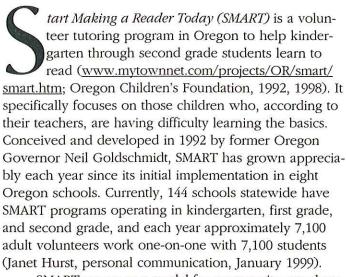# When less may be more: A 2-year longitudinal evaluation of a volunteer tutoring program requiring minimal training

Scott Baker

Russell Gersten

Thomas Keating

Eugene Research Institute, Eugene, Oregon, USA

Start Making a Reader Today (SMART) is a volunteer tutoring program in Oregon to help kindergarten through second grade students learn to read (www.mytownnet.com/projects/OR/smart/smart.htm; Oregon Children's Foundation, 1992, 1998). It specifically focuses on those children who, according to their teachers, are having difficulty learning the basics. Conceived and developed in 1992 by former Oregon Governor Neil Goldschmidt, SMART has grown appreciably each year since its initial implementation in eight Oregon schools. Currently, 144 schools statewide have SMART programs operating in kindergarten, first grade, and second grade, and each year approximately 7,100 adult volunteers work one-on-one with 7,100 students (Janet Hurst, personal communication, January 1999).

SMART serves as a model for community members who wish to be more actively and positively involved with their local schools. In a series of town meetings throughout Oregon, Governor Goldschmidt consistently encountered a deep sense of disconnection between adults and the schools their children attended. What made SMART unique was that, from its inception, it attempted to reconnect communities and schools by asserting two basic premises. The first underlies virtually all tutoring programs (Juel, 1996; Shanahan, 1998; Wasik, 1998): Adults can make a vital difference in the lives of young students by spending time reading to them and teaching them to read. Even the best instructional environments for first graders in a public school setting, with one expert teacher responsible for teaching 20–30 students, cannot match the educational intensity of a one-to-one interaction. When an adult sits down with a child and shares in the pleasures of reading and then helps the child build literacy skills, progress accelerates.

SMART's second basic premise is that adults receive benefits as great as the students from the experience of meaningful involvement in the life of a young child (Oregon Children's Foundation, 1992, 1998). Not only would children with special needs become better readers, but adults, actively involved in the education of those children, would gain a better understanding of school life and, as they watched their students become better readers, would emerge from the experience with a sense of real accomplishment.

From the beginning, SMART was designed for rapid, wide-scale implementation. A major concern was that the program be low cost and feasible to implement and expand. Although SMART is in many ways similar to

When less may be more: A two-year longitudinal evaluation of a volunteer tutoring program requiring minimal training

The purpose of the current study was to evaluate the effects on reading achievement of a low-cost, widely implemented volunteer reading program that has been expanding rapidly throughout the state of Oregon. Eighty-four beginning first grade students at risk of reading difficulties were randomly assigned to experimental and comparison groups. Adult volunteers tutored students in the experimental group in 30-minute sessions two times per week in first and second grade. At the end of grades 1 and 2, students were administered a number of standardized reading measures, including measures of individual word reading, reading comprehension, word comprehension, and reading fluency. Analyses revealed that students in the experimental group made greater growth on a word identification measure than students in the comparison condition; they also made more growth than a group of average-achieving students who were from the same classrooms as the students in the experimental and comparison groups. Students in the experimental group also scored higher than students in the comparison condition on measures of reading fluency and word comprehension at the end of second grade. Differences were not statistically significant on passage comprehension. Findings are discussed in the context of the reading achievement effects that other adult volunteer reading programs have attained. We suggest that in establishing adult volunteer reading programs it is important to consider how to balance the intensity of training reading volunteers to achieve measurable impact on reading achievement with real world realities of the volunteer tutoring experience and goals for the extensiveness of implementation.

Cuando menos puede ser más: Una evaluación longitudinal de dos años de un programa de tutorías voluntarias que requería un mínimo entrenamiento

El propósito del presente estudio fue evaluar los efectos de un programa de tutorías, de bajo costo y muy difundido que se ha expandido rápidamente en el estado de Oregon, sobre los progresos en lectura. Ochenta y cuatro estudiantes de primer grado considerados en riesgo de fracaso en lectura fueron asignados al azar a grupos experimental y de control. Los voluntarios adultos realizaron tutorías con los estudiantes del grupo experimental en sesiones de 30 minutos, dos veces por semana en primero y segundo grado. Al final de primero y segundo grado se administró a los estudiantes un conjunto de medidas estandarizadas de lectura, que incluían lectura de palabras individuales, comprensión lectora, comprensión de palabras y fluidez en lectura. Los análisis revelaron que los estudiantes del grupo experimental progresaron más que los estudiantes del grupo de control en una medida de identificación de palabras; asimismo progresaron más que un grupo de estudiantes promedio provenientes de las mismas aulas que los estudiantes de los grupos experimental y de control. Los estudiantes del grupo experimental también se desempeñaron mejor que los del grupo de control en medidas de fluidez en lectura y comprensión de palabras al final del segundo grado. En comprensión de textos las diferencias no fueron estadísticamente significativas. Los hallazgos se discuten en el contexto de los efectos en el progreso en lectura obtenidos por otros programas de tutores voluntarios adultos. Sugerimos que al establecer programas de tutorías con voluntarios adultos es importante considerar la forma de equilibrar la intensidad del entrenamiento de los voluntarios para obtener un impacto mensurable en el progreso en lectura con la realidad concreta de la experiencia de las tutorías voluntarias y objetivos para hacer extensiva la implementación.

Wenn wenig mehr sein kann: Eine zweijährige Langzeitbewertung eines Nachhilfeprogramms mit freiwilligen Helfern bei minimalen Trainingsanforderungen

Der Sinn der laufenden Studie bestand in der Bewertung der Auswirkungen auf die Leseleistungsverbesserungen aufgrund eines kostensparenden, von freiwilligen Helfern weitläufig angewandten Leseprogrammes, das in schneller Folge überall im US-Staat Oregon verbreitet wurde. Vierundachtzig Schüleranfänger der ersten Klasse mit Risiken von Leseschwierigkeiten wurden wahllos den Experimentier- und Vergleichsgruppen zugeteilt. Erwachsene als freiwillige Helfer gaben den Schülern in der Experimentiergruppe jeweils für 30 Minuten zweimal die Woche Nachhilfeunterricht in den ersten und zweiten Klassen. Am Ende der 1. und 2. Klasse wurde den Schülern eine Anzahl standardisierter Lesebewertungsmaßnahmen auferlegt, einschließlich der Bemessung individueller Lesefertigkeiten von Worten, dem Leseverständnis, dem Wortverständnis, und der flüssigen Lesebeherrschung. Analysen ergaben, daß Schüler in der Experimentiergruppe größere Fortschritte bei der Bemessung der Worterkennung machten als Schüler unter Vergleichskonditionen; auch machten sie mehr Fortschritte als eine Gruppe von Durchschnittsschülern, die aus dem gleichen Klassenraum kamen wie Schüler in den Experimentier- und Vergleichsgruppen. Die Schüler der Experimentiergruppe erzielten auch höhere Noten als Schüler unter den Vergleichsbedingungen bei der Bemessung der Leseflüssigkeit und dem Wortverständnis am Ende der zweiten Klasse. Im Verständnis von Passagen bzw. Absätzen waren solche Differenzen statistisch nicht signifikant. Die Ergebnisse werden im Zusammenhang von Leseleistungseffekten diskutiert, die andere freiwillige Erwachsenenhelfer mit ihren Leseprogrammen erzielt haben. Wir schlagen vor, daß bei der Schaffung von freiwilligen Lesenachhilfeprogrammen durch Erwachsene es wichtig ist zu berücksichtigen, wie die Intensität der Leseschulung freiwilliger Helfer ausgewogen gestaltet wird, um einen meßbaren Einfluß auf die Leseleistung unter wirklichkeitsnahen Bedingungen an Erfahrungen und Zielen von freiwilligen Helfern zur weitgehendsten Verbreitung zu erreichen.

other tutoring programs such as those described by Shanahan (1998) and Wasik (1998), several important differences set it apart. Except for the use of Americorps volunteers as coordinators at some schools, SMART is entirely a private-sector enterprise. From the outset, the Oregon business community has played a large role in supporting the program by funding operating costs and paying for books, as well as by actively encouraging their employees to become reading volunteers and by facilitating their involvement as part of their paid employment. The individuals who conceived SMART believed that keeping government support to a minimum and relying primarily on support from local business and community organizations would increase its chances for survival.

SMART also differs from other programs in its approach to volunteer training and its minimal demands on teachers. Volunteer training is brief and focuses as much on the logistics of tutoring (e.g., where books are located, public school safety) as it does on reading instruction techniques. Tutors are provided with a broad framework to use during sessions, rather than specific techniques. SMART's approach to training contrasts sharply with the rather extensive training many educators suggest volunteers need to effectively tutor students in reading (e.g., Juel, 1994; Roller, 1998).

SMART's approach developed in part from the expectation that volunteer turnover from year to year was likely to be high and therefore intensive training would not be cost-effective. At approximately 50% per year, the turnover rate of SMART volunteers has proved to be an important training issue (Hurst, personal communication, January 12, 1999). There was also a sense that it would be far easier to recruit tutors to begin with if it were clear they were not expected to either know or acquire specialized instructional skills. Instead, the emphasis of recruiting was on asking volunteers, frequently under the auspices of their employers, to simply show up twice a week to read with students.

Similarly, to implement the program easily on a large scale, SMART was intentionally designed to place minimal demands on teachers whose students were being tutored and on coordinators supervising tutors and providing ongoing training. Parents and teachers were told that SMART's purpose was to supplement the daily reading instruction provided by the classroom teacher. There was no attempt to coordinate this supplemental program with the core reading program of each school or each classroom. This was due in part to the complex logistics that would have been necessary for such an arrangement, and in part because of the desire to keep implementation simple. Teachers were asked only to identify the students they felt needed extra support in reading.

What is the relative impact of different approaches to tutoring?

Given the high turnover rate typical of most volunteer programs and the extraordinary cost of training tutors, any program that is self-sufficient and serves a large number of students bears closer scrutiny. Compared with the volunteer reading programs evaluated in a recent review by Wasik (1998), SMART is low cost, serves a large number of students in predominantly low-income schools, and requires minimal training. Little systematically collected information is available, however, on the impact of volunteer reading programs on reading ability. The purpose of the current study was to evaluate the impact of the SMART tutoring program on the reading abilities of students deemed at risk for failure.

Only three previous studies of the impact of volunteer reading programs have been conducted that have used controlled experimental-comparison group designs (Wasik, 1998). These three programs—the Howard Street Tutoring Program (Morris, Shaw, & Perney, 1990), the Intergenerational Tutoring Program (American Academy of Arts and Sciences and Boston Partners in Education, 1999), and the School Volunteer Development Project (U.S. Department of Education, 1979, as cited in Wasik, 1998)—are drastically different than SMART in several ways. In these programs, training of tutors is much more intensive, lengthy, and highly structured. These programs also had an impact on a far smaller number of students, with less than 150 students being served at the time they were evaluated (American Academy of Arts and Sciences & Boston Partners in Education, 1999; Morris et al., 1990; U.S. Department of Education, 1981).

At the time we conducted our evaluation of SMART, over 7,000 students were being served. We believed it was crucial to understand the impact on reading achievement of programs like SMART, which provide minimal training and rely primarily on the judgment and instincts of literate adults to tutor struggling readers. We wished to examine whether these effects were comparable to the effects achieved by tutoring programs that are more intense and costly to implement. This information could be critical for policy makers and others who want to balance program effectiveness and breadth, trying to meaningfully serve as many students as possible who need assistance in the primary grades.

A secondary purpose of the study was to determine the impact of SMART on the referral and placement of students with reading problems in special education. A major goal of contemporary educational policy at both the federal and state or local levels (e.g., California Department of Education, 1998; Texas Reading Initiative, 1997) is to provide intensive beginning reading instruction—with any necessary support—in the primary grades to reduce un-

necessary special education placement in later years. To our knowledge, this study is the only one conducted that has examined the impact of a volunteer reading program on special education referral and placement.

Before presenting the results of the evaluation, we briefly describe relevant research on tutoring, with particular attention to the three previous studies of volunteer tutoring programs that used controlled experimental designs. We then describe the SMART program, including the training of volunteers. Presentation of results and a discussion of implications follow.

Relevant research on tutoring

Considerable research indicates that one-to-one tutoring in which teachers and other paid professionals serve as tutors produces more substantial gains than any other dyad combination, including tutoring by peers, parents, or volunteers (Shanahan, 1998). Wasik and Slavin (1993) examined five of the most popular programs involving one-on-one tutoring by trained adults: Reading Recovery, Success for All, Prevention of Learning Disabilities, the Wallach Tutoring Program, and Programmed Tutorial Reading. They analyzed 16 studies of first-grade tutorials for children at risk of reading failure and found that the overall effect size was .51 standard deviation units, suggesting that the tutored children gained substantially more than untutored comparison students.

Most relevant to the SMART program and the current study is research on the effectiveness of tutoring by volunteers, which is typically conducted outside the normal classroom setting. Volunteer tutoring is increasingly popular because of the vast number of potential volunteers who could dramatically augment the amount of direct, one-to-one reading instruction students receive in the early grades. The popularity of volunteer tutoring has been enhanced considerably by the America Reads Challenge. While campaigning, President Clinton (1996) promoted this challenge by calling for the mobilization of "a million volunteer reading tutors all across America . . . to help every eight-year-old learn to read" [because] . . . "we know that individualized tutoring works."

Wasik (1998) comprehensively summarized research on the impact of volunteer tutoring on early reading achievement. She identified 17 programs that met the following criteria: (a) adult volunteers were used as tutors, (b) the tutoring was in reading, and (c) the children being tutored were in kindergarten through Grade 3. Wasik concluded that empirical analysis of the impact of those programs is complicated since 5 of the 17 programs reviewed presented no evaluation data at all, and only 3 of them used designs of sufficient rigor to allow causal statements to be made about program effectiveness.

Although three programs had adequate outcome data in reading to determine their effects compared to a comparison group, they were quite different from SMART, most notably in the nature and extent of volunteer training and the number of children served. We review these three programs briefly and then provide a more lengthy description of SMART. We compare the effects all four of these programs had on reading outcomes later in the article.

Howard Street Tutoring Program (Morris et al., 1990). The Howard Street Tutoring Program provides one year of one-to-one tutoring for poor readers in Grades 2 and 3. The program began in 1979 and initially served approximately 20 students per year. It has remained a small-scale program. By 1990, at the time of the formal study, it served approximately 50 students per year. A paid reading specialist is essential to the program and trains nonpaid volunteers in groups of two or three during approximately four 1-hour sessions. Volunteers then tutor children at the conclusion of the school day, following lessons individually planned for each child by the reading specialist. In addition to ongoing planning for each child, the reading specialists assist volunteers who need or want special training during the course of the year. Lessons focus on reading of connected text by students, developing alphabetic understanding, writing, and reading by tutors. The program resulted in statistically significant gains relative to the comparison group, on measures of word reading, accuracy of passage reading, and spelling.

Intergenerational Tutoring Program (American Academy of Arts and Sciences & Boston Partners in Education, 1999). The Intergenerational Tutoring Program represents a joint effort among Initiatives for Children of the American Academy of Arts and Sciences, Boston Partners in Education, and the Boston Public Schools to improve the reading skills of first graders who are identified by their teachers as having reading difficulties. At the time of the evaluation, 70 children in low-income Boston schools were being tutored. A certified teacher coordinates the program and does the training. Volunteer tutors receive four 3-hour training sessions prior to the start of tutoring, and ongoing support and training every 2 weeks once tutoring begins. Initial training addresses the basic format of the tutoring sessions, which are conducted three times per week for 45 minutes. Tutoring sessions address letter recognition, word study, phonemic awareness, printing and writing, and guided reading. Once tutoring begins, ongoing training for volunteers covers learning new reading activities and games, sharing tutoring experiences with one another, problem solving, and giving feedback to the Program Coordinator. Tutors keep daily written logs on each of the students they are tutor-

ing. Preliminary analysis indicates the program has had impact on letter identification, but not on measures of word reading, phonemic awareness, or reading connected text.

School Volunteer Development Project (U.S. Department of Education, 1979 as cited in Wasik, 1998). This program was cited in the National Diffusion Network as an exemplary program, even though it was implemented in just two schools in Florida and was terminated during the 1980s. We were able to obtain only a brief description of this project and have relied to some extent on Wasik's (1998) review to help describe the program. The program was developed in Dade County, Florida, for children in Grades 2 through 6 who were functioning 1 or more years below grade level. Community volunteers tutored children for 30 minutes per day, 4 or 5 days per week. Tutors were trained prior to tutoring in a variety of skills. In addition, tutors worked with a reading specialist on the skills they were tutoring. Wasik reported that the program resulted in an overall effect size of .50 on a global measure of academic achievement, the Metropolitan Achievement Test (1984).

Although Wasik (1998) included 14 additional programs in her review of volunteer tutoring programs, none of the 14 had been evaluated by means of an acceptable research design. Salient differences between these 14 programs and Oregon's SMART program are that the training of coordinators and volunteers, and the content of tutoring sessions, are generally far less complex and structured in the SMART program.

The Start Making a Reader Today (SMART) tutoring program

Students are designated for participation in SMART by their teachers, who are asked to choose students who they believe are at risk for reading failure. Students attend tutoring sessions for 30 minutes twice a week throughout the school year, and they may take home two books each month to keep for themselves for home reading. Popular books to read and take home in first grade include *The Very Hungry Caterpillar* (Carle, 1984), and *The Grouchy Ladybug* (Carle, 1996). Popular books in second grade include *A Pocket for Corduroy* (Freeman, 1980) and any number of books from the Arthur series (e.g., *Arthur Babysits* and *Arthur Accused*, Brown, 1997, 1998).

Volunteer tutors represent a diverse group, although considerable emphasis was placed on recruiting members of the business community. Two thirds of all SMART volunteers have been in the program less than 2 years. The greatest proportion of volunteers (33%) is in the 30–45-year-old age group, with the 45–65 age group the next largest (29%). One fifth are over 65.

Volunteers can be trained either in the fall before tutoring has begun or any time during the school year. An initial training session is held at the beginning of the year at a central location, such as the school district central office. The training lasts 1–2 hours, during which 30–40 minutes is devoted to actual reading strategies volunteers can use with students. The remaining time goes to orientation and discussion of logistical and administrative issues, school rules, and safety protocols. Training emphasizes the importance of reading to students and having students read. Volunteers are encouraged to try to increase students' interests in reading, to make the tutoring sessions fun, and to ask students questions about the material they read. After the initial 30-minute training session, volunteers are free to begin working with children on their own, and most receive no additional training.

SMART volunteers may also sign up to be tutors after the school year begins. In this case, training is conducted at the school in impromptu sessions organized by the coordinator. A common training activity is for the coordinator to model a few strategies during a reading session with a student before the volunteer takes over. Approximately half the volunteers are trained in this impromptu fashion.

The key resource for volunteers is the volunteer handbook (Oregon Children's Foundation, 1992, 1998). The handbook indicates that children will improve their reading if (a) they are provided with necessary background to appreciate the story being read, (b) they have opportunities to hear different types of books being read (i.e., some fiction, some science books, some biography, some poetry), (c) they learn something about letter-sound relationships to read unknown words, (d) they make predictions about the story, and (e) they derive meaning from illustrations.

To help children improve in these areas, four reading strategies a volunteer can use with the child are described: (a) the volunteer reads to the child, (b) the volunteer and child read together (e.g., at the same time), (c) the volunteer reads a section of text that the child then rereads, and (e) the volunteer asks the child questions during reading.

The handbook says that these strategies will work best if volunteers review books carefully before reading them with children by relating the content of a book to the child's experiences before reading it, skimming the book as a warm-up activity, and looking at and talking about the illustrations in the book as a way of engaging children in dialogue. A section of the handbook gives volunteers sample questions they can ask children before, during, and after reading, such as What are some words that might be in this story? (before); Is this what

you expected to happen? (during); and Who was your favorite character? (after).

Each school has a half-time SMART coordinator who manages the program in that building. Most coordinators are Americorps volunteers or instructional assistants with no formal training in reading instruction or elementary education. Their coordinator training for SMART amounts to approximately 1 full day per year. The coordinator's main responsibilities are recruiting volunteers, setting up a place in the school for tutoring sessions, making sure there are a sufficient number of books, and working with teachers to identify tutoring times.

The present evaluation of SMART differs from evaluations of other volunteer tutoring programs in that students were in tutoring for 2 years at the time of the evaluation, compared to 1 year for the other programs. SMART also departs from two of the three other programs evaluated in that it provides tutoring to students in first grade. The Intergenerational Tutoring Program also tutored students in first grade, but the other two programs began tutoring students in second grade. The final difference is the number of students being served at the time of the evaluation. In SMART, the number of students tutored was far greater than in the other programs.

There are several important similarities in the evaluations of SMART and the other three volunteer tutoring programs that warrant their grouping as similar studies for comparative analysis. All of the evaluations examined effects on reading outcomes, and each expected its tutoring program to have a positive, measurable impact on reading achievement. All of the evaluations randomly assigned students to treatment and comparison groups. In all cases, the treatment was one-to-one tutoring and the comparison condition in all cases was no tutoring. It is important to note that in all of the programs, students in both treatment and comparison conditions continued to receive regular classroom reading instruction during the volunteer tutoring.

Method

Design

We used an experimental design with random assignment of eligible students within each classroom to either a SMART or comparison group. This type of design is considered optimal for field research (Cook & Campbell, 1979). Pairs of students in each classroom were matched on a salient pretest variable, Rapid Letter Naming (Kaminski & Good, 1996), and randomly assigned to treatment and comparison groups. Letter naming was used because it is one of the best predictors of subsequent reading achievement (Adams, 1990; Bond &

Dykstra, 1967; Chall, 1967; Snow, Burns, & Griffin, 1998). For example, in Bond and Dykstra's classic first-grade study, letter naming at the beginning of first grade was the best predictor of end-of-year reading achievement. Summaries of the research on beginning reading by Chall (1967), Adams (1990), and Snow et al. (1998) have all concluded that letter naming is the single best predictor of beginning reading achievement.

Sample

Sampling procedures. In fall of the first evaluation year, all first-grade classrooms (24 total) in six schools across four school districts provided children for the study. These six schools were selected because they were all in the first year of implementing a SMART program, ensuring that none of the children nominated to participate had been in the SMART program in kindergarten. Typical of the procedure used to select students for SMART in classrooms across the state, approximately one quarter of the students in each classroom (four to six students) were selected by their teacher because they demonstrated reading difficulties. Teachers considered two criteria when nominating their students to participate in SMART: (a) the students' reading skills were among the lowest in their classrooms, and (b) in the teachers' opinions, it was likely the students had relatively few academic literacy experiences with adults or others in the home. The survey that teachers used to nominate students for the SMART program is included in Appendix A.

Teachers were asked to split their nominations approximately evenly between males and females. Teachers also nominated a group of four to six other students whom they believed had about average reading and language skills and who were likely to have had frequent academic literacy experiences in the home. Teachers were explicitly told to select average-ability readers, not high-ability readers (see Appendix A). This group served as a standard for assessing relative reading progress. After teachers' nominations, letters were sent to the parents of nominated children requesting permission for their child to participate in a study explained as an evaluation of a program designed to help students in the early grades become better readers. Parents of two students out of 129 nominated by teachers to participate in SMART declined permission.

When consent was obtained, all students were administered a battery of pretest measures (described in the Measures section). Rapid Letter Naming (Kaminski & Good, 1996; O'Connor, Notari-Syverson, & Vadasy, 1996) was used to evaluate students nominated by their teachers to participate in SMART and assign them to the SMART program or to the comparison group. In each classroom, the two students nominated for SMART who

scored lowest on Rapid Letter Naming were paired. The same procedure was followed with the next two lowest scoring students until all students on the teacher nomination list were paired. Then, one member of each pair was randomly assigned to either the SMART group or the comparison group. In this way, random assignments were made at the classroom level, so there is no reason to expect that quality of classroom reading instruction differed for students in the experimental and comparison groups. The only difference was that students assigned to the SMART group received 1 hour of tutoring per week.

We were concerned about denying services to students who their teachers believed needed tutoring. But because all of the schools were in their first year of SMART implementation, there were not enough tutors available to serve all eligible students. Therefore, those students in the comparison group, who were eligible for tutoring but did not receive it, may not have received tutoring even if the study had not been conducted.

Sample attrition. The original sample—those students who were tested at the beginning of Grade 1—included 64 assigned to the SMART program and 63 assigned to the comparison group. Attrition rates over the 2 years were 33% in the SMART program and 35% in the comparison group. The final samples of students in the SMART and comparison groups were virtually identical on all of the measures administered at pretest.

Students in the comparison group were dropped from the study if they moved to a school outside of the two counties in which the study was being conducted. Students in the SMART group were dropped if they moved to a school outside one of the participating counties or if they moved to a school inside a participating county that did not have a SMART program.

Description of the sample. Students in the final sample were those students who participated in the full 2 years of the evaluation: 43 students in the SMART group and 41 students in the comparison group. All six participating schools were Title I schools located in two of the largest counties in the state. The schools represented a diverse range of communities, from low income/large city to working class/moderate size-city to rural. The communities were representative of the Title I school population of western Oregon. Student ethnicity was as follows: European American (47%) African American (30%), American Indian (10%), Asian American (6%), and Latino (6%). There were 44 female students in the sample and 40 males.

Average-achieving students. A third group of 36 average-ability students was also part of the evaluation. Data for these children were used to assess the progress SMART students made relative to a normative sample of students at the participating schools. Teachers were accurate in selecting average-achieving students as opposed to high-performing students. That is, the scores of these students on the reading measures reflect average as opposed to exceptional performance (see Table 1).

SMART tutoring procedures

Students in the SMART group received one-to-one tutoring for 6 months each year in first and second grade. Tutoring occurred in 30-minute sessions 2 days per week. Over the 2-year period, the number of one-to-one sessions per student ranged from 49 to 98, with a mean of 73 (and a standard deviation of 10.9).

All students in the SMART and comparison groups (as well as the average-achieving group) received regular classroom reading instruction throughout the 2 years. We

Table 1 Pretest means, standard deviations, and percentiles for SMART, comparison, and average-achieving groups

Measures	SMART (N = 43) M (SD)	Percentile	Comparison (N = 41) M (SD)	Percentile	ES(Δ)	Average achieving (N = 36) M (SD)	Percentile
Word Identification, WRMT-R (W score)	354.9 (12.6)	5th	357.5 (14.5)	6th	-.18	389.5 (30.0)	52nd
Letter Naming Fluency	27.7 (14.2)		25.2 (15.1)		.17	47.8 (15.1)	
Phonemic Segmentation	11.7 (7.5)		12.6 (8.8)		-.10	19.1 (7.4)	
Expressive One Word Picture Vocabulary Test-Revised (raw score)	43.3 (16.4)	19th	43.8 (18.3)	23rd	-.03	55.8 (11.4)	57th

made no attempt to influence or interfere with school practices or decisions about how to teach reading to students, nor did we offer advice as to whether students in the study should receive any kind of specialized reading instruction or programs. Likewise, we made no attempt to influence student referral or placement decisions in special education.

Assessment procedures

Students were tested three times in the study: at the beginning of Grade 1 (October, 1996), the end of Grade 1 (May, 1997), and the end of Grade 2 (May, 1998). The pretest battery took 20 to 30 minutes to administer; it was done in one session. First-grade posttesting took 40 to 60 minutes to complete and was done over the course of two sessions separated by no more than 2 days. Second-grade posttesting was completed in one session that took approximately 45 minutes.

As detailed below, all measures possess strong psychometric characteristics for the population of students in the evaluation and have been used in published research studies. Measures administered in the study, except for the commonly utilized measures—that is, the three subtests of the Woodcock Reading Mastery Test-Revised (1998) and the Expressive One Word Picture Vocabulary Test–Revised (1990)—are presented in Appendix B.

Certified teachers and graduate students in school psychology were trained by the first author to administer the battery of measures. Test administrators were kept blind as to which of the three groups students belonged. The reliable administration and scoring of each measure were established for each test administrator before any student in the study was tested. Each new examiner administered the test battery with the first author to their first student included in the study. Both adults scored student performance independently. Reliability was calculated by dividing the lowest raw score by the highest raw score. A checklist was used to ensure reliable administration procedures. When reliability was at least .95, the examiner tested students independently. In all cases, examiners tested students on their own after co-administering the battery with the first author with no more than 2 students. After that, two times per week, each examiner would determine reliability with another examiner using the same procedure. Reliability checks continued throughout the testing period and remained above .95 throughout.

Measures

Four types of measures were administered to students during the evaluation: (a) prereading measures, which included phonemic awareness and alphabetic understanding; (b) reading accuracy and fluency measures, which included word identification and reading fluency;

(c) reading comprehension; and (d) vocabulary, which included a word comprehension measure and an expressive picture vocabulary measure. Table 2 summarizes the schedule for administration of measures at different phases of the study.

Because students' reading abilities change extensively during the course of first and second grade, only one measure (the Word Identification subtest of the Woodcock Reading Mastery Test-Revised) was administered at all three assessment periods. Measures were selected and administered at times when most students were considered to have sufficient skill to provide meaningful information. For example, the reading fluency measure was not administered until the end of first grade because, in the fall, most first graders—especially those considered at risk for reading difficulty—are not yet reading well enough to provide meaningful information in this area. Some early literacy measures were included in the first-grade battery to assess emerging phonemic awareness skills, but were dropped from the second-grade battery, because they no longer fit the children's level of reading development.

Prereading measures

In the fall and spring of first grade, children were administered two measures to assess the core underlying processes in learning to read: Phonemic Segmentation and Rapid Letter Naming (Adams, 1990; Kaminski & Good, 1996; O'Connor et al., 1996; Snow et al., 1998; Torgesen, Morgan, & Davis, 1992).

Phonemic Segmentation (O'Connor et al., 1996). On this measure of phonemic awareness, examiners orally presented 3-phoneme words to students one at a time. Students responded by saying the individual phonemes in each word. For example, the examiner would say "*make.*"To answer correctly, children would say "/m/ /a/ /k/." As specified in the testing procedures, the task was modeled and practiced prior to administration. During administration, children received 1 point for each correct phoneme they produced (i.e., 0 to 3 points per word). The measure took 3–5 minutes to administer. Alternate-form reliability on a similar measure of phonemic segmentation (Kaminski & Good, 1996) was reported at .88, and predictive validity with reading measures that ranged from .73 to .91.

Rapid Letter Naming (Kaminski & Good, 1996). On this measure, students were presented with randomly ordered upper- and lowercase letters arranged in rows on a sheet of paper. They were asked to name as many letters as possible in 1 minute. The number of correctly named letters per minute was calculated. Reliability of the measure has been reported at .93 by Kaminski and Good (1996), who also reported 1-year predictive validity coef-

Table 2 Administration schedule of primary measures used in the evaluation

Measure	Fall first grade	Spring first grade	Spring second grade
Prereading			
Phonemic Segmentation	X	X	
Rapid Letter Naming	X	X	
Reading Accuracy and Fluency			
Word Identification: WRMT-R	X	X	X
Oral Reading Fluency First-Grade Passage		X	X
Oral Reading Fluency Second-Grade Passage			X
Reading Comprehension			
Passage Comprehension: WRMT-R		X	X
Vocabulary Knowledge			
Expressive One Word Picture Vocabulary Test-Revised	X	X	
Word Comprehension: WRMT-R			X

Note. WRMT-R stands for the Woodcock Reading Mastery Test-Revised

ficients with reading criterion measures that ranged from .72 to .98.

Reading accuracy and fluency

We measured two aspects of reading accuracy and fluency: (a) reading isolated words correctly and (b) reading connected text fluently. The Word Identification subtest of the Woodcock Reading Mastery Test-Revised was used to assess accuracy of word reading. Oral Reading Fluency was used to assess ability to read words fluently (Shinn, 1998).

Word Identification Subtest of the Woodcock Reading Mastery Test-Revised (1998). The Word Identification subtest measures a student's ability to read words in isolation. The test begins with simple words and gradually becomes more difficult. The test takes from 2 to 15 minutes to administer, depending on a student's reading ability. This subtest was administered at all three testing times. Split-half reliability estimates are reported to be .98 for first graders. According to the examiner's manual, the correlation between the Word Identification subtest and the *Woodcock-Johnson Total Reading* score is .82 for first-grade students.

Oral Reading Fluency (Shinn, 1998). Oral Reading Fluency has been used in educational research and practice as a measure of reading proficiency for more than 15 years. Research has demonstrated consistently that the number of words students read correctly in 1 minute provides a reliable and valid measure of overall reading ability (Fuchs, Fuchs, & Maxwell, 1988; Potter & Wamre, 1990; Shinn, Good, Knutson, Tilly, & Collins, 1992). Standardized procedures were used for administration of this measure (Shinn, 1989). Each student read aloud a story written at either a first- or a second-grade level. The

reading passages were taken from a basal reading series and were used by the first author on numerous occasions in schools as a method to assess the reading skills and progress of beginning readers. The number of words the student read correctly in 1 minute provided an index used in data analysis. The first-grade passage was administered during spring of first and second grade. The second-grade passage was administered only during the spring of second grade.

Estimates of the internal consistency, test-retest, and interscorer reliability for Oral Reading Fluency have ranged from .89 to .99. Correlations with other measures of reading, including measures of decoding and comprehension, have ranged from .73 to .91 (Shinn, Tindal, & Stein, 1988). Correlations between Oral Reading Fluency and standardized measures of reading comprehension are typically above .80 (Marston, 1989). Shinn et al. (1992) conducted a confirmatory factor analysis of Oral Reading Fluency and concluded that in the early grades the measure was as valid an indicator of reading comprehension as it was an indicator of decoding ability.

Reading comprehension

The Passage Comprehension subtest of the Woodcock Reading Mastery Test-Revised was used to assess reading comprehension. This subtest provided an indication of the child's ability to comprehend short written text. The child read a portion of text silently and then supplied a missing word appropriate to the context of the passage. Administration time for this test ranged from 10 minutes to 25 minutes, depending on the ability of the student. Passage Comprehension was administered at the spring testing in first and second grade. According to the test manual, split-half reliability estimates for this measure

were .94 for first graders. The correlation between Passage Comprehension and the Woodcock-Johnson Total Reading score was .63.

Vocabulary knowledge

We measured two aspects of vocabulary knowledge. The Word Comprehension subtest of the Woodcock Reading Mastery Test-Revised was used to assess word comprehension. The Expressive One Word Picture Vocabulary Test-Revised (1990) was used to assess expressive vocabulary.

Word Comprehension: Antonyms, Synonyms, and Analogies Subtests of the Woodcock Reading Mastery Test-Revised. Word Comprehension assesses a student's reading vocabulary. On the Antonyms subtest, students read individual words out loud and state a word that means the opposite. On the Synonyms subtest, students read individual words and state another word with the same meaning. On the Analogies subtest, students read three words, two of which are related to each other, and are asked to supply a fourth word that completes the analogy. These subtests were administered at the end of second grade. Split-half reliability estimates for this subtest were .95 for first graders. The correlation of Word Comprehension with the Woodcock-Johnson Total Reading score for first graders was .82.

Expressive One Word Picture Vocabulary Test-Revised (EOWPVT-R). The Expressive One Word Picture Vocabulary Test-Revised (EOWPVT-R) is a measure of expressive language, an important component in reading comprehension. On this test, examiners asked children to name individual pictures (e.g., apple) or to tell what was happening in a picture (e.g., eating). Median split-half reliability coefficients for the EOWPVT-R are reported at .90. Criterion related validity coefficients with Peabody Picture Vocabulary Test-Revised, a measure of receptive language, was reported at .59. The measure was administered for two reasons: as a possible predictor of reading acquisition, and because it was hypothesized that the dialogic nature of SMART tutoring might result in improved vocabulary knowledge of students. The EOWPVT-R was administered in the fall and spring of Grade 1 only. Interim analyses (i.e., spring of Grade 1) revealed the SMART and comparison groups were virtually identical at both pretest and posttest. Consequently, the EOWPVT-R was not administered at the end of second grade.

Results

As seen in Table 2, data collection occurred in the fall of first grade, spring of first grade, and spring of second grade, with the measures varying somewhat at each point. Pretest data are presented first, confirming that the experimental and comparison samples are statistically equivalent. We then present an analysis of the impact of tutoring on reading achievement as measured through covariance procedures and growth curves, followed by comparison of rates of student placement in special education. Finally, we present a supplementary analysis of volunteers' perceptions of the impact of tutoring on their students and on their own view of schools.

Pretest data

Table 1 presents means, standard deviations, and percentiles for the SMART, comparison, and average-ability groups, on the pretest battery: the Word Identification subtest of the Woodcock Reading Mastery Test-Revised, Rapid Letter Naming, Phonemic Segmentation, and the EOWPVT-R. Effect sizes are also presented using Glass's Δ (Cooper & Hedges, 1994), a commonly used measure of effect size. Glass's Δ was computed by subtracting the mean of the comparison group from the mean of the SMART group and dividing by the standard deviation of the comparison group (Cooper & Hedges, 1994).

There were no statistically significant differences between the SMART and comparison groups on any of the pretest measures: WRMT-R, $t(82) = -.89$, $p = .37$; Rapid Letter Naming, $t(82) = .78$, $p = .44$; Phonemic Segmentation, $t(82) = -.53$, $p = .60$; EOWPVT-R, $t(82) = -.14$, $p = .89$). Glass's Δ in Table 1 shows the high degree of comparability between the two groups at pretest. On all four measures, effect sizes are very small, and vary in a nonsystematic fashion. The mean effect size of -.05 is close to zero.

Differences in achievement at end of first and second grades

Two reading measures were administered at the end of first and second grade, Oral Reading Fluency (First-Grade Passage) and the Passage Comprehension subtest of the Woodcock Reading Mastery Test-Revised. Analysis of covariance (ANCOVA) was used to analyze performance differences between SMART and comparison students at the end of second grade. The two covariates were pretest scores (beginning of first grade) on (a) Phonemic Segmentation and (b) the Word Identification subtest of the Woodcock Reading Mastery Test-Revised. Results indicated a statistically significant effect favoring SMART on the Oral Reading Fluency measure (First-Grade Passage); $F(1, 80) = 7.61$, $p = .007$. The effect on the Reading Comprehension subtest approached significance, $F(1, 80) = 3.46$, $p = .067$.

For the reading measures administered at the end of second grade only, ANCOVA results were statistically sig-

nificant both for Oral Reading Fluency (Second-Grade Passage), $F(1, 80) = 6.37$; $p = .014$, and the Word Comprehension subtest of the Woodcock Reading Mastery Test-Revised, $F(1, 80) = 5.20$; $p = .025$.

Effect sizes for the different reading measures, using Glass's Δ were Word Identification, .44; Oral Reading Fluency (First-Grade Passage), .48; Oral Reading Fluency (Second-Grade Passage), .53; Word Comprehension, .43; and Passage Comprehension, .32. Except for Passage Comprehension, these effects are considered moderate in magnitude (Cohen, 1988). The effect on Passage Comprehension is considered small.

Supplemental analyses: Impact on predictive measures during Grade 1. Because SMART tutoring did not in any way stress phonemic awareness, we expected there to be no effect on that aspect of reading ability. As predicted, Phonemic Segmentation and Rapid Letter Naming

were not affected by the SMART tutoring. ANCOVAs revealed no statistically significant differences on these measures at the end of the first grade. Effect sizes on Phonemic Segmentation and Rapid Letter Naming for SMART versus comparison groups were .07 and -.06, indicating virtually identical performance. In addition, an ANCOVA showed that there was no difference between SMART and comparison students on the measure of expressive language (EOWPVT-R) at the end of first grade. The effect size of .12 favoring SMART was small.

Analysis of reading growth over time (word identification only)

Table 3 presents the results on measures of reading proficiency, at both the interim assessment (i.e., end of first grade), and for the final assessment (end of second grade).

Table 3 Means, standard deviations, and percentiles for SMART and comparison groups at the end of first and second grades

Measures	SMART (N = 43) M (SD)	SMART (N = 43) Percentile	Comparison (N = 41) M (SD)	Comparison (N = 41) Percentile	ES(Δ)	Average achieving (N = 36) M (SD)	Average achieving (N = 36) Percentile
			Interim scores (end of first grade)				
Reading accuracy and fluency							
Word Identification, WRMT-R (W score)	409.2 (29.7)	33rd	398.9 (24.4)	21st	.42	438.6 (30.2)	69th
Oral Reading Fluency First-Grade Passage	27.8 (22.8)		18.7 (17.3)		.53	57.0 (34.2)	
Reading comprehension							
Passage Comprehension, WRMT-R (W score)	449.3 (24.4)	23rd	443.2 (14.2)	15th	.43	466.1 (16.0)	55th
			Posttest scores (end of second grade)				
Reading accuracy and fluency							
*Word Identification, WRMT-R (W score)	449.4 (30.2)	29th	437.9 (25.9)	21st	.44	470.4 (22.1)	47th
**Oral Reading Fluency First-Grade passage	71.3 (35.2)		55.9 (32.1)		.48	98.8 (35.1)	
*Oral Reading Fluency Second-Grade passage	61.5 (35.5)		45.9 (29.5)		.53	90.5 (38.3)	
Reading comprehension							
*Word Comprehension, WRMT-R (W score)	472.3 (17.3)	31st	465.4 (16.2)	19th	.43	487.8 (9.7)	69th
Passage Comprehension, WRMT-R (W score)	468.9 (16.0)	28th	464.7 (13.1)	22nd	.32	481.7 (10.9)	53rd

Note. Means for the SMART and comparison groups are adjusted for pretest performance on the Phonemic Segmentation test and the Word Identification subtest of the WRMT-R using analysis of covariance. Significance tests are between the SMART and comparison groups. Percentiles and effect sizes are determined using the adjusted means. Means for the average-achieving group are unadjusted.
* $p < .05$. ** $p < .01$.

The Word Identification subtest of the Woodcock Reading Mastery Test-Revised was administered at all three test phases (i.e., the beginning of first grade, the end of first grade, and the end of second grade). We used individual growth curve methodology (see Bryk & Raudenbush, 1992, or Stoolmiller, 1995, for more details and relevant example applications) to analyze change in student performance on this measure. This method allowed us to estimate (a) the mean rate of change and the extent to which individual children varied about the mean rate of change and (b) correlates of individual variability in change, which in this investigation focused on group status. Maximum likelihood estimation was used for nested chi-square tests of variance components using the Mplus program (Muthen & Muthen, 1998). Restricted maximum likelihood estimation was used for estimation of group differences using the LME procedure in the S-plus 4 software package (Mathsoft, 1998).

We were interested in two questions regarding the growth parameters of students in these three groups. First, we predicted that students in SMART would have growth rates that would surpass the word reading growth of students in the matched comparison group. Second, we predicted that students in SMART would have growth rates that were similar to the growth rates of students in the average-achieving group.

To characterize the pattern of change over time, we first fit models to determine whether growth was linear or a curvilinear 2nd degree polynomial (i.e., a combination of linear and quadratic). We computed a nested chi-square statistic to assess the importance of the quadratic factor in understanding rate of growth in all 3 samples (SMART, the comparison group of at-risk readers, and the average-ability group). Within-subject error variance was fixed at about 2% of the total time 1 variance and $\chi 2 =$ 1953, $df = 4$, was statistically significant. Inspection of the z test statistics associated with the individual quadratic growth factor parameters revealed that all 5 were highly statistically significant, the minimum z being -3.31. Thus, growth in word reading was best described by a curvilinear model (i.e., a combination of linear and quadratic trends), indicating that growth for all 3 samples was greatest in Grade 1 and tapered off in Grade 2.

To test for group differences in growth rates among the three groups, dummy coded group membership variables using two contrasts—SMART versus matched comparison and SMART versus average achieving—were added to the growth model. Differences were statistically significant among the instructional groups in terms of intercept and slope. The pooled within-group standard deviation of the linear growth rate factor was 1.63.

The SMART, matched comparison, and average-ability groups had estimated linear growth rates of 5.94, 5.20, and 5.20 words per month, respectively. The differences between SMART and the other groups were statistically significant in both cases ($z = -2.07$, one-tailed $p = .019$ for matched comparison; $z = -2.00$, one-tailed $p = .023$ for average ability). The effect sizes for the differences were .45 pooled standard deviation units for both comparisons.

Student placement in special education

Relative rates of student referral and placement in special education were compared through chi-square analysis. Data were collected at the six evaluation schools in the fall of Grade 3 on special education referral and placement. We asked two questions of special education teachers at each school: (a) While the target students were attending that particular school, had they ever been referred for special education services? (b) While the target students were attending that particular school, had they ever been placed in special education?

At the beginning of third grade, 38 of the original 43 students in the SMART group (88%) were attending the same school they attended in Grades 1 and 2. Of the 41 students in the comparison group, 32 still attended the same school (78%). For these samples of students, we were able to trace special education referral and placement. By the beginning of third grade, 15 of the 38 students in the SMART group had been referred for special education (39%). Of these 15 students, 10 were actually placed in special education (26%). For the comparison group, 18 of the 32 were referred for special education placement (56%). Of these 18, 14 were placed in special education (44%).

Chi-square analysis showed that the difference in rate of special education placement by fall of grade 3, 39% for SMART versus 56% for the comparison group, approached statistical significance ($p = .12$). Power in this instance was limited by the relatively small sample size for chi-square analysis. We believe the lower special education placement rate for students in the SMART group is a potentially important finding. It is also worth noting that the special education placement rates are very high for *both* groups, providing additional evidence that SMART serves students who need as much support as possible learning to read.

Overall, the results of the evaluation suggest that in terms of reading achievement, students in SMART benefited a great deal from their participation in the tutoring program. Before discussing in detail the nature of those improvements and possible explanations, we examine the results of a survey of the program's volunteers.

Volunteers' perceptions of the impact of tutoring

In 1993 and 1994, 3 years prior to beginning the longitudinal study of SMART reading outcomes, we asked

all SMART volunteers in Oregon to complete a survey questionnaire in order to learn their impressions of the volunteer experience. A total of 903 volunteers submitted responses in 1993 and 986 in 1994. The responses were not significantly different statistically across the 2 years. About half of the SMART respondents rated their participation as an "excellent and valuable experience." Another 45% felt it was "a worthwhile experience." Ninety-five percent of volunteers felt the role of the SMART volunteer was a challenging experience.

Eighty-two percent felt their training was either excellent or good, while 17% felt the training did only a fair job of preparing them. About half of the volunteers indicated they "would like more guidance" in learning additional techniques and tools for improving children's reading and understanding. The need for more guidance, as opposed to extensive training, was frequently voiced. Volunteers desired information such as how to deal with students who have short attention spans, students who tell them "they did not want to read," or students who appeared unhappy or angry. Some requested more age-specific training, including information on what to expect at a given age or grade level. Several asked for ideas for alternative strategies or activities to increase motivation.

In open-ended responses, volunteers indicated that they would like more opportunities to meet with teachers, primarily to receive guidance on reading instruction, and to get a sense of how the teacher viewed the child's reading progress. The theme of emotional bonding consistently emerged from the open-ended responses. Many tutors described the intense nature of the one-on-one tutoring situation and the depth of feeling toward their students. For example, one volunteer noted that "so many of them are so angry and frustrated they can't listen or learn." Many volunteers indicated they would like to know more about their students' lives or more about the specific nature of their students' learning problems. Many volunteers complained about the fact that tutoring sessions had been canceled due to an array of school activities (such as assemblies or plays) and that tutors were rarely given advance notice. The quality and quantity of space provided were also a common concern.

Volunteers were very positive about the books that were available for students through SMART. About 80% indicated that they often or always found books appropriate for their students. Volunteers with Spanish-speaking students indicated a need for Spanish-language books and books appropriate for older children who were not yet fluent in English. Several suggested that books be organized according to reading or grade level. (This was later done.)

Volunteers disagreed somewhat about ways to improve SMART tutoring. Some indicated they wanted more games, puzzles, and other activities, while others believed these distracted from the reading tasks. Some volunteers indicated their roles should be simply to promote the joy of reading, while others felt it more important to teach reading skills.

Perhaps the most interesting survey finding was volunteers' responses to how the SMART experience changed their views of school. Many expressed an increased understanding of the challenging job teachers have and recognized the limited resources available to schools. One volunteer summarized the challenges by commenting how reading problems often are "coupled with emotional problems" in young learners. Most were impressed with the efforts schools directed towards these challenges.

One person summarized the intensity and importance of the tutoring experience this way: "As a parent with four kids, I have long viewed schools as the most critical battleground in our society. SMART rubs my face in it. SMART tutoring should be required of all adults, especially those voting against school funding."

Discussion

This study found that Oregon's SMART volunteer tutoring program improved the reading abilities of students deemed at risk for failure in reading. On most measures of reading, the performance of students in SMART was statistically higher than was the performance of students in a randomly assigned, matched comparison sample. Statistically significant differences were found on three aspects of reading: word reading, reading fluency, and word comprehension (i.e., reading vocabulary). The impact of the intervention on passage comprehension was not statistically significant, but the difference favored students in the SMART group and approached statistical significance ($p = .07$).

Effect sizes on all reading measures indicated the impact was at the level of educational importance. Effect sizes ranged from a low of .32 on reading comprehension to a high of .53 on the second-grade passage of Oral Reading Fluency. Taken together, the analysis indicated that SMART had a clear, positive impact on the reading achievement of students who received tutoring.

On the Word Identification measure (the subtest of the Woodcock Reading Mastery Test-Revised), the data were analyzed using growth curve analysis in order to determine the relative rates of growth for students in SMART, the comparison group, and average-achieving readers. The growth rate of students in the SMART group surpassed the growth of students in the comparison group, as well as the growth rate of students in the average-achieving group. The SMART group's greater growth

compared with the average-achieving group is particularly important given what we know about the ever-expanding gap between good and poor readers over time unless intensive early intervention takes place (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Stanovich, 1986).

Despite the greater growth rate of students in SMART versus the other two groups, at the end of second grade their level of performance was still much lower than the performance of students in the average-achieving group. At the end of second grade, the mean score for students in SMART corresponded to about the 30th percentile across the three subtests of the WRMT-R, compared to a mean score falling between the 47th and 69th percentiles for students in the average-achieving group. Thus many students in SMART remained at risk of reading-related difficulties in their subsequent school careers. At the end of second grade, students in the comparison group, with a mean score corresponding to about the 20th percentile across the three WRMT-R subtests, were at even greater risk for reading-related difficulties.

The data indicate that reading-related difficulties began to surface for some students in the study by third grade. For example, 44% of students in the comparison group had been placed in special education by the fall of third grade, compared to 26% for students in SMART. This difference is not statistically significant but suggests a possible trend that should be further investigated. It does show clearly, however, that students in SMART and students in the comparison group remain at considerable risk of reading-related difficulties. Current research is aimed at determining which children in SMART made the greatest reading growth, and a means for determining which students require a more intensive intervention than SMART beginning in first grade to avoid serious reading difficulties (Baker, Stoolmiller, & Gersten, 2000).

Strength of effects of SMART versus other volunteer reading programs

It is important to place the findings of this study in the context of other research on tutoring. Recent reviews by Wasik (1998) and Shanahan (1998) clearly indicated that there is a dearth of well-controlled research investigating the effects of volunteer tutoring programs on student reading. Because of this, Wasik and Shanahan indicated that all they actually could present were hypotheses about best practice. Neither