Representative Educators concerning Early Reading." Even before any extensive empirical research was undertaken on the *when* and *how* of beginning reading instruction, Huey cited the views of "our foremost and soundest educators" who were "profoundly dissatisfied" with reading as it was then carried on. Among the things that displeased them were "premature reverence for books . . . and a neglect of own [sic] thinking which has atrophied the naïve originality of the children and made them slaves to 'what is written. . .'"

Huey continues, "As child nature is being systematically studied the feeling grows that these golden years of childhood, like the Golden Age of our race, belong naturally to quite other subjects and performances than reading; and to quite other objects than books; and that reading is a 'Fetish of Primary Education' which only holds its place by the power of tradition and the stifling of questions concerning it. It is believed that much that is now strenuously struggled for and methodized over in these early years of primary reading will come of themselves with *growth*, and when the child's sense organs and nervous system are stronger; and that in the meantime he should be acquiring [his] own experiences and developing wants that will in time make reading a natural demand and a meaningful process, with form and book always secondary to [his] own thought."

Huey then quotes Dewey (New York Teachers' Monographs, November, 1898) and comments as follows: "Professor Dewey says that while there are exceptions, present physiological knowledge points to the age of about eight years as early enough for anything more than an incidental attention to visual and written language-form. . . . While the fetish of Greek is passing, there remains, he says (in "The Primary Education Fetish," *Forum*, vol. XXV) the fetish of English, that the first three years of school are to be given largely to reading and a little number work. This traditional place was given to reading in an early century, when the child had not the present environment of art gallery, music, and industrial development, but when reading was the main means of rising and was the only key to culture. Reading has maintained this traditional place in the face of changed social, industrial, and intellectual conditions which make the problem wholly different.

"Against using the period of six to eight years for learning to read and write, Professor Dewey accepts the opinion of physiologists that the sense organs and nervous system are not adapted to such confining work, that such work violates the principle of exercising the fundamental before the accessory, that the cramped positions leave their mark, that writing to ruled line forms is wrong, etc. Besides, he finds that a certain mental enfeeblement comes from too early an appeal to interest in the abstractions of reading." (pp. 301-305) Thus, a later start and a meaning emphasis seem to have been recommended as early as 1908, long before the 1920s and 1930s when the research findings were gathered.

as it was yesterday's future. He runs grave risks if he ignores what has gone before.

A sense of perspective has also been lacking, to an even greater extent than in the researchers themselves, in the interpreters and summarizers of research on beginning reading. They have tended to overgeneralize findings, especially those that fitted into prevailing educational views. They have paid only slight attention to studies that did not support these views. And they have tended to find all kinds of justification for their stands, even when faced with contradictory evidence.

Too often, the interpreters and summarizers have forgotten that no single research study can stand by itself, that each must be seen in the light of other relevant research on the reading and learning process. If we are to gain lasting benefits for theory and practice, the research must be interpreted periodically as new evidence comes to the fore, not only in reading but also in psychology, linguistics, medicine, sociology, and other disciplines. This, unfortunately, has not been done often enough in the field of reading in general and beginning reading in particular.

These are strong statements. Let me illustrate what I mean with a few examples.

We are still debating when is the best time to begin reading instruction. The study that has had the greatest impact on the prevailing theory and practice was one reported by Mabel Morphett and C. Washburne in 1931. From a comparison of the mental abilities and reading achievement of first graders in the Winnetka, Illinois, schools, Morphett and Washburne concluded that a mental age of six and one-half was probably the best age to start formal reading instruction.

Subsequent studies by Washburne and others appeared to confirm the conclusion that a start at this relatively late age produced better results. In fact, considering a mental age of six and one-half ideal for beginning became part of the conventional wisdom as expressed in most textbooks for teachers on the teaching of reading and most published reading programs in the United States. Even Fred J. Schonell, whose textbooks on the teaching of reading have probably been the most widely used in England, seemed to favor a mental age of six (1948), while all around him five-year-olds were learning to read in the English infant schools, and nannies in upper-class homes were teaching youngsters to read even before they entered these schools.

Did the research really "prove" that a child must reach a mental age of six or six and one-half before he can benefit from beginning reading instruction? Yes, if we ignore the study reported by Arthur Gates in 1937, only six years after the Morphett and Washburne study. Gates found in

one of four classes he studied that children with mental ages as low as five could cope successfully with first-grade reading. Who was right?

If we examine both studies, we must conclude that both were intrinsically valid; yet both were relevant only for the particular situations studied. The Morphett-Washburne findings make sense when we realize that the children they studied were an intellectually superior group: The median mental age at the beginning of grade 1 was about seven. The standard of success (for their first graders) was quite high, and instruction was based mainly on independent silent reading. In such a situation, a child with a mental age below six, not quite ready to do most of the learning on his own and aware that he was less able than the other children, might well be judged a "failure" according to the stringent definition of reading success used by the investigators.

The classes Gates studied had a more normal IQ distribution in terms of the general population. Also, they used easier materials, and the children seem to have received more direct and differentiated instruction. Thus, a child with a lower mental age found himself in a better position to learn. Furthermore, Gates's criterion of success was less stringent.

AND THERE IS A TIME FOR RESEARCH . . .

What is of even greater interest, however, is that the Gates study received only minor attention in the summaries of research and in the textbooks on methods of teaching reading. Why? Probably because Gates's findings ran counter to the current mainstream of educational thought and practice. In the 1930s and 1940s the ideal school was a child-centered institution; its purpose was to provide activities that would encourage the child's inner growth and his emotional and social development. The more informal the program, the better. Educators put little faith in specific training, relying mostly on activities they felt would foster growth and maturation. In fact, they actually feared imposing training on the child lest it frustrate him and produce not only small or no gains in learning but also permanent dislike of the activity taught and permanent personality damage.

Now the climate of opinion has reversed. The first sputnik propelled the United States into a great push toward academic achievement. Then research by Dolores Durkin (1964) and others² indicated that younger

² See Fowler (1962) for a post-sputnik summary of pre- and post-sputnik research on early learning. Significantly, although much of the evidence on early learning has been in existence for a long time, it has really only been put to use after the first sputnik was launched.

children can and do learn to read and that an earlier start may be better than a later one.

We are less concerned at present with social and emotional development and more with academic achievement. The conventional wisdom appears to be changing from "the later the start, the better," to "the earlier the start, the better." This trend is also seen in the various head-start programs for culturally disadvantaged preschool children and in the current renaissance of the Montessori schools, some of which begin reading instruction at age four.

Which position is correct? Time—and the failures that this new set of assumptions may bring—might provide some answers. But it will not be easy to find out unless we follow the children in these new "experiments" over a long period of time. We will be debating this issue again and again unless we undertake—with the kind of financial support that has generally been lacking—long-term longitudinal studies designed to test out fully the effects of either an early or a late start.

It is interesting to ponder whether the studies of Durkin and others would have engaged the interest of investigators during the 1930s and 1940s. I venture to guess that had they carried out their studies then, their findings would have been largely ignored or would have gotten lost in the shuffle, and thus their research would have "said" nothing to us.

That this supposition is not unfounded is suggested by what happened to the writings of Bloomfield and other linguistic scientists on how to begin reading instruction. During the past several years one could hardly pick up an educational journal or attend a conference on reading without finding some discussion of linguistics and its potential contribution to beginning reading instruction. As reported in Chapter 1, beginning reading materials based on "linguistic principles" are coming off the presses at a fast and furious pace, and at least two major books on linguistics and reading have appeared since 1962—Fries (1962) and Lefevre (1964).

Is it that the linguists have only recently completed significant research that they are now making available to us? No. Their findings and recommendations have been known for a long time. As early as 1942 Bloomfield published his two articles in *Elementary English Review* criticizing the then prevailing beginning reading methods and materials, which relied on vocabulary selected primarily on a meaning-frequency basis, and recommending instead a vocabulary based on spelling regularity. Even earlier than that, in *Language* (1933), he made his basic recommendations on the application of linguistics to the teaching of reading and spelling. The Bloomfield teaching materials, which have only recently been published (1961 and 1963), were also written about that time.

Why were Bloomfield's recommendations and materials ignored for so long? The reason is probably to be found in the educational preference of the time. In 1942, when Bloomfield outlined his criticism of beginning reading instruction, most reading researchers and practitioners were convinced that a less formal approach to education in general, and to reading instruction in particular, was desirable. They were veering away from formal teaching. Bloomfield's recommendations and his instructional materials smacked too much of formality, of dull drill, and of the old-fashioned phonics from which they were trying to escape.

Over and over again, if we look back into the research on beginning reading, we find that almost every issue debated and experimented with today has been the subject of study and experimentation at some time in the past. The studies made by Donald Durrell and his students (1958) on the importance of knowing the alphabet (identifying letters by name) for beginning reading success were preceded by the studies of Frank Wilson and C. W. Flemming published in 1938 (see Chapter 5). Durrell came to essentially the same conclusions as Wilson and Flemming. But somehow these earlier works were "lost" and were seldom mentioned in summaries of research.

Even Pitman's ITA is not new in principle. Using modified alphabets as the first step in learning to read has been tried and researched in the past (Winch, 1925). Yet this research appears to have been lost from view during certain periods, only to be revived again at a later date, often under the guise of new discoveries and under new labels.

WHY DON'T WE LEARN FROM THE PAST?

There are many reasons why we don't. Probably the most important one is that most researchers have been concerned more with answering specific practical questions than with accumulating tested knowledge and developing theory. Indeed, most of the research in beginning reading has been the work of teachers and administrators who were interested—and rightly so—in answering practical, timely questions. They undertook one small study either as a doctoral or master's thesis or as "action research." Few of the really major figures in the field of reading instruction have done the major research.

The reading field has also suffered from a dearth of synthesizers and theorists—people who pull together the evidence from the hundreds and thousands of small studies and try to build theories. Of course, the task of synthesizing is always difficult. But it has proved particularly so in the area of beginning reading instruction. Even on the most basic issues of beginning reading—the very definition of beginning reading, what it is and what its goals should be—clarity and agreement have been rare, as I

have shown in my overview of current innovations (Chapter 1) and in my analysis of the views of various proponents on issues in the current debate (Chapter 2). Much of the heat expended in the present debate stems from disagreement on the definition and goals of beginning reading and from the many related issues arising from this disagreement. Unfortunately, while researchers have implicitly agreed on the ultimate goals of reading instruction, no one has conducted experiments to test how effectively a method develops the qualities of mature reading implicit in these goals. Thus, a critic could always state that although a given method might result in the child's recognizing words or understanding a passage better than another method, it was not more effective in terms of attaining a goal of reading that he believed to be more important.

A good illustration of this is the study of the value of systematic phonics in early reading instruction reported by Donald Agnew in 1939. Agnew found that an early, relatively heavy, and consistent emphasis on phonics in grades 1 to 3—as compared with a later, lighter, and less consistent emphasis—resulted in significantly higher scores on tests of phonic ability, accuracy of oral reading, word pronunciation, eye-voice span, vocabulary, and some aspects of silent reading comprehension (ability to follow directions and recall details). However, a heavier phonic emphasis did not significantly improve accuracy and speed of silent reading or other aspects of silent reading comprehension (predicting outcomes and reading for general significance).

Agnew concluded from his findings:

If the basic purpose in the teaching of primary reading is the establishment of skills measured in this study (namely: independence in word recognition, ability to work out the sounds of new words, efficiency in word pronunciation, accuracy in oral reading, certain abilities in silent reading, and the ability to recognize a large vocabulary of written words), the investigations would support a policy of large amounts of phonetic training.³ *If, on the other hand, the purposes of teaching primary reading are concerned with "joy in reading," "social experience," "the pursuit of interests," etc., the investigations reported offer no data as to the usefulness of phonetic training.* (p. 47, emphasis mine)

Tracking down the different interpretations of Agnew's conclusions proved to be an interesting bit of detective work. Agnew's study, some writers reported, indicates that an early emphasis on phonics promotes attainment of some important goals of beginning reading. Others, however, said that although Agnew found early, heavy, and consistent phonics to be a significant advantage in terms of word recognition, it did not affect interest!

The point is that no method has ever been systematically tested in

³ Phonics training.

terms of pupil interest. Yet over and over again, a particular method is said to be more interesting, to produce readers who are more interested in reading, and therefore to be preferable. As would be expected, the "more interesting" approach is often the one favored by the researcher himself.

Until there is some agreement on the major goals of beginning reading, it will always be possible to reject experimental findings by conveniently shifting goals. It is easy to say, "Yes, you found greater gains for some goals of reading, but not for those I think important."

Imprecise terminology is another cause of difficulty in synthesizing research findings. Throughout the experimental literature, methods are called *look-say*, *sight*, *word*, *phonic*, *alphabet*, *sentence*, *natural*, *organic*, *eclectic*, and now *linguistic*. Seldom are these labels fully defined, and what one author calls a *sight* method another may call a *combination sight and phonic* or an *eclectic* approach. What W. H. Winch (1925), for example, called *look-say* is closer to the Bloomfield-Barnhart (1963) *linguistic* approach of today. Even the same terms have undergone historical changes; a *word* method of 1920 seems quite different from a *word* method of 1960.

A third difficulty is that researchers in reading, like researchers in other scientific fields, have looked to acclaimed leaders in their field to set the framework for the design and interpretation of their investigations. It is difficult to understand why, in the face of a great deal of experimental evidence showing more favorable results for a code emphasis, so many researchers either have failed to acknowledge their own and others' findings or have suggested that a code emphasis fulfills minor objectives only. Yet phenomena of this type are quite common in the natural and physical sciences as well.⁴

⁴ See Kuhn (1962) for a discussion of how resistance to "new scientific truth" has characterized the history of the natural and physical sciences. He writes: "In the past [this strong resistance] has most often been taken to indicate that scientists, being only human, cannot always admit their errors, even when confronted with strict proof. I would argue, rather, that in these matters neither proof nor error is at issue. The transfer of allegiance from paradigm to paradigm is a conversion experience that cannot be forced. Lifelong resistance, particularly from those whose productive careers have committed them to an older tradition of normal science, is not a violation of scientific standards but an index to the nature of scientific research itself. The source of resistance is the assurance that the older paradigm will ultimately solve all its problems, that nature can be shoved into the box the paradigm provides. Inevitably, at times of revolution, that assurance seems stubborn and pigheaded as indeed it sometimes becomes. But it is also something more. That same assurance is what makes normal or puzzle-solving science possible. And it is only through normal science that the professional community of scientists succeeds, first in exploiting the potential scope and precision of the older paradigm and, then, in isolating the difficulty through the study of which a new paradigm may emerge." (pp. 150-151)

Some light is thrown on this situation by the writings of the late William S. Gray, acknowledged leader of, and spokesman for, reading experts for four decades; major summarizer and interpreter of reading research; and author of America's leading basal-reader series, the Scott, Foresman & Company series.⁵ In 1948, in *On Their Own in Reading*, a book dedicated to "all who guide children in their efforts to acquire independence in word attack in reading," Gray views with alarm the swing, since 1900, from one extreme ("an undue emphasis on learning the form and sound of separate words")⁶ to its opposite ("guessing from context with little attention to the visual form of words").⁷ At the same time, however, he dramatically poses the question: "Shall we, in response to public demand, reinstate the old mechanical phonic drills and content that inevitably result in dull, word-by-word reading?"

You can guess the answer. To Gray, and to other leaders in reading, phonics meant a return to "drill" and the "dull content" of the phonics readers of the early 1900s. Gray explains his objection as follows:

In the very nature of things, reading material constructed on this artificial basis [words selected on phonic elements previously taught] was certain to lack continuity of thought. Indeed, pages of such primers and first readers may be read almost as effectively by beginning with the last sentence and reading to the top of the page as by reading in the usual way from top to bottom. (p. 19)

He reproduced a page from the Beacon primer, a phonic reader of 1912, to illustrate his points.

I reproduce the same page, together with a page from Gray's own primer, which selects words and content for "meaningfulness" and "continuity of thought," and I invite the reader to try both pages from bottom to top, as well as from top to bottom.

In my opinion, both are improved by being read from bottom to top. They have more punch! At the same time, I imagine we could debate at length which is the duller.⁸

What is essential, however, is that Gray assumed that an early emphasis on phonics (or decoding, in our terminology) led inevitably to

⁵ Gray was coauthor, with William Elson, of the 1930 edition and senior author of the 1940 and 1950 editions.

⁶ My code emphasis.

⁷ My meaning emphasis.

⁸ Actually, it is possible to write interesting readers with either kind of restriction. Dr. Seuss has done so, and so have many others who have followed his lead. See his *Hop on Pop*, a veritable phonic or linguistic reader that controls words on a common phonic-element or spelling-pattern principle, and *Green Eggs and Ham*, a book that contains only fifty high-frequency words common to primers and first-grade readers in the conventional basal-reading series. Both are published by Beginner Books, Random House.

Reproductions of:

Page 50 of the Beacon Primer, 1912
(a phonic reader)*

Page 41 of *The New Fun with
Dick and Jane*, 1956 (a meaning
reader)†

*Black and white ink drawing of a
horse*

*Full-color illustration of Dick jump-
ing over blocks with entire family
looking on*

My name is Dick.

Do What I Do

I am a big horse.

“See me jump,” said Dick

You may pat me.

“Oh, my! This is fun.

You may ride me.

Come and jump.

Will you ride on my back,
Tom?

Come and do what I do.

I will not run fast.

Look, look!

I will not kick.

Who can jump?

I will not jump.

Who can do what I do?”

I will stand still.

I like to have Tom ride on
my back.

I can run like the wind.

* From J. H. Fassett, *Beacon Primer*, Ginn and Company, Boston, 1912.

† From W. S. Gray et al., *The New Fun with Dick and Jane, Primer of The
New Basic Readers*, Scott, Foresman & Company, Chicago, 1956.

dull content and that it also made for mechanical, "word-by-word" reading, which impedes comprehension and enjoyment. As an alternative, he gave his endorsement to the "meaning-first, word-analysis later" approach that was adopted by his own and most other basal-reading series. This endorsement had formidable effects. As A. Sterl Artley noted in his foreward to the 1948 edition, "[The] worth [of this approach] has been demonstrated in our schools so convincingly that today his philosophy, ideas, and procedures—even his nomenclature—are not only generally accepted but are being incorporated into most instructional materials."

I have painted a rather dismal picture here, and some perspective is in order. The confusions and difficulties that arise when generalizations are derived from existing research are not peculiar to the field of beginning reading. They characterize other fields in education as well.

Until quite recently researchers in reading have had to do their thinking and investigating on a shoestring. When we consider that most of the early research was done by professors, teachers, and administrators in their spare time and with the occasional help of a doctoral or a master's student, it is surprising that we know as much as we do.

The support that is now being made available for educational research by private foundations and government is reason to take a more optimistic view of the future.⁹ Even more encouraging is the growing trend toward buttressing the research with a firmer theoretical basis from other disciplines, particularly psychology and linguistics.¹⁰

Finally, I certainly am not saying that the research has been valueless. If this were in any way the case, there would be no point in analyzing it at length, as I do in the remainder of this section.

⁹ It will not all be clear sailing, however. More research may engender more confusion if it is not synthesized and put into a meaningful framework. Furthermore, while large research budgets may give us the wherewithal to uncover more facts, a sense of humor and simple honesty are needed if negative findings are to be reported from, say, a ½- or 1-million-dollar research project carried out over a five-year period, or even longer. Indeed, more research will require more people to read, sift through, and make sense out of the findings resulting from it. We will need, most of all, the few who have the imagination and insight to create the theories which may then be confirmed or denied by years of careful research.

¹⁰ See, for example, the *Project Literacy* reports.