

Ralph A. Hanson

*Hanson Research Systems, Garden Grove, California, USA*

Donna Farrell

*University of Science and Arts of Oklahoma, Chickasha, USA*

Reading Research Quarterly  
Vol. 30, No. 4  
October/November/December 1995  
©1995 International Reading Association  
(pp. 908-933)

# The long-term effects on high school seniors of learning to read in kindergarten

When to begin teaching children to read is an educational policy issue of continuing public interest and controversy. Recent research suggests that the quality and timing of early reading instruction may considerably improve the reading ability and schooling achievement of children in the early elementary grades (Beck, 1973; Burton & Jones, 1982; Chall, 1983; Gersten & George, 1990; Mason, 1984). However, such research has not addressed the major criticism of this practice: that, over the long term, it will have a serious negative impact on the reading skills and attitudes of students (Morphett & Washburne, 1931).

The lack of solid empirical data to confirm or refute this argument has resulted in educational policies and practices that vary widely among countries, states/provinces, and districts, and even within school districts. In the United States, for example, most children begin their initial reading instruction in first grade, while a few begin in kindergarten (Bennett, 1986). Still others are reported to receive it as early as preschool (Beck, 1986).

These differences can be traced to various sources, but state/province, district, and school policies are clearly influential. States and provinces specify the legal age at which children are allowed to enter public school; districts and schools have various policies regarding retention or placement of kindergarten-age children in "developmental" or "transitional" classes. In addition, and not insignificantly, parents also contribute to the beginning reading process, such as by attempting to teach

reading, by selecting a preschool program, and by choosing when to enroll their children in kindergarten, since parents often have the option of delaying, but not accelerating, the start of formal schooling. All these decisions can and do determine when a child begins to receive reading instruction. (See Hall, 1963; Siegel & Hanson, 1991; and Wolf & Kessler, 1987, for a further discussion of these factors.)

The result of these differences is that some children begin receiving reading instruction as early as age 4, while others are as old as 8 years before such instruction begins. This difference in when children begin receiving reading instruction is the fundamental issue addressed by this study. The essential research question is, "Do differences in the age at which children begin receiving formal reading instruction have any measurable impact on their subsequent schooling experiences, reading achievement, attitudes, and literacy levels as young adults?"

The information for this inquiry was gathered in a U.S. national follow-up study ( $N = 3,959$ ) of high school seniors, supported by a grant from the U.S. Department of Education (Hanson, 1984). Some of these seniors were enrolled in kindergarten in 1973 and attended specific elementary schools that provided formal reading instruction in kindergarten. The study compares the subsequent reading and school-related experiences of these students, from kindergarten through their senior year in high school, with those of students from the same districts who did not receive the kindergarten reading instruction.

## Research and development of the Beginning Reading Program

The Southwest Regional Laboratory (SWRL) is one of a number of regional educational laboratories created in the 1960s as part of the U.S. "Great Society" legislation. At the outset, a major goal of SWRL was to provide a new generation of research-based educational programs for use by schools using advanced research and development (R&D) procedures (see, e.g., Baker & Schutz, 1971; Schutz, 1970).

The first specific program SWRL set out to produce was a beginning reading program for use in kindergarten. This effort led to the development of a program eventually referred to as the Beginning Reading Program (BRP) which is the basis for the research reported here. (It is referred to by the full name, the acronym "BRP," or simply "the Program").

Beyond the fact that it was a formal reading program designed for kindergarten children, the BRP represented a major departure from existing programs in a number of ways. Although the development of this program began rather modestly, it came to be a focus of extensive work by SWRL staff for more than 15 years. Indeed, it became the means by which SWRL defined and operationalized a philosophy and approach to educational research and development that became known as "programmatic R&D." (See Schutz, 1970, for a full discussion of this approach.)

Consistent with that approach, the specific goal of the BRP project was to develop a state-of-the-art beginning reading program that could teach all children in kindergarten to read without major alterations to the existing kindergarten classroom structure (Hanson & Schutz, 1975). Although this goal was clearly accomplished, as verified by many of the nationwide studies eventually conducted on it (Hanson & Schutz, 1975, 1976, 1978, 1986), the R&D effort required to design, develop, and field test the Program was, and still remains, unprecedented (e.g., Durkin, 1990).

This extensive R&D effort focused on four specific areas consistent with the programmatic R&D strategy developed by Schutz (1970). In the case of the BRP, this R&D effort occurred in four ordered, but overlapping, phases:

### *1. Identification of the program content parameters and instructional requirements.*

The research in the first phase was conducted primarily by reading specialists and linguists and was initiated at the onset of the Program's development, in the middle 1960s. Both analytical and empirical research was carried out on many aspects of beginning reading, such as investigating the specific words and elements to be

taught, the optimum lexical pace and order of presentation for a beginning reading program, and the size of the type that was most comfortable for young children to read (e.g., Hanson & Ross, 1973). The most well-known individual among the many specialists involved in this work was Frank Smith (1967, 1969a, 1969b); however, many others contributed to this effort (e.g., Follett, 1971; Koehler, 1971, 1972; Lott, 1969; Smith, Lott, & Cronnell, 1969; Sullivan, 1973). One hallmark of this collective effort to design and define the BRP was that no single philosophy or approach was dominant. Thus, the focus of this work was kept on developing children's reading skills, defined as the ability to read and understand simple stories in the most direct and efficient manner.

### *2. Research and development on the optimum strategies for teaching beginning reading skills.*

During the second phase, using the design specifications provided, a development staff—composed primarily of educational psychologists—worked with artists, story writers, and others to create prototype lessons and units and, eventually, a prototype program (Niedermeyer, 1972; Niedermeyer & Fischer, 1971; Quellmalz, 1973; Sullivan, 1968, 1973). Once again, various types of research were involved and ranged from single lesson tryouts (at SWRL) to field tests of units in actual school settings, either conducted by SWRL staff or with trained SWRL observers present (Hylton & Quellmalz, 1970). This research also included studies of other components designed to help supplement regular classroom instruction through the use of parents and tutors (e.g., Hanson, Kaplan, & Yaman, 1971; Niedermeyer & Ellis, 1970; Sullivan & Labeaune, 1971).

### *3. Research on the necessary installation requirements for effective functioning of the program in actual school settings.*

The research in the third phase included the development of training and installation materials and a district/state evaluation/information system. These components were designed to (a) train teachers to use the Program, (b) provide ongoing data on the implementation process, and (c) obtain measures of the direct effects on students from using the Program. This work began in the late 1960s with the full-year tryouts of the Program (at that time called the First Year Communication Skills Program). Between 1968 and 1971, full-year tryouts of the Program were conducted each year in 10 to 20 districts in southern California (e.g., Hanson & Berger, 1971; Hanson & Resta, 1970; Hanson & Ross, 1973; Resta & Niedermeyer, 1970). The work included extensive data gathering, training of the schools' kindergarten staff, and formal and informal observations and

testing of kindergarten children using a variety of instruments (Hanson, McMorris, & Bailey, 1986; Niedermeyer, 1971; Niedermeyer & Giguere, 1972).

#### *4. Quality assurance evaluations during the national implementation of the Program.*

The previously mentioned R&D resulted in a simple, efficient, and effective product that could be readily implemented in virtually any kindergarten classroom and was now referred to as the SWRL/Ginn Beginning Reading Program (SWRL, 1972). In 1971 the Program was designated as an exemplary educational product and was therefore deemed eligible to be purchased by schools using Title III funds. (Title III was a U.S. government program that provided funds to states to purchase, use, and evaluate designated new programs.) This offer applied to 17 designated states as well as Guam and the Virgin Islands. By separate contractual agreement, SWRL staff provided the training of district staff at regional meetings. Also, SWRL staff were contacted as part of the Title III evaluation requirement to gather and compile information on the implementation and effects of the Program and to provide reports to various audiences. These audiences included federal officials, district and school staff, and classroom teachers. This work led to a series of "schooling effects" studies that described the findings from this national effort (Hanson & Schutz, 1975, 1976, 1978, 1986).

### ***Description of the BRP resources***

The final version of the Program embodied the carefully researched design and content specifications mentioned previously that sought to provide optimal sequencing and presentation of the critical sounds and words needed to gain initial competence in reading (e.g., Adams, 1990). Children were introduced to a basic sight vocabulary and decoding skills that allowed them to begin reading short stories after only the first week of instruction.

These beginning reading skills were introduced and taught using specially developed materials and developmentally appropriate instructional procedures. These procedures involved using group flashcard lessons, playing simple games, and reading from a set of 52 special story booklets. The story booklets, which were the central focus of instruction, consisted of illustrated stories about a series of animal characters that capitalized on the children's natural affinity for animation. The children read them aloud and, to allow for additional practice and to increase parental involvement, each child was provided with the full set of these consumable booklets to take home. At the back of the booklets were discussion questions about the story characters and plot, which were

used by the teacher or parent to verify the child's understanding and encourage discussion.

Teacher training workshops, using carefully developed audiovisual components followed by discussion sessions, were carried out in each district before implementing the Program. The general goal of this teacher training was to familiarize teachers with the Program's content and instructional methods in order to allow mastery of the beginning reading skills by all students. These workshops placed great emphasis on the importance of encouraging each child to participate as he or she became ready. Teachers were told to positively acknowledge every effort on the part of a child—including non-responses. With the focus on responses from each child and the division of the Program into 10 units (each followed by a brief skills assessment), teachers had the means to confirm that each child had learned each skill and concept. For children having difficulties, additional individual instruction was recommended and special materials provided.

The BRP was widely disseminated during the early 1970s using Title III monies. Elementary schools representing about 15% of U.S. kindergarten pupils adopted it in both the 1972-73 and 1973-74 school years. During the 1973-74 school year, more than 2,000 elementary schools in 400 school districts across the United States implemented it in their kindergarten classes for teaching their students to read. These included both public and private school systems, most of which served large numbers of disadvantaged students (Hanson & Schutz, 1975, 1976, 1978). The BRP was extremely well received and, even though Title III funding ended in 1974, many districts continued to use the Program and the accompanying testing and information system for years afterward (Hanson, Lehman, & Bailey, 1981).

During the national implementation of the Program, both the implementation and direct effects of the BRP were documented through data provided by the unique evaluation and monitoring system embodied in the Program (Doherty & Hanson, 1974; Hanson, 1979; Hanson & Bailey, 1971; Hanson & Berger, 1971). The simple, criterion-referenced assessments administered upon completion of each unit provided detailed measures of the reading ability of every kindergarten class, school, and district that participated in the Program in each of the Title III years. These same tests also provided explicit information to the teachers on the reading progress of each of their students.

A series of SWRL technical reports documented the reading skills acquired by the children who used the BRP (e.g., Hanson, Lehman, & Bailey, 1981; Hanson & Schutz, 1975, 1976). These national reports clearly indicated that the children's reading ability was directly

linked to the Program implementation practices employed by teachers and schools. That is, factors such as the date instruction began and the time the teacher spent each week using the Program predicted the number of units completed by the students (Hanson, Lehman, & Bailey, 1981; Hanson & Schutz, 1978). Thus, although there were wide differences in the number of BRP units completed within each classroom, both within and across districts, the data clearly showed that these differences were due to decisions made by school personnel regarding the time and effort devoted to teaching the Program rather than to either the biosocial (e.g., ethnic background, gender, or social class) or behavioral (e.g., entering language skills) characteristics of the kindergarten students and classes.

For example, about 25 weeks were required, on average, to complete all 10 units of the BRP; however, although some schools and teachers used it for the full 25 weeks and completed the entire 10 units, others opted to spend fewer weeks and provide less instruction. Thus, although all the children who participated in the Program acquired some reading skills, the differences in the instructional time spent led to differences in the number of BRP units completed by students at the class and school levels. This resulted in substantial differences between classes and schools in the reading abilities of their students at the end of their kindergarten year (see Hanson & Schutz, 1978).

Because the progress made by each class, school, and district that participated in the BRP was carefully documented in the series of national evaluation studies mentioned earlier, a large, descriptive database was compiled. Collectively, these BRP studies involved data from more than one-half million kindergarten students from 500 school districts across the United States over 10 years (1969–79). The accessibility of this extensive database made it possible to identify the BRP schools and conduct a follow-up study 12 years later.

### ***Prior research on early reading***

The original series of BRP studies of the 1970s and 1980s had clearly demonstrated that teaching children to read in kindergarten could be readily accomplished with a modest expenditure of resources and minimum alterations to the regular kindergarten curriculum. Kindergarten children from virtually all backgrounds and school contexts became proficient beginning readers when only 20 to 30 minutes of daily instructional time was devoted to the teaching of reading in the context of an otherwise standard kindergarten environment. In essence, this research refuted the traditional notion that pupils from at-risk (disadvantaged and minority) backgrounds are not

ready to read upon entering kindergarten. The only differences observed in their rate of progress in learning to read, as compared to their more advantaged peers, was a slight increase in the instructional time required to complete the first several units of the Program.

These findings were not based on a single study; rather, they were extensively replicated in schools and districts from all over the United States across a number of school years (Hanson, Lehman, & Bailey, 1981). Further, they were also replicated with different kindergarten reading readiness programs in a special program-fair evaluation study (Hanson, Schutz, & Bailey, 1977).

Although these findings, as well as the reactions of parents, teachers, and administrators, were all supportive of teaching reading in kindergarten, they still did not answer the most serious criticism of teaching children to read early: early reading will have detrimental long-term effects on children's reading achievement, interests, and general attitude toward school. That proposition is supported in theory by developmental psychologists such as Gesell and Piaget (e.g., Ames, 1950; Gesell, 1954; Ilg & Ames, 1964; Inhelder, 1962), as well as many of the current early childhood curriculum professionals (e.g., Durkin, 1987a; Kuczen, 1987; Willis, 1993). However, it was first popularized in the 1930s by Carlton Washburne, a prominent leader of the progressive education movement. Washburne was involved in several influential studies that provided the only empirical support for this notion (i.e., Morphett & Washburne, 1931; Washburne, 1936). Most notable is the Morphett and Washburne (1931) study, which concluded that teachers could decrease the chance of reading failure by postponing reading instruction for students until they reached a mental age of 6½ years (about the middle of first grade). Mental age was measured using standard IQ tests, which were gaining widespread use at that time (Arthur, 1925).

Over the years, many researchers have heavily criticized this research and dismissed this single empirical study showing support for *not* beginning formal reading instruction until the age of 6½ (e.g., Beck, 1986; Downing, 1963; Gates, Bond, & Russell, 1936, 1939; Gray, 1937; Prescott, 1955; Tyler, 1964). Yet, this policy continues in most elementary schools today. As others have noted, the subsequent criticism of this study, as well as the research conducted on this issue for more than 50 years, has failed to overcome the unique impact of the original Morphett and Washburne study (Durkin, 1966, 1968, 1987b; Wolf & Kessler, 1987). Downing (1963) summarized these feelings by stating:

It is obvious that generalization from this limited study to the world-wide reading situation was quite unjustified, yet this is what happened. (p. 27)

Thus, in spite of this one flawed study (see Durkin, 1987b; Wolf & Kessler, 1987), extensively criticized for 50 years, and the more recent controversy over the use of standardized IQ tests by schools in subsequent decades, the enduring legacy of the Washburne recommendation continues to manifest itself in the beginning reading policies and practices in the majority of U.S. elementary schools. As Tyler so aptly noted in a thorough review of the research on this issue years ago, "Old myths neither die nor fade away" (Tyler, 1964, p. 226).

Meanwhile, support for providing reading instruction *prior* to first grade continues to emerge from a number of sources. These include findings from earlier research reviews such as those cited in Tyler (1964); the more recent, extensive work by Durkin (1987b) and others on children taught to read before entering first grade (e.g., Beck, 1986; Blatchford, 1987; Mason, 1984; Smith, 1976; Stewart, 1985; Teale, 1980; Teal & Sulzby, 1987; Tobin & Pikulski, 1987); and the emerging findings about changing behavior and skills of young children (e.g., Hanson & Kelble, 1985; White, Alexander, Wilson, & Fuqua, 1985). Collectively, this body of research addresses the conceptual, operational, and practical problems associated with teaching preschool children to read. It demonstrates that children are able to, and often do, learn to read well before entering kindergarten and seem to retain their advantage in subsequent years. Also, it shows that children do not have to be from advantaged homes to learn to read early.

Although most children who come to kindergarten with reading skills are typically from advantaged backgrounds (Durkin, 1987b), children from all types of backgrounds can learn to read at least during kindergarten (before age 6) (Brzieński, 1967; Carmine, Carmine, Karp, & Weisberg, 1988; Englemann & Englemann, 1966; Hanson & Schutz, 1978; Ulrich, Louisell, & Wolfe, 1971; Weisberg, 1987). Further, there are many signs that early reading skill acquisition facilitates later skills in reading (e.g., Ball & Blachman, 1991; Beck, 1973; Burton & Jones, 1982; Chall, 1983; Durkin, 1966, 1968; Gersten & George, 1990; Lundberg, Frost, & Peterson, 1988; Rudolph & Cohen, 1984). These later studies link early reading competence to later competence levels, though most do not define the specific intervention responsible for it. It is this issue of long-term effects that this study sought to address—more specifically, to see if adult reading levels could be linked to teaching children to read in kindergarten.

### ***The BRP follow-up study***

During the 1985–86 school year, with the support of funds from the U.S. Department of Education

(Hanson, 1984), data were gathered from 3,959 high school seniors who were attending high schools in 24 school districts across the country, 23 of which had implemented the kindergarten BRP in either some or all of their elementary schools during the 1973–74 school year. Follow-up data were obtained from the high school seniors in these 23 school districts, including not only students who had received kindergarten reading, but also other seniors in the same high schools who had similar schooling experiences, but had not received BRP instruction. (Data were also inadvertently obtained from high school seniors in one school district that had not implemented the BRP in any of their elementary schools.) These differences in kindergarten experiences, along with current follow-up data, allow the examination of questions about the effects of providing formal reading instruction in kindergarten on the reading skills, educational experiences, and attitudes of high school seniors.

### ***Study design***

Although the logic for doing the follow-up study of the kindergarten BRP students was clear, the procedures needed to carry out such a study were not. This is due in part to the fact that, although there is much discussion of the long-term benefits of various educational programs, policies, and practices, few longitudinal and follow-up studies involving programs for preschool children are actually carried out (see Hanson & Siegel, 1991). Even more limiting was the fact that the majority of the reading studies focused on programs that began in first grade or above (e.g., Juel, Griffith, & Gough, 1986).

Some famous follow-up studies have been carried out on other attributes using other approaches. Especially noteworthy are the various longitudinal studies of young children that began in the 1930s and have been widely used since (e.g., Jones & Bayley, 1941; Sontag, Baker, & Nelson, 1958). Another is the widely cited follow-up studies by Terman and Ogden (1947, 1959) of adults who were considered gifted as children.

Some other studies carried out more recently or currently under way, such as those surrounding the High School and Beyond and Head Start projects, may also prove to be beneficial (e.g., Consortium for Longitudinal Studies, 1979; Meek & Armstrong, 1983; Office of Educational Research and Improvement, Center for Statistics, 1986). However, due in part to the long time period and extensive cost and requirements needed to carry them out, longitudinal and follow-up studies are the exception in education and are still relatively few in number.

A second problem encountered was the scarcity of practical procedures and designs to guide such efforts. Recommended procedures are usually grounded in clas-

sical experimental design methods with recommendations such as "random assignment of subjects to treatment groups" and other requirements that, either for ethical or logistical reasons, are extremely difficult to fulfill in applied settings (e.g., Boruch & Pearson, 1985; Goldstein, 1979). Thus, the existing literature was of limited value in guiding the study design.

However, despite the lack of precedents, the procedures required to carry out this particular type of follow-up study appeared quite simple at one level. We only needed to obtain data on the reading skills, habits, and attitudes of high school seniors from a sample of the districts and schools that had implemented the BRP in kindergarten 12 years earlier and compare them to other high school seniors from the same districts who had not received the kindergarten BRP instruction.

After dismissing the logistical complexities associated with such a strategy, the critical question of how to make meaningful comparisons remained; specifically, "How can comparison groups be formed to assess the effects of kindergarten reading on the skills, attitudes, and other characteristics of these high school seniors?" This is the familiar issue of defining the appropriate control or comparison group(s), and it is clearly different in this type of research from that in classical experimental designs (e.g., Campbell & Stanley, 1966).

To address this issue, an evaluation research strategy was defined in which factors using characteristics of the program, participants, and settings were used to form ad hoc and post hoc comparison groups (see Hanson, 1986; Hanson & Siegel, 1991). A central issue in this type of evaluation research is to identify naturally occurring program implementation differences and use them as the basis for forming comparison groups. Using such comparison groups, logically related effects variables can be analyzed for differences. Any finding can then be replicated across (a) specific student subpopulations of interest, (b) different settings (classes, schools, and districts), and (c) logically related dependent variables. The credibility of this type of research depends on both the replicability and the generalizability of the results produced by this cross-validation process within and across study subpopulations.

These subpopulations are defined in the ways mentioned previously and also by special interest groups (gender, ethnicity, and social class) and by settings (classes, schools, and districts). An even more ideal arrangement is to replicate the results across successive studies carried out across several years and/or cohorts of students (see Hanson, 1986; Hanson, Lehman, & Bailey, 1981). In this way, the research strategy embodies elements of operations research in combination with more traditional methods to evaluate policy alternatives.

This general approach was employed in the original, large-scale inquiries into the kindergarten BRP effects (Hanson, Lehman, & Bailey, 1981; Hanson & Schutz, 1978). In those inquiries, substantial numbers of participating students provided early reading achievement information via the implementation and assessment instruments embodied in the Program that were developed for this purpose. This information was then used to generate management reports for various audiences (such as school, district, and federal documents) and analyzed using the ad hoc and post hoc comparison group strategy described previously for research and policy analysis purposes (e.g., Hanson, 1986). Results were replicated across many groups and settings, as well as across the same settings over successive years (see Hanson, Lehman, & Bailey, 1981).

Another issue in using this analytic strategy was the definition of the characteristics of students that might be affected by learning to read. In the original series of BRP studies, the skill assessment instruments were designed to be function-specific and, therefore, sensitive to any differences that might be present (Bloom, 1967; Hanson, McMorris, & Bailey, 1986). The present inquiry began with this same concept of function-specific assessment (Hanson & Siegel, 1988a) and expanded upon it. Thus, defining the potential effects variables was a major task of the study.

This effort resulted in the careful development and use of three separate instruments eventually compiled and presented in a single survey booklet. These instruments sought to provide information on the reading skills and attitudes of students as high school seniors. However, they also sought to provide information on other characteristics within a model of reading development involving both the home and school (see Siegel, 1987, 1990). Put simply, reading was viewed as a skill that is learned across development (from birth through adulthood).

Across this period of development, there are many indicators of how skill development is progressing, including test scores, academic grades in school, remedial reading needs, and general educational achievement indicators in both elementary and high school. If the kindergarten reading experience had an impact on students at or near age 6, then it might be reflected in the students' reading skills and attitudes, not only as high school seniors, but also in other indicators of reading progress experienced across development. That is, the effects of early reading should show a clear pattern of impact on factors throughout, not only at the end of formal schooling. Accordingly, the study sought to provide a coherent set of effects variables to be evaluated as part of the overall analytic strategy.

In summary, every effort was made in the BRP fol-

low-up study design to obtain information that would facilitate and complement the analytical and measurement strategies developed in the original BRP inquiries. Critical to this design was the availability of the original data that could be linked to a follow-up sample composed of a large number of participants representing all the diversity in the original population in terms of different districts, schools, social classes, ethnic groups, and categories of kindergarten BRP implementation (i.e., number of BRP units completed). Collectively, such information would create a follow-up study database that would be used to define comparison categories reflecting the different kinds of kindergarten reading experiences students had. Then, within these categories, various effects variables, in logically related clusters, would be examined. Any significant result would be cross validated by examining other related variables and comparison groups. In this way, a more coherent picture of the results than those derived from any single study would be produced. In essence, this inquiry can be viewed as a single study that is replicated many times.

### ***Instrumentation, participants, and design strategy***

Conducting the BRP Follow-up Study required the completion of three related groups of tasks. One set of tasks was the design, development, pilot testing, and final preparation of the data collection instruments. This set of tasks resulted in the development of two instruments: The Reading Biographer and the Reading Vocabulary Test, and acquiring permission to use the reading comprehension items from the Academic Instructional Measurement Systems (AIMS). All three instruments were eventually incorporated into a 16-page data-collection pamphlet called the Student Booklet.

A second set of tasks was securing the participants and gathering the follow-up data. This involved identifying the BRP follow-up schools and districts, orchestrating the various activities required to obtain and maintain their participation, and performing the actual data collection, coding, and entry process.

The third and final set of tasks was the formulation of the study design strategy, analysis of the data, and documentation of the results. In the following sections, the procedures followed to complete each set of tasks are summarized. A more complete account of each task is given in Hanson and Siegel (1988b).

### **Study instrumentation**

Three study instruments were designed to carefully assess the reading ability and major factors, from

preschool through high school, related to the reading competence of high school seniors. Assessing reading ability was accomplished through the use of two achievement measures and assessing the major factors was accomplished through the use of a self-report survey instrument. Each of the three instruments are described in the following sections.

1. *The AIMS Reading Test.* A standardized test of reading comprehension that is part of the Academic Instructional Measurement System (AIMS) was the first of two reading measures. This test, which was used with permission of the publisher, Charles Merrill, was developed by Darrell Sabers (1985) and contains 45 items. It provided a standardized test of students' vocabulary comprehension that could be easily interpreted and also included the two reading comprehension skill components, literal and inferential, needed for categorizing students in terms of Chall's (1983) model of reading development.

2. *The Reading Vocabulary Test.* The second reading measure was a specially developed test of reading vocabulary (see Broach, 1988). This test was designed to provide an estimate of each student's vocabulary size. Additionally, it would provide the reading vocabulary estimate needed to categorize students into one of Chall's five reading development stages. The Reading Vocabulary Test accurately estimates the reading skill level of students based on their knowledge of a carefully selected and calibrated set of reading vocabulary items.

To create this measure, a stratified sample of words, based on both frequency of occurrence and use across subject areas, was used. In Chall's model, vocabulary levels shift from frequent, concrete words to more infrequent, abstract words across development. The vocabulary test design corresponds to that model with the items being selected from the *American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971) by their frequency of occurrence and dispersion across subjects as suggested by the theory of receptive vocabulary (Stenner, Smith, & Burdick, 1983). This theory suggests that item (word) difficulty is a function of the likelihood of the word having been encountered and the range of contexts in which the word is likely to have been previously encountered.

To determine the frequency strata to be sampled, estimates of vocabulary sizes by ages from Anderson and Freebody (1981) were interpolated. For adolescents, this procedure resulted in an estimate for vocabulary size in the 50,000 to 80,000 range. Because of the overrepresentation of probable underachievers in the sample, the decision was made to start at 5,000 and sample at intervals of 20,000 words.

Several versions of a prototype of the test were refined in studies on two separate samples ( $N = 19$ ;  $N =$

55) drawn from local high school classes. Items were evaluated, and those either too easy or too hard, relative to their frequency of occurrence, were revised or eliminated. The final version of the test consisted of 25 items, with 5 items drawn from each of 5 frequency bands (i.e., 5–25K, 25–45K, 45–65K, 65–85K, and 85K+). Each word was embedded in a simple declarative sentence with foils that followed a prescribed pattern and would make grammatical sense when inserted into the stem.

After use in the study, the technical quality of the Reading Vocabulary Test was confirmed and one change made to the scoring. The latter pertained to the students in the highest vocabulary level category (85K+ words). Because it contained few words, this category was collapsed with the next highest level into a single category, defined simply as greater than 65,000 words (65K+).

The internal consistency reliability (KR20) of the test was estimated at .73, which is considered an acceptable overall estimate for a test of this length (25 items). However, a more appropriate estimate is the extent to which the test accurately and consistently placed students in the designated vocabulary categories (1 to 4). One indicator of that is the extent to which the score distributions within each vocabulary category corresponded to expectations (i.e., discriminated between those students in the adjacent categories). The mean number of items correct for students in each of the four vocabulary categories were, respectively, 11.9, 15.7, 18.2, and 21.3, showing the expected discrimination pattern.

The overall validity of the approach was also evident in various ways. Among these were the fact that overall item difficulties corresponded to the word's decreasing frequency of use. Also, the test sorted students into the predicted developmental stage categories. Finally, as expected, the scores on the vocabulary test were highly correlated with scores on the reading comprehension skills test ( $r = +.64$ ).

These two instruments, the AIMS Reading Comprehension Test and the Reading Vocabulary Test, were used both alone and in combination to measure the reading competence of the study participants. When the tests were used in combination, students could be placed in one of five reading categories corresponding to Chall's (1983) reading development stages.

**3. The Reading Biographer.** To achieve the second goal—assessing the major factors in students' experiences related to their reading competence as high school seniors—a self-report questionnaire, referred to as The Reading Biographer, was developed (Hanson & Siegel, 1988a). The Reading Biographer sought to measure, retrospectively, the major events in a student's life that are linked to the development of reading competence. The factors are drawn from home, school, and extracurricular

activities (Siegel, 1987, 1990). Prototype items for The Reading Biographer were drawn from prior research, analytic methods, and professional reviews.

Each item was developed following carefully defined measurement procedures. These included:

1. use of only operationally defined responses designed to provide maximum discrimination between the behavioral categories of interest;
2. use of multiple items aggregated to form each variable (to enhance both validity and internal consistency reliability);
3. evaluation/revision of all items with high school seniors in actual school settings in a series of three pilot studies.

This measure, consisting of 38 items, provided information useful in formulating many of the independent and dependent variables, only a small sample of which are reported in this paper. The specific dependent (i.e., effects) variables used in this study that were measured by The Reading Biographer were (a) measures of students' current reading attitudes and interests, (b) amount of remedial instruction received, and (c) academic grades and tracking status. Some other independent variables it provides are measures of background information on social class, preschool experiences, schools attended, and gender/ethnic group membership.

The three assessment instruments (reading comprehension, vocabulary, and The Reading Biographer) were compiled into the 16-page Student Booklet, which took less than two class periods for all students to complete. These Student Booklets, along with specific directions to teachers for administering and returning them, were mailed to participating schools.

## Participants

Successful completion of this study required a strategy for locating the high school seniors across the United States who had entered specific elementary schools in 1973–74. Although school districts were provided the names of the elementary schools and the target year, locating students that still remain in a school system as high school seniors is a difficult and time-consuming task. Finding the time for students to take a test and complete a questionnaire is also considered a major inconvenience for high schools, which already have heavy testing requirements for their seniors. Allocating further testing time for this study was a considerable favor to ask of both the teachers and students.

To simplify both issues, the Student Booklet was used as a survey and administered to all the seniors in the high school, regardless of which elementary school(s) they attended. This approach was the most useful for the study design requirements; it provided

**Figure 1** Number of students and a description of each BRP Follow-up Study design comparison group

Kindergarten experience comparison groups Design 1		
(1) No kindergarten BRP/ no BRP elementary school <i>n</i> = 1,161	(2) No kindergarten BRP/ attended a BRP elementary school <i>n</i> = 871	(3) Kindergarten BRP/ attended a BRP elementary school <i>n</i> = 1,453
Kindergarten experience comparison groups Design 2		
(1) No kindergarten BRP <i>n</i> = 2,425	(2) Some kindergarten BRP <i>n</i> = 443	(3) Most/all kindergarten BRP <i>n</i> = 1,091

*Note:* Student is the unit of analysis and all students are included. High school is used as a control factor in the two-way ANOVA.

useful comparison group information without excessive additional costs.

Of course, this would not have been possible without the agreement and cooperation of the school districts; in addition to the study having the advantage of a large pool of districts to contact (420 districts), the districts exhibited considerable interest and professional goodwill. Most were very interested in knowing the results of the study to help guide current policy decisions about what to teach in kindergarten. Although space limits a discussion of the full set of considerations that went into deciding which districts to contact, a major concern was to obtain a representative sample of the original pool of districts and elementary schools to fulfill the basic design requirements. At the same time, the number of participants was limited by the funds available for materials, liaison, postage, data processing, and personnel costs.

In the end, a large and representative sample of 3,959 high school seniors was obtained. They were from a total of 23 BRP school districts and 1 non-BRP district in 10 different states representing all regions of the country. The majority of the elementary schools represented in the follow-up study had been designated as Title I eligible and served disadvantaged populations in 1973–74. Only 252 students, or 16% of the follow-up sample, were from non-Title I schools.

In the original kindergarten BRP inquiries, the overall sample of participating schools included a slightly larger sample of Title I schools (54%) and pupils (53%) than would normally be found in the general population. As compared to the original participants, the follow-up sample included an even larger percentage of both schools and students who were Title I eligible. However, obtaining a large number of Title I students was considered desirable since a major objective of the BRP

Follow-up Study was to assess the effects of the kindergarten reading effort on at-risk student populations.

### Design strategy

In addition to data from these high school seniors, the study contained documentation of each elementary school's effort and degree of success in teaching kindergarten children to read using the BRP during the 1973–74 school year (Hanson & Schutz, 1976). This baseline information could be used to categorize those students who attended kindergarten in the elementary schools that implemented the BRP into various comparison groups. However, the full follow-up study sample included many high school seniors who had not received instruction with the BRP. Specifically, there were those who (a) attended BRP elementary schools but not in kindergarten (i.e., entered in first grade or later) or (b) attended some other elementary school (usually in the same district) that did not implement the BRP in kindergarten. Collectively, these differences in the students' kindergarten and elementary experiences allowed for the formation of the different comparison groups for the study.

Specifically, the information was used to create two designs. In Design 1 there were three comparison groups defined as follows: (a) those who did not attend one of the BRP schools (i.e., assumed not to have received formal reading instruction in kindergarten); (b) those who began attending a BRP school in some grade above kindergarten (i.e., assumed to have had the same or similar elementary schooling experiences as the kindergarten BRP students but did not take part in the kindergarten BRP); and (c) those who began attending a BRP elementary school in kindergarten and, therefore, received the BRP instruction. Note that only those students in the third comparison group received any

kindergarten reading instruction. Comparisons between students in groups 2 and 3 would be especially interesting; with the exception of the kindergarten BRP, these two groups shared the same or similar schooling experiences and social class composition (see Design 1 in Figure 1).

The data for these three comparison groups were evaluated in two ways: first, by simply comparing differences on potential effects variables among the three groups using a one-way analysis of variance (ANOVA) design; and second, by analyzing a two-way, main effects ANOVA design with the high school attendance center as one factor and the three Program implementation categories as the second factor. Thus, variability in the effects variables due to differences among high schools (and also districts) could be evaluated independently of the kindergarten BRP experience variable.

Design 2 also comprised three comparison groups that more precisely examine the students in group 3 in Design 1. The three comparison groups in Design 2 are defined as follows: (a) those who did not receive kindergarten reading instruction (i.e., groups 1 and 2 in Design 1); (b) those who received some kindergarten reading instruction (i.e., from 1 to 7 units); and (c) those who received most or all of the kindergarten reading instruction (i.e., from 8 to 10 units). These groups provide a more rigorous comparison of effects because two of the groups received at least some kindergarten reading instruction. Here the expectation would be that group 1 would provide the baseline measure, group 2 would show some effects, and group 3 would show a larger difference on the effects variables. As with Design 1, the results were evaluated using both a one-way ANOVA and a two-way ANOVA, with the high school attendance center being the second main effects factor.

One final note regarding the two BRP Follow-up Study designs: because the designs use student-level data, other variables from the student questionnaire (The Reading Biographer) could be used to further examine the results for the effects variables. The specific biosocial variables used for this purpose were social class, gender, and ethnic background. Presentation of the results using these additional variables allows for closer examination of any differences found among the three comparison groups in each design. Also, they can indicate if any findings that emerge for the overall population can be generalized across these subpopulations.

## Results

The study was extremely fortunate in several ways. First, the major design requirement of obtaining a large representative sample of the original districts to provide

**Table 1** Description of BRP Follow-up Study participants by gender, ethnic/racial background, and social class

Factor	Category	Number	Percent
Gender	1 Female	1,888	48.0
	2 Male	2,048	52.0
Ethnic background	1 Asian	303	8.3
	2 Black	687	18.8
	3 Hispanic	110	3.0
	4 Native American	375	10.0
	5 All others	2,177	59.6
Social class	1 Lowest	428	11.0
	2 Lower middle	1,070	27.5
	3 Middle	1,430	36.8
	4 Upper middle	711	18.3
	5 Highest	245	6.3

data on the study instruments was met. The districts included students from a diverse set of elementary schools within the original districts, some of which had implemented the kindergarten BRP and others which had not. They also included a large number of students from elementary schools that had used most or all of the kindergarten BRP and others that had used only portions of it. Together, these subpopulations provided enough diversity in kindergarten experiences to form the post-hoc comparison groups with a large number of subjects within each group.

Second, the carefully defined and developed function-specific instruments (Hanson, McMorris, & Bailey, 1986; Hanson & Siegel, 1988a) worked perfectly. The variability they provided on these measures allowed a series of analyses to be carried out that collectively provide a strong basis for evaluating the long-term effects on students of beginning formal reading instruction in kindergarten.

The students from the 24 school districts and 10 states that provided follow-up data are described in Table 1 in terms of their gender, ethnic/racial categories, and socioeconomic background. As these data show, there were slightly more males than females, and the sample was quite diverse in terms of ethnic and social class background. In terms of ethnic background, minority groups make up about 40% of the total sample, with Black students accounting for about 20% of this total. The ethnic group percentages do show one unusual number—the higher than expected percentage (10%) of Native Americans in the sample. However, because of the study's emphasis on examining results for at-risk groups, their overrepresentation was considered a positive feature of the study.

**Table 2** The number and percent of students included in the full BRP Follow-up comparison categories

Kindergarten/elementary school category	Students		High schools	Districts
	Number	Percent		
1. No BRP, no BRP elementary school	1,549	39	41 of 43	23 of 24
2. No BRP, some BRP elementary school	867	22	42 of 43	23 of 24
3. BRP and BRP elementary school	1,534	39	41 of 43	23 of 24
4. No information	9	—	—	—
Total	3,959	100		

As might be anticipated, the social class background data (parents' occupations and education) show somewhat larger proportions in the lower middle to middle class categories than the population at large. This was consistent with their representation in the original kindergarten BRP studies where lower social class groups and minorities were also overrepresented. Overall, these descriptive statistics show the diversity required by the study in a large sample of high school seniors with only a few unusual percentages.

Another important concern is the extent to which the sample included students who received the kindergarten BRP reading instruction in 1972. Some data on this issue are given in Table 2, which shows the number of students, high schools, and districts providing follow-up data by the three kindergarten/elementary school group categories.

As Table 2 indicates, overall, a total of 3,959 students from 43 high schools in 24 districts participated in the follow-up study. The table also shows that of this total, 1,549, or 39% of the seniors, were in the first group (i.e., those who did not receive the kindergarten BRP and did not attend an elementary school that offered it). That means that they either attended another elementary school in the same district for kindergarten or transferred into the district after elementary school. In the second group there were 867 (22%) high school seniors. The seniors in this category did not receive the kindergarten BRP but did attend an elementary school that offered it (i.e., they enrolled in a BRP elementary school sometime after kindergarten). The third group contained 1,534 high school seniors (39%), who both attended a BRP elementary school and received the kindergarten BRP instruction. The fourth and final group included 9 seniors for whom no information was available on their elementary and kindergarten experience. The substantial numbers of high school seniors in the first three groups was consid-

ered a positive result; these three groups, as well as the subcategories within each, would provide the basis for forming the student groups used for comparison purposes in evaluating the effects variables.

In Table 3, the overall means and standard deviations are given for the effects variables, which are grouped into four sets. These data confirm that there is substantial variability in the full student population on all these variables. The mean values (which are also given) do not fall near either the maximum or the minimum values. Thus, they provide substantial differentiation between the seniors on the measures. This result was further confirmed by the full frequency distributions (not reproduced here), which showed substantial differences on all variables with few students attaining either the highest or lowest values.

Table 4 presents the simple correlations among the effects variables for all seniors in the follow-up sample. As expected, the four measures of reading skills (comprehension, vocabulary, reading stage, and illiteracy) are highly intercorrelated. Note that the variables in the second cluster (reading attitude, books read this year, and time spent reading) all have moderate correlations with one another but lower relationships to the four reading skills variables.

In the third cluster, remediation has a very low negative relationships to all variables, while both of the other two variables in this cluster, grades and attendance and high school academic track, have substantial positive relationships to each other and to all measures of reading competence. Finally, in the fourth cluster, the variables social class and parents' education are highly related and have quite low relationships to all other variables, which was not anticipated.

Collectively, the data in Table 4 show the expected pattern of high intercorrelations for the variables within the same clusters with more modest correlations be-

**Table 3** Descriptive statistics on the effects variables for the full BRP Follow-up Study sample

Effects variable	Mean	SD	N	Possible value	
				Minimum	Maximum
<i>Current reading competency</i>					
Comprehension	12.8	4.0	3,945	0	19
Vocabulary	17.1	3.7	3,958	0	24
Reading stage	2.7	1.0	3,943	0	1
Illiteracy	.15	.36	3,943	0	1
<i>Reading behaviors</i>					
Attitude	4.3	1.3	3,796	2	6
Books read this year	2.4	1.1	3,959	0	4
Time spent reading	1.6	1.1	3,959	0	4
<i>Schooling history</i>					
Remediation	.4	1.1	3,828	0	9
Grades and attendance	5.3	1.6	3,039	2	8
High school academic track	2.3	.6	2,777	1	3
<i>Family background</i>					
Social class	2.8	1.1	3,884	1	5
Parents' education	5.9	2.6	3,789	1	12

**Table 4** Correlation among the BRP Follow-up Study effects variables (decimals omitted)

Effects variable	1	2	3	4	5	6	7	8	9	10	11	12
<i>Current reading competency</i>												
1 Comprehension												
2 Vocabulary	.64											
3 Reading stage	.82	.81										
4 Illiteracy	-.59	-.65	-.71									
<i>Reading behaviors</i>												
5 Attitude	.29	.31	.31	-.16								
6 Books read this year	.09	.12	.11	-.02	.49							
7 Time spent reading	.06	.09	.06	-.03	.38	.38						
<i>Schooling history</i>												
8 Remediation	-.09	.08	-.09	.06	-.12	-.15	-.16					
9 Grades and attendance	.41	.38	.41	-.23	.34	.17	.11	-.22				
10 High school academic track	.39	.38	.39	-.22	.28	.13	.06	-.51	.46			
<i>Family background</i>												
11 Social class	.19	.24	.24	-.14	.08	.02	.05	-.06	.18	.23		
12 Parents' education	.03	.04	.06	.01	.03	.03	.00	-.02	.09	.17	.70	

tween the variables in different clusters. Aside from differences due to scaling (e.g., illiteracy) only two major exceptions are observed. The first concerns remediation, which was expected to show high negative relationships to achievement, as well as to variables in the other clusters, but did not. Instead, it shows little or no relationship to any of the other variables in the matrix.

The second exception concerns the variables in the family background cluster. Here the expectation would be that the social status variables (social class and par-

ents' education) would be highly related not only to each other (which they are) but also to the reading competency measures. The latter did not emerge. Though the correlations are positive, they are much lower than expected. A possible reason for this result emerged later in the study, and this result will be discussed further.

*Design 1 results: Differences between kindergarten BRP students and students in other kindergarten programs*

The general analysis strategy was to evaluate each

**Table 5** Summary of the ANOVA results and the effects variable means for Design 1 groups based on the elementary school attended and kindergarten BRP instruction

Effects variables	Kindergarten BRP experience groups			ANOVA significance level
	(1) No BRP/ no BRP school ( <i>n</i> = 1516)	(2) No BRP/ BRP school ( <i>n</i> = 846)	(3) BRP/ BRP school ( <i>n</i> = 1409)	
<i>Current reading competency</i>				
Comprehension	12.6	12.7	13.2	.00
Vocabulary	17.0	16.9	17.4	.00
Reading stage	2.7	2.7	2.8	.00
Illiteracy	.17	.18	.11	.00
<i>Reading attitude and behavior</i>				
Attitude	4.3	4.3	4.4	.34
Books read this year	2.4	2.4	2.4	.58
Time spent reading	1.5	1.6	1.6	.01
<i>Schooling history</i>				
Remediation	.55	.28	.23	.00
Grades and attendance	5.2	5.2	5.4	.00
High school academic track	2.3	2.2	2.3	.24
<i>Family background</i>				
Social class	2.9	2.8	2.7	.00
Parents' education	2.7	2.5	2.4	.00

effects variable within each cluster within each of two study designs composed of the three predefined and logically related comparison groups discussed previously. Differences in effects were then determined by using standard ANOVA procedures for each variable in a cluster. Next, the patterns of significant differences that emerged for all variables were reviewed. The assumption was that if the subgroup results showed a clear and consistent pattern of effects, only then could the result be considered important.

In Design 1, three groups of students were compared: (a) those who did not attend a BRP elementary school, (b) those who entered and attended a BRP elementary school after kindergarten (i.e., first grade or later), and (c) those who participated in the kindergarten BRP and attended a BRP elementary school. A special concern was to find out whether the attainment levels for students who had received the kindergarten BRP (group 3) were comparable to those students who had clearly different kindergarten experiences but otherwise comparable schooling histories (i.e., group 2, who attended the same elementary school but not the same kindergarten). Also important was to compare the BRP students to those students who differed in both their kindergarten and other elementary schooling experiences (i.e., group 1, who attended other elementary schools).

Differences among these three comparison groups on the full set of the effects variables discussed earlier were evaluated using simple ANOVA procedures. A summary of these analyses is presented in Table 5.

These results show that significant mean differences on all four reading competency variables were found among the three comparison groups. Further, when one looks at the means for each group, a consistent source for these differences emerges; it is between the children who received kindergarten reading (group 3) as compared to the other children (i.e., those in groups 1 and 2). In other words, the students who received the BRP instruction in kindergarten attained clearly higher scores on all four measures of reading competency than either those who entered the BRP school after kindergarten or attended another elementary school that did not offer the kindergarten BRP. Note also that the mean scores of the students in groups 1 and 2 are nearly identical. This supports the contention that the differences observed were due to the kindergarten reading experience rather than to some other factor associated with their subsequent schooling.

Among the effects variables in the second cluster, reading attitude and behavior, only one difference emerged as significant. The students in both groups who attended the BRP schools reported spending more reading (in hours per week) outside of school. For the

**Table 6** Summary of two-way ANOVAs on each effects variable with Design 1 (i.e., the comparison among the three kinds of kindergarten and elementary schooling experiences) being one factor (F1) and the high school being the other (F2)

Effects variable	Design 1 (3 comparison groups)		×	High School ( <i>n</i> = 43)	
	<i>F</i>	Significance		<i>F</i>	Significance
<i>Current reading competency</i>					
Comprehension	6.2	.00		8.3	.00
Vocabulary	9.2	.00		11.3	.00
Reading stage	5.9	.00		10.0	.00
Illiteracy	13.5	.00		8.0	.00
<i>Reading attitude and behavior</i>					
Attitude	.60	.58		1.7	.00
Books read this year	1.8	.17		1.2	.23
Time spent reading	2.7	.08		2.0	.00
<i>Schooling history</i>					
Remediation	3.0	.05		1.6	.01
Grades and attendance	3.1	.04		2.5	.00
High school academic track	6.5	.00		5.6	.00
<i>Family background</i>					
Social class	.18	.19		7.2	.00
Parents' education	5.1	.00		6.2	.00

schooling history variables, two of the three were significantly different across groups. The two variables, measuring the amount of remedial instruction received and the grades and attendance pattern in school, respectively showed significant differences in the anticipated direction. Clear mean differences favoring those having the kindergarten BRP, as compared to the students in the other two groups, emerged on both these important variables of school success. The third variable, high school academic track, did not show a significant difference.

Finally, significant differences were found among the students in the three comparison groups in terms of the social class of their families and their parents' education. However, in both cases, these differences are in a direction opposite from that suggested by the other reading competence differences. Students who received the kindergarten reading instruction came from families with a lower social class background than those who did not. This is an astounding result but consistent with the fact that most districts that chose to implement the BRP in the 1970s chose to implement it in their poorest schools. In essence then, not only did the students who received formal reading instruction in kindergarten exhibit a clear pattern of (a) showing superior current reading skills, (b) having higher grades and better attendance in school, and (c) needing and receiving significantly less remedial instruction in both elementary and secondary school, but they were also from families with a significantly lower social class status and parent education as compared to those in the other two comparison groups.

This type of finding was considered extraordinary, so additional analyses were carried out to examine these data at a more detailed level. In a second analysis, the effects variables on students in each of the three Design 1 comparison groups were broken down by high school within a two-way, main effects, ANOVA design. The purpose of using the high school as an additional main effect was not necessarily to see if differences between them would emerge, since such differences were clearly expected; rather, it was to see if the differences on the dependent variables observed among the three groups in Design 1 would still be significant after controlling for the differences between districts and high schools. In essence, it was to further confirm that the differences observed were not due to some unusual sampling characteristic. A summary of these ANOVAs is presented in Table 6.

For each of the individual variables, the result of a two-way ANOVA is summarized with the three Design 1 comparison groups as the first factor and the student's high school as the second main effect. As anticipated, the data in Table 6 show that for nearly every variable, significant differences were found associated with high schools. The data also show that all the findings for the three comparison groups were confirmed for all variables except time spent reading. Particularly noteworthy is the fact that all reading competence measures continued to show significant effects across the three comparison categories, favoring the kindergarten BRP group. In addition, one effects variable, high school academic track, which did not show a significant difference when

**Table 7** Summary of the ANOVA results and the effects variable means for Design 2 groups based on the elementary school attended and amount of kindergarten BRP instruction received

Effects variable	Type of school: Level of BRP implementation	Non-BRP	BRP	BRP	ANOVA significance level
		None	1-7 units	8-10 units	
<i>Current reading competency</i>					
Comprehension		12.6	12.8	13.3	.00
Vocabulary		17.0	17.1	17.5	.00
Reading stage		2.7	2.7	2.8	.00
Illiteracy		1.17	.14	.10	.00
<i>Reading attitude and behavior</i>					
Attitude		4.3	4.4	4.4	.36
Books read this year		2.4	2.4	2.4	.59
Time spent reading		1.6	1.6	1.6	.35
<i>Schooling history</i>					
Remediation		.45	.26	.21	.00
Grades and attendance		5.2	5.3	5.4	.00
High school academic track		2.2	2.3	2.3	.11
<i>Family background</i>					
Social class		2.9	2.7	2.8	.01
Parents' education		2.6	2.5	2.4	.00

evaluated alone, showed a significant difference favoring the kindergarten BRP group in the two-way ANOVA. In other words, the students receiving the kindergarten reading instruction ended up being in higher academic tracks when between high school differences were removed.

An additional point about the data in Table 6 concerns the social class variable. Note that the differences are not significant for the comparison group factor, but they are significant for the high school factor. This is what would be expected since social class differences are usually associated with high school attendance centers. Finally, these results explain the earlier "lower-than-expected" correlation between social class and reading competence. The impact of the early reading experience on some of the lower social class students was to raise their reading levels as high school seniors and, thereby, reduce the reading competence-family background correlations. In summary, these analyses directly support the earlier proposition that early reading impacts the effects variables as expected. Further, the differences observed on the key effects variables emerge even when controlling for differences between high schools and districts.

*Design 2 results: Differences among student groups receiving none, some, and much/all kindergarten BRP*

Design 2 focuses on those students who received different amounts of kindergarten BRP instruction. This focus pursued the finding from the original BRP inquiries, which showed that the more kindergarten BRP

units completed in each classroom, the higher the students' reading skills were at the end of kindergarten. Accordingly, the first group includes all students who did not receive any kindergarten BRP instruction, regardless of where they attended elementary school. In the second group are all students who attended elementary schools implementing part of the BRP (between 1 and 7 units), and the third group includes those completing most or all of the Program (between 8 and 10 units). In this design, group 1 (None) can be viewed as a large, diverse baseline group composed of more than 2,400 students who, it was assumed, did not receive any kind of formal reading instruction in kindergarten. The second and third groups were those who received either some or all of the kindergarten BRP instruction. The expectation was that students in the second and third groups (which included all the kindergarten BRP students) would differ from those in group 1, which serves in this design as a baseline comparison group.

The results of this analysis, including differences in the effects variable means and one-way ANOVA summaries, are given in Table 7. These results show that the pattern of mean differences and ANOVA results for the four reading competence measures are significant and show the expected pattern of differences. Similarly, the important schooling history variables of remediation and grades and attendance follow this pattern. Only high school academic track and the reading attitude and behavior variable did not show significant differences in these analyses.

**Table 8** Summary of two-way ANOVAs on each effects variable with Design 2 (level of kindergarten BRP implementation) being one factor (F1) and the high school being the other (F2).

Effects variable	Design 2		High School	
	F1	Significance	F2	Significance
<i>Current reading competency</i>				
Comprehension	6.4	.00	8.2	.00
Vocabulary	8.9	.00	11.2	.00
Reading stage	6.2	.00	9.9	.00
Illiteracy	13.6	.00	8.0	.00
<i>Reading attitude and behavior</i>				
Attitude	.1	.89	2.4	.00
Books read this year	1.4	.25	1.2	.23
Time spent reading	1.9	.14	1.9	.00
<i>Schooling history</i>				
Remediation	2.9	.05	1.6	.01
Grades and attendance	3.4	.03	2.5	.00
High school academic track	5.7	.00	5.5	.00
<i>Family background</i>				
Social class	.76	.47	7.3	.00
Parents' education	3.7	.03	6.1	.00

By way of contrast, the social class and parents' education variables were significant, again with the kindergarten BRP students having the lowest mean values on the variables. Collectively, these results are consistent with those presented previously for Design 1 and show the major effects variables to be sensitive to differences in the amount of kindergarten reading instruction provided to students. The results also show that this instruction was effective enough to reverse the usually positive relationship between parental social class and achievement.

To examine the data further, the two-way ANOVA analysis was used with the high school as one factor and the three comparison groups as the other factor. The results are reported in Table 8. Here, the high school factor emerged as a significant main effect for all the effects variables but one. More noteworthy is the fact that all the effects variables that emerged as significant in the one-way ANOVA (Table 7) were also significant in this analysis. These included the four measures of reading competence and three variables measuring schooling history.

Collectively, these Design 2 results extended the earlier results provided in Design 1. They show that not only are there effects associated with students receiving formal reading instruction in kindergarten, but also with the student groups defined by the amount of kindergarten reading instruction. Hence, not only is kinder-

garten reading important, but the more reading instruction, the better. Put another way, the effects variables, measured 12 years after the students' kindergarten BRP experience, showed reliable differences corresponding to differences in the amount of reading instruction the students received in kindergarten.

### Kindergarten reading and adult literacy

One final set of analyses using the Design 1 and 2 groups was carried out to further examine the differences obtained among students within the three biosocial classifications (i.e., ethnic, gender, and social class groups). One-way ANOVAs were computed for each group and design on each dependent variable, and the means were plotted. Close examination of these data revealed that there was a remarkable degree of correspondence between the results obtained within each of the categories formed by the biosocial variables and those shown for the full sample. Within nearly every category of these biosocial factors, the means for the students who received the kindergarten BRP were clearly higher. No matter how the sample was dissected, the differences remained.

Because of space limitations, all these analyses are not reproduced for this report (see Hanson & Siegel, 1988b). However, one set of analyses using a broad measure of illiteracy is given. This illiteracy variable was derived from Chall's reading stage. The Chall reading stage

**Table 9** Percent of students in each kindergarten and elementary experience comparison group (Design 1) who can be classified as functionally illiterate as seniors in high school (i.e., reading at or below the 5th grade level)

Biosocial variable	(1) No BRP/ no BRP school	(2) No BRP/ BRP school	(3) BRP/ BRP school	Combined or composite percent
<i>Ethnic background</i>				
Asian	13	10	1	9
Black	44	39	28	37
Hispanic	35	31	26	31
Native American	20	22	18	20
Other	8	8	6	7
<i>Gender</i>				
Female	15	17	11	14
Male	18	18	11	15
<i>Social class</i>				
Low	21	25	14	19
Medium	18	16	11	15
High	7	10	5	7
Overall	17	18	11	15

variable placed each student into one of five categories by using the combination of their reading vocabulary and performance on two types of reading comprehension items (literal and inferential). To create the illiteracy variable used here, those students in the two lowest categories, corresponding to functioning at or below a standardized fifth-grade reading level (i.e., functionally illiterate for a high school senior), received a one (1). Thus, illiteracy simply indicates whether a student is in either of these two categories. When it is summarized for a student group (and multiplied by 100), the mean gives the percent of students in each group in these low categories; that is, the percent of students who are clearly classified as functionally illiterate as high school seniors.

The results of the analysis for this variable are presented in Table 9 and show the practical impact of kindergarten reading instruction on this one global measure of adult literacy. The entries in this table indicate the percentage of students in each comparison group who were reading at or below the fifth-grade level as seniors in high school; these data are given by ethnic, gender, and social class populations.

Perhaps more directly than any other analysis, these data show the practical effects associated with beginning formal reading instruction in kindergarten. The students who learned to read in kindergarten fared better in all groups than those who did not. In virtually every category the data show the same consistent pattern of results: the percentage of functionally illiterate students is lower for groups who participated in the

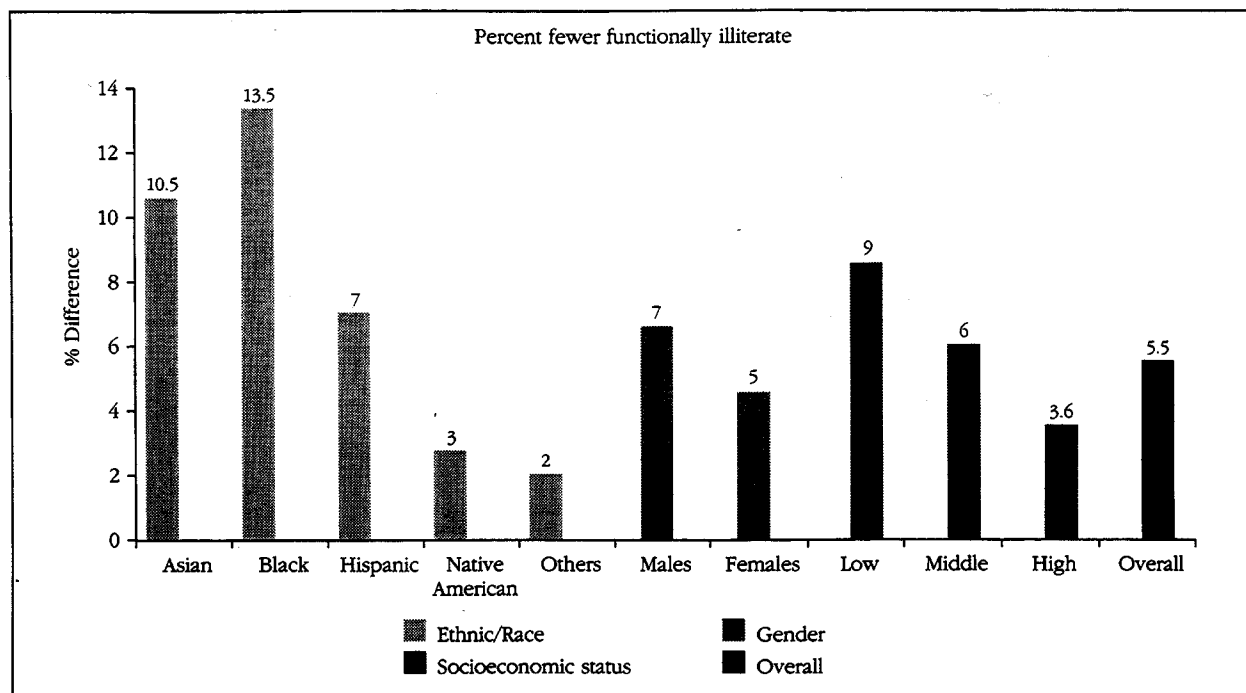
kindergarten reading program. Depending on the specific group, these differences varied from 2% to as much as 16%. Note, however, that the results for each group always favor the kindergarten readers. As Table 9 indicates, in the full sample, there were about 7% fewer of these poor readers among students who had the kindergarten reading program as compared to those who did not (i.e., 11% versus 18%). Put another way, there were about one-third fewer functionally illiterate high school seniors among those who received reading instruction in kindergarten as compared to those who did not.

Couple these results with the fact that the schools that taught reading in kindergarten included more students from disadvantaged backgrounds (lower social class), and these results are even more impressive. For these students to be even comparable to higher social class groups, who did not receive kindergarten reading instruction, would be a significant accomplishment. The fact that they read better overall and that there were substantially fewer students in the functionally illiterate group is quite surprising.

A conservative estimate of this effect for each group can be obtained by comparing each group receiving the BRP in terms of the percentage of low-level readers to the average for that group that did not receive the BRP (average of first two columns). These data are summarized in Figure 2 and show that these differences vary from 1% to 13.5%, depending on the specific group.

The largest percentages are in the groups that are typically from disadvantaged backgrounds and who usu-

**Figure 2** Illustration of the mean differences in the percent of “functionally illiterate” students favoring those students who participated in the kindergarten Beginning Reading Program (by ethnicity/race, gender, social class, and overall)



ally have high illiteracy rates as adults. These include those students in the ethnic minorities, males, and those from lower social class backgrounds. These are exactly the groups that it was anticipated would be most likely to show effects in the original proposal (see Hanson, 1984; Hanson & Siegel, 1988b). However, the data also show that the effects extend beyond these groups to students in virtually all groups. In terms of the proportion of poor readers within any given group, the kindergarten reading experience is valuable for advantaged as well as for disadvantaged populations in reducing the number of poor readers. The at-risk groups showed the greatest percentage of differences, but they also had the largest percentages of students in the lowest reading category.

### ***Summary, policy implications, and conclusion***

This study set out to examine a most controversial school policy issue: whether it is more advantageous to begin formal reading instruction in kindergarten or to delay it until first grade. To accomplish this task, this study used an extensive set of data gathered from a large national sample of schools that implemented the Begin-

ning Reading Program in their kindergarten classes during the 1973–74 school year. In 1986, 12 years later, a portion of the original BRP population, along with other students who had different kindergarten experiences, completed the Student Booklet, an instrument designed to measure current reading skills and other potential effects variables. A special concern of the research was to see if such instruction had any impact on at-risk students. Such students are always designated as having the lowest literacy rates as well as the highest school dropout rates.

At a conceptual level, the notion that learning to read earlier in school would impact subsequent reading skill makes excellent sense. It is supported by a number of motivational, learning, and social competence theories. These theories suggest that being able to gain competence in a critical skill such as reading should be very important to a child's self-confidence and subsequent learning in any area. It would also be consistent with current knowledge about socialization processes that occur in schools, which suggest that schools label and sort children very early into the “competent” and “less competent” categories and that these labels persist into adulthood (Shephard & Smith, 1989). In a skill area as central

to schooling success as reading, being labeled as competent during the early school years should have a lasting, positive impact. Further, recent studies have shown that early reading ability consistently correlates with later reading ability (e.g., Hays & Cangelosi, 1985; Juel et al., 1986; Williams & Silva, 1985). That is, students who are good readers at an early age (such as kindergarten and first grade) tend to be better readers in later grades.

On the other hand, persuasive rhetoric, based on old research studies and the philosophies of certain developmental theorists, still cautions against beginning formal reading instruction in kindergarten (e.g., Durkin, 1987b; Kuczen, 1987; Willis, 1993). This point of view is usually espoused by various early childhood professionals, both inside and outside of schools. Their major contention is that learning to read before the age of 6½ leads to initial frustration and anxiety that will result in long-term, negative effects for students.

To address this contention empirically required locating comparative data on a large, national sample of children who had been systematically taught to read in kindergarten, measuring their reading skills and attitudes as seniors in high school, and comparing them to other seniors who had different kindergarten but comparable elementary schooling experiences. In spite of a variety of obstacles, plus additional ones associated with the timing and funding of the project (Hanson & Siegel, 1988b), the study was successfully completed. More significantly, the results obtained refute the prevailing conventional wisdom about schools, programs, and particularly, current policy regarding teaching reading in kindergarten. In addition to showing a positive effect on such school-related factors such as grades, attendance, and the need for remedial instruction, this study demonstrated a remarkably clear and consistent pattern of increased reading competency for high school seniors as a result of receiving formal reading instruction in kindergarten.

How can such effects be understood? One immediate notion is that the results were not due to the kindergarten reading instruction alone, but rather to a combination of the instruction and the "halo" effect that having early reading skills provided. Students entering first grade with the ability to read are probably held in higher esteem than nonreaders and, thus, are labeled as "smart" by parents, teachers, and peers.

It should be noted that such convincing evidence favoring kindergarten reading instruction was totally unexpected. In the original grant proposal (Hanson, 1984), the expectation was that the early reading experience might show some positive, long-term impact for at-risk groups. Not only did the disadvantaged students benefit from the kindergarten reading instruction, but so did the advantaged groups.

Perhaps even more remarkable is the fact that, collectively, the high school seniors who participated in the kindergarten reading program had a lower social class rating than those who did not. Thus, in spite of an overall lower social class level, the students who received the kindergarten reading program still outperformed the higher social class students who did not. It is only in rare circumstances where a group with a lower social class rating outperforms one with a higher social class rating on a norm-referenced test of reading achievement. Further, the fact that these differences can be linked to an educational intervention makes them even more extraordinary.

### Policy implications

What then are the policy implications for schooling practices and future research in this area? One policy implication can be stated as follows: Schools having a well-developed reading program in kindergarten that is effectively and appropriately implemented should produce positive long-term benefits for all students, regardless of their background characteristics. Accordingly, support should be provided for schools to begin formal reading instruction at the kindergarten level in a manner consistent with the implementation methodology and instructional strategies of the Beginning Reading Program.

In this regard, it is important to recall that this Program required between 20 to 30 minutes of group instruction time per day for about 25 weeks. Thus, it did not preclude any of the other developmental, skill-building, or play activities that usually occur as part of the traditional kindergarten program. This should make reading instruction more appealing to those concerned about maintaining a "balanced" kindergarten curriculum for the "total" child.

The study also has important policy implications for future research on kindergarten reading, as well as for issues related to schooling effects and evaluation methodology. Naturally, it is important for future research to see if these results can be confirmed with other groups and within other contexts. Such replication and extension of study results are always important in behavioral research and especially when the findings are so at odds with prevailing notions.

This study also had some limitations that should be avoided, if possible, in the future. One of these is that many of the data, specifically from The Reading Biographer, are self-reported and were gathered at a single, later point in time. It is hoped that the considerable time and effort that went into designing and pilot testing the instruments, as well as the specific directions to the school districts involved, helped in this regard (e.g., carefully examining the students' ability to recall and to readi-

ly provide the information requested in The Reading Biographer and requiring that the Student Booklet be administered in a standardized and timely fashion). The ideal situation, of course, would have been to have a full set of longitudinal data gathered across *each* student's development; this was, however, never an option. Had the districts been able to provide longitudinal data on their schools' programs and the relevant experiences of their students, the level of detail and accuracy of this information would no doubt have been greatly enhanced.

Two other related limitations were the inability to account for those students who either (a) attended schools in which beginning reading instruction, other than the BRP, was part of the kindergarten curriculum or (b) dropped out of high school. Unfortunately, there was no way of obtaining this sort of information from the schools. Thus, no empirical data were available to answer the two questions that are typically asked about this study: "Might the results obtained have been influenced by students in non-BRP kindergarten comparison groups who were provided with kindergarten reading instruction other than with the BRP?" and "Were the results affected by a differential dropout rate across the comparison groups?"

The response in both cases is probably, "Yes." However, in this particular study, the effects would serve to bias the findings on the conservative side. For example, if those students in the non-BRP comparison groups, who were given other kinds of beginning reading instruction in kindergarten, had been identified and either removed from the study or placed in with one of the BRP comparison groups, the results probably would have shown *greater* effects for the BRP groups.

Likewise, in regard to dropouts, students who typically drop out of high school are those with lower overall achievement levels. Because all the non-BRP comparison groups had lower achievement levels than the BRP groups, one would have expected more dropouts (i.e., lower achieving students) in the non-BRP groups. Thus, the inclusion of dropouts in the study would probably have had the effect of *increasing*, not decreasing, the observed differences in achievement. Once again, only a full set of longitudinal data on all study participants would allow empirical confirmation of these results.

The current study is also simplistic when compared with the full set of possibilities that such a methodology would present with full longitudinal data. However, it also shows that the use of follow-up data can still be used to define schooling effects more precisely than before. Such data can show how schooling effects can be translated directly into cost information by comparing groups in terms of the actual amount of remedial education required by students in the three groups. Coupled

with information on the costs of remedial efforts, the effects of kindergarten instruction can be translated into a direct cost-benefit figure (e.g., Hanson & Siegel, 1989). In a similar fashion, such figures can be generated using the illiteracy information. For example, how much does it save society when the proportion of illiterate high school seniors is reduced by one-third?

Although it has been noted that schools are generally not able to provide precise data on their programs and other experiences of students who progress through them, the study also found that such a capability is emerging rapidly in some districts. Schools and other social service institutions are improving in their ability to handle information and maintain databases. Accordingly, better and more complete resources for policy information on schooling should emerge in the near future. Such information can then be used to routinely monitor the long-term effects of given practices and programs such as kindergarten reading, both within and across districts (Hanson & Siegel, 1991). This type of research, using new data sources and evaluation methods rather than either the follow-up approach employed here or, worse yet, one-shot studies such as those most often reported, should result in major breakthroughs in the ability of schools to validate program costs and effects. Programs, people, processes, and their interrelations are dynamic, not static, entities. Hence, policy researchers, school administrators, and other consumers and producers of educational policy need to monitor them constantly. This perspective and, in particular, its usefulness for understanding the costs and effects of school programs, has been developed elsewhere (Hanson, 1986).

## Conclusions

The major finding of this study, briefly stated, is: Students who learned to read in kindergarten were found to be superior in reading skills and all other educational indicators measured as seniors in high school. Further, this finding held up across districts and schools, as well as ethnic, gender, and social class groups. Also, there was absolutely no evidence of any negative effects from learning to read in kindergarten. Collectively, the results provide full support for the policy of teaching reading in kindergarten. Thus, any district with a policy that does not support kindergarten reading should be ready to present new and compelling reasons to explain why not—beyond the old and now refuted myth that it has long-term, adverse effects on students' reading skills, attitudes, and behaviors.

## REFERENCES

- ADAMS, M.J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: The MIT Press.

- AMES, L.B. (1950). Developmental trends in reading behavior. *Journal of Genetic Psychology*, 77, 291-312.
- ANDERSON, R., & FREEBODY, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.
- ARTHUR, G. (1925). A quantitative study of the results of grouping first grade classes according to mental age. *Journal of Educational Research*, 12, 173-185.
- BAKER, R.L., & SCHUTZ, R.E. (Eds.) (1971). *Instructional product development*. New York: Van Nostrand Reinhold.
- BALL, E., & BLACHMAN, B. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, 26, 49-66.
- BECK, I. (1973). *A longitudinal study of the reading achievement effects of formal reading instruction in the kindergarten: A summative and formative evaluation*. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.
- BECK, J. (1986). *How to raise a brighter child*. New York: Pocket Books.
- BENNETT, W.J. (1986). *First lessons: A report on elementary education in America*. Washington, DC: U.S. Department of Education.
- BLATCHFORD, P. (1987). Associations between pre-school reading-related skills and later reading achievement. *British Educational Research Journal*, 15(1), 15-23.
- BLOOM, B.S. (1967, December). *Toward a theory of testing which includes measurement-evaluation-assessment*. Proceedings of the Symposium on Problems in the Evaluation of Instruction (ESEIP Occasional Report No. 9). Los Angeles: University of California, Los Angeles.
- BORUCH, R.F., & PEARSON, R.W. (1985). *The comparative evaluation of longitudinal surveys*. New York: Social Science Research Council.
- BROACH, D. (1988, April). *Estimating vocabulary size: An example of function specific testing*. Paper presented at the annual meeting of the National Council for Measurement in Education, New Orleans, LA.
- BRZIEJSKI, J.E. (1967). *Summary report of the effectiveness of teaching reading in kindergarten*. Denver, CO: Denver Public Schools.
- BURTON, N.W., & JONES, L.V. (1982). Recent trends in achievement levels of black and white youth. *Educational Researcher*, 11(4), 10-15.
- CAMPBELL, D., & STANLEY, J. (1966). *Experimental and quasi-experimental design for research*. Boston, MA: Houghton Mifflin.
- CARNINE, D., CARNINE, L., KARP, J., & WEISBERG, P. (1988). Kindergarten for economically disadvantaged children: The direct instruction component. In C. Warger (Ed.), *A resource guide to public school early childhood programs* (pp. 73-98). Alexandria, VA: Association for Supervision and Curriculum Development.
- CARROLL, J., DAVIES, P., & RICHMAN, B. (1971). *The American Heritage word frequency book*. New York: American Heritage.
- CENTER FOR STATISTICS. (1986). *Annotated bibliography of studies using data from high school and beyond: A national longitudinal study for the 1980's* (Contractor Report No. 300-82-0273). Washington, DC: U.S. Government Printing Office.
- CHALL, J. (1983). *Stages of reading development*. New York: McGraw-Hill.
- CONSORTIUM FOR LONGITUDINAL STUDIES. (1979). *Lasting effects after preschool summary report: Further analysis of longitudinal studies* [DHEW publication No. (OHDS) 79-30179]. Washington, DC: U.S. Government Printing Office.
- DOHERTY, W., & HANSON, R. (1974). *Development of BRP end-of-year assessment for the 1972-73 ESEA Title III Quality Assurance Program* (Technical Memorandum 5-70-01). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development.
- DOWNING, J. (1963). Is a mental age of six essential for reading readiness? *Educational Research*, 6, 16-28.
- DURKIN, D. (1966). *Children who read early*. New York: Teachers College Press.
- DURKIN, D. (1968). When should children begin to read? In H.M. Robinson (Ed.), *Innovation and change in reading instruction, The 67th yearbook of the National Society of Education, Pt II* (pp. 30-71). Chicago, IL: University of Chicago Press.
- DURKIN, D. (1987a). Testing in the kindergarten. Special issue: The state of assessment in reading. *The Reading Teacher*, 40, 766-770.
- DURKIN, D. (1987b). *Teaching young children to read*. Boston, MA: Allyn & Bacon.
- DURKIN, D. (1990, September). Reading instruction in kindergarten: A look at some issues through the lens of new basal reader materials. *Early Childhood Research Quarterly*, 5(3), 299-313.
- ENGLEMAN, S., & ENGLEMAN, T. (1966). *Give your child a superior mind*. New York: Simon & Schuster.
- FOLLETTIE, J. (1971, March). *Word attack model*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 109 620)
- GATES, A.I., BOND, G.L., & RUSSELL, D.H. (1936, May). Reading readiness. *Teachers College Record*, 37, 679-685.
- GATES, A.I., BOND, G.L., & RUSSELL, D.H. (1939). *Methods of reading readiness*. New York: Bureau of Publications, Teachers College, Columbia University.
- GERSTEN, R., & GEORGE, N. (1990). Teaching reading and mathematics to at-risk students in kindergarten: What we have learned from field research. In C. Seefeldt (Ed.), *Continuing issues in early childhood* (pp. 245-259). Columbus, OH: Merrill.
- GESELL, A. (1954). The ontogenesis of infant behavior. In L. Carmichael (Ed.), *Manual of child psychology* (pp. 335-373). New York: Wiley.
- GOLDSTEIN, H. (1979). *The design and analysis of longitudinal studies: Their role in the measurement of change*. New York: Academic.
- GRAY, W.S. (1937). *The teaching of reading: A second report, 36th Yearbook of the NSSE, Part I*. Bloomington, IL: Public School Publishing Co.
- HALL, R.V. (1963). Does entrance age affect achievement? *Elementary School Journal*, 63(7), 391-396.
- HANSON, R. (1979). *Programmatic research and development and the dissemination/utilization of educational product systems. Proceedings of the Conference on Research on Educational Dissemination and Utilization*. San Francisco, CA: Far West Regional Laboratory.
- HANSON, R. (1984). *The long-term effects on disadvantaged students of learning to read in kindergarten*. Proposal to the U.S. Department of Education. Tulsa, OK: University of Tulsa.
- HANSON, R. (1986). Program evaluation and improvement based on cost-effectiveness analyses. *Studies in Educational Evaluation*, 12, 89-108.
- HANSON, R., & BAILEY, J. (1971). *The development of the 1970-71 quality assurance procedures* (Technical Memorandum 5-71-54). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development.
- HANSON, R., & BERGER, R. (1971). *Quality assurance in large scale installation of criterion-referenced instructional programs* (Professional Paper No. 18). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 059 547)
- HANSON, R., KAPLAN, S., & YAMAN, N. (1971). *Comparison of the 1969 and 1970 tryouts of the Summer Reading Program* (Technical Memorandum 5-71-07). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development.
- HANSON, R., & KELBLE, E. (1985, January). *The relationship of home*

- and preschool learning to children's language development and progress in kindergarten. Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas, TX.
- HANSON, R., LEHMAN, J., & BAILEY, J. (1981). *A six year longitudinal study of an instructional innovation*. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.
- HANSON, R., MCMORRIS, R., & BAILEY, J. (1986). Differences in instructional sensitivity between item formats and between achievement tests items. *Journal of Educational Measurement*, 23(1), 1-12.
- HANSON, R., & RESTA, P. (1970). *First-year Communication Skills Program: Interim report* (Technical Memorandum 5-76-01). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 057 993)
- HANSON, R., & ROSS, S. (1973). *Relationship of instructional time allocation to program pacing* (Technical Memorandum 5-73-02). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development.
- HANSON, R., & SCHUTZ, R.E. (1975). *The effects of Programmatic R&D on schooling and the effects of schooling on students: Lessons from the first year installation of the SWRL/Ginn Kindergarten Program* (Technical Report 53). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 126 018)
- HANSON, R., & SCHUTZ, R.E. (1976). *Instructional Product implementation and schooling effects: Lessons from the second year installation of the SWRL/Ginn Kindergarten Program* (Technical Report 56). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 124 925)
- HANSON, R., & SCHUTZ, R.E. (1978). A new look at schooling effects from programmatic research and development. In D. Mann (Ed.), *Making change happen?* (pp. 120-149). New York: Columbia University Press.
- HANSON, R., & SCHUTZ, R.E. (1986). A comparison of methods for measuring achievement in basic skills program evaluation. *Educational Evaluation and Policy Analysis*, 8(1), 1-13.
- HANSON, R., SCHUTZ, R.E., & BAILEY, J. (1977). *Program-fair evaluation of instructional programs: Initial results of the kindergarten reading readiness inquiry* (Tech. Rep. 57). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 165 519)
- HANSON, R., & SIEGEL, D.F. (1988a). *Design and development of The Reading Biographer: An application of function specific assessment*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- HANSON, R., & SIEGEL, D.F. (1988b). *The effects on high school seniors of being taught to read in kindergarten* (Tech. Rep. 1). Garden Grove, CA: Hanson Research Systems. (ERIC Document Reproduction No. ED 341 967)
- HANSON, R., & SIEGEL, D.F. (1989, April). *Cost-benefit analysis of kindergarten reading instruction*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- HANSON, R., & SIEGEL, D.F. (1991, April). *Formative, summative, and confirmative: The three phases of program evaluation*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- HAYS, B.L., & CANGELOSI, J.S. (1985, Winter). Kindergarten reading instruction and students' reading achievement trends. *Reading Improvement*, 22(4), 305-315.
- HYLTON, J., & QUELLMALZ, E. (1970, December). *The development of an observer training video tape sequence*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. EJ 031 613)
- ILG, F., & AMES, L.B. (1964). *School readiness*. New York: HarperCollins.
- INHENDER, B. (1962). Some aspects of Piaget's generic approach to cognition: Thought in the young child. *Monographs of the Society for Research in Child Development*, 27(2).
- JONES, H., & BAYLEY, N. (1941). The Berkeley growth study. *Child Development*, 12, 167-173.
- JUEL, C., GRIFFITH, P., & GOUGH, P. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology*, 78(4), 243-225.
- KOEHLER, J., Jr. (1971, July). *The effects of serial pattern discrimination and mixed word identification training on sight word acquisition and retention*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 110 916)
- KOEHLER, J., Jr. (1972). *Novel word decoding in kindergartners as a function of decoding strategy and segmentation-blending training form*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 067 647)
- KUCZEN, B. (1987). *Childhood stress: How to raise a healthier, happier child*. New York: Dell.
- LOTT, D. (1969, July). *Visual word recognition: Its implications for reading research and instruction*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 035 515)
- LUNDBERG, I., FROST, J., & PETERSON, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, 23, 263-284.
- MASON, J. (1984). Early reading from a developmental perspective. In P.D. Pearson (Ed.), *Handbook of reading research* (pp. 505-543). New York: Longman.
- MEEK, M., & ARMSTRONG, S. (1983). *Achieving literacy: Longitudinal studies of adolescents learning to read*. Boston, MA: Routledge & K. Paul.
- MORPHETT, M., & WASHBURNE, C. (1931). When should children begin to read? *Elementary School Journal*, 31, 469-503.
- NIEDERMEYER, F.C. (1971). *Development of teacher-administered tests for the SWRL Reading Program*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 109 608)
- NIEDERMEYER, F.C. (1972, September). *Observational procedures for classroom tryout of instructional materials and procedures*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 106 886)
- NIEDERMEYER, F.C., & ELLIS, P. (1970, May). *The development of a tutorial program for kindergarten reading instruction*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 057 994)
- NIEDERMEYER, F.C., & FISCHER, K. (1972, July). *Prototype development of an instructional improvement kit for the Beginning Reading Program*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 108 138)
- NIEDERMEYER, F.C., & GIGUERE, C. (1972, July). *Criterion-referenced tests for the SWRL Kindergarten Curriculum*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 114 404)
- OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT, CEN-

- TER FOR STATISTICS. (1986). *Annotated bibliography of studies using data from high school and beyond: A national longitudinal study for the 1980's* (Contractor Report No. 300-82-0273). Washington, DC: U.S. Government Printing Office.
- PRESCOTT, C.A. (1955). Sex differences in Metropolitan Readiness Test results. *Journal of Educational Research*, 48, 605-619.
- QUELLMALZ, E. (1973, Summer). Functions of status testing in the development of an instructional program. *Improving Human Performance*, 2. (ERIC Document Reproduction No. EJ 087 866)
- RESTA, P.E., & NIEDERMEYER, F.C. (1970). *Behavioral analysis of First-year Communication Skills teacher administered instruction*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 037 412)
- RUDOLPH, M., & COHEN, D. (1984). *Kindergarten and early schooling* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- SABERS, D. (1985). *Curriculum referenced tests of mastery: Technical report*. Columbus, OH: Charles Merrill.
- SCHUTZ, R.E. (1970). *Programmatic instructional development*. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis. (ERIC Document Reproduction No. ED 055 432)
- SHEPHARD, L., & SMITH, M.L. (1989). *Flunking grades: Research and policies on retention*. New York: Falmer.
- SIEGEL, D.F. (1987). *Identification and validation of process factors related to the reading achievement of high school seniors: A follow-up study*. Unpublished doctoral dissertation, The University of Tulsa, Tulsa, OK.
- SIEGEL, D.F. (1990, July/August). The literacy press: A process model for reading achievement. *Journal of Educational Research*, 83(6), 336-347.
- SIEGEL, D.F., & HANSON, R. (1991). Kindergarten educational policies: Separating myth from reality. *Early Education and Development*, 2(1), 5-31.
- SMITH, F. (1967). *Understanding reading: A psycholinguistic analysis of reading and learning to read* (3rd ed.). New York: Holt, Rinehart and Winston.
- SMITH, F. (1969a). Familiarity of configuration vs. discriminability of features in the visual identification of words. *Psychonomic Science*, 14(6), 261-263.
- SMITH, F. (1969b). The use of featural dependencies across letters in the visual identification of words. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 215-218.
- SMITH, F., LOTT, D., & CRONNELL, B. (1969). The effect of type size and case alternation on word identification. *American Journal of Psychology*, 82(2), 248-253.
- SMITH, M.L. (1976). *The effect of preschool experiences on achievement in reading*. Paper presented at the annual meeting of the College Reading Association, Miami. (ERIC Document Reproduction No. 130 256)
- SONTAG, L. BAKER, C., & NELSON, V. (1958). Mental growth and personality development: A longitudinal study. *Monographs of the Society for Research in Child Development*, 23, 1-143.
- SOUTHWEST REGIONAL LABORATORY (SWRL) FOR EDUCATIONAL RESEARCH AND DEVELOPMENT. (1972). *SWRL/Ginn and Company Beginning Reading Program*. Lexington, MA: Ginn and Company.
- STENNER, A., SMITH, M., III, & BURDICK, D. (1983). Toward a theory of construct definition. *Journal of Educational Measurement*, 20, 305-316.
- STEWART, I.S. (May-June 1985). Kindergarten reading curriculum: Reading abilities, not reading readiness. *Childhood Education*, 61(5), 356-360.
- SULLIVAN, H. (1968, May). *Improving learner achievement through evaluation by objectives*. Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development. (ERIC Document Reproduction No. ED 065 505)
- SULLIVAN, H. (1973). *Considerations in selecting and using instructional objectives*. Los Angeles, CA: Instructional Objectives Exchange. (ERIC Document Reproduction No. ED 177)
- SULLIVAN, H., & LABEAUNE, C. (1971, February). Parents: Summer reading teachers. *Elementary School Journal*, 71. (ERIC Document Reproduction No. EJ 032 510)
- TEALE, W. (1980). *Early reading: An annotated bibliography*. Newark, DE: International Reading Association. (ERIC Document Reproduction No. 190 995)
- TEALE, W., & SULZBY, E. (1987). Literacy acquisition in early childhood. In D. Wagner (Ed.), *The future of literacy in a changing world Vol. I* (pp. 120-129). New York: Pergamon.
- TERMAN, L., & ODEN, M.H. (1947). *Genetic studies of genius IV: The gifted child grows up*. Stanford, CA: Stanford University Press.
- TERMAN, L., & ODEN, M.H. (1959). *Genetic studies of genius V: The gifted group at mid-life: Thirty-five years follow-up of the superior child*. Stanford, CA: Stanford University Press.
- TOBIN, A.W., & PIKULSKI, J.J. (1987). A longitudinal study of reading achievement of early and non-early readers through sixth grade. *National Reading Conference Yearbook*, 37, 39-58.
- TYLER, F. (1964). Issues related to readiness to learn. In E.R. Hilgard (Ed.), *National Society for the Study of Education, 63rd yearbook* (pp. 210-239). Chicago, IL: University of Chicago Press.
- ULRICH, R., LOUISELL, S.E., & WOLFE, M. (1971). The learning village: A behavioral approach to early education. *Educational Technology*, 11, 32-45.
- WASHBURN, C. (1936, February). Ripeness. *Progressive Education*, 13, 125-130.
- WEISBERG, P. (1987). *Accelerating the reading achievement of poverty-level preschoolers by extensive training in synthetic phonics*. Tuscaloosa, AL: University of Alabama.
- WHITE, S., ALEXANDER, P., WILSON, V., & FUQUA, D. (1985). *Young children's ability to perform geometric analogy problems*. Paper presented at the annual meeting of Southwest Educational Research Association, Austin, TX.
- WILLIAMS, S.M., & SILVA, P.A. (1985). Some factors associated with reading ability: A longitudinal study. *Educational Research*, 27(3), 159-168.
- WILLIS, S. (1993, November). Teaching young children: Educators seek "Developmental Appropriateness." *ASCD Curriculum Update*, 1-8. (Stock No. 611-9311)
- WOLF, J.M., & KESSLER, A.K. (1987). *Entrance to kindergarten: What is the best age?* (ERS Monograph). Arlington, VA: Educational Research Services.

Received August 27, 1993

Final revision received October 16, 1994

Accepted November 23, 1994

#### AUTHOR NOTES

This follow-up study represents the culmination of research activity that began at the Southwest Regional Laboratory (SWRL) in 1969 and concluded at The University of Tulsa in 1987. Complete information on the study procedures and data sources can be found in the Final Report to U.S. Department of Education for Grant No. 8410091. The grant provided the support for this project through the effort of Dr. Terrell Bell, Secretary of Education at that time. It should be noted that this article provides substantial additional analyses of the data not

available for the Final Report. The authors also take full responsibility for the results and interpretations of them.

Special thanks are due to Dr. Bell for authorizing the grant support and to Dr. Thomas Enderlein, who administered the grant. Thanks are also due to The University of Tulsa administration and

staff for all their support and assistance during the study. Finally, a most appreciative and sincere thanks are in order to all the districts, staff, and students who took the time to provide the study information. Without their help and professional goodwill, the study would have been impossible to complete.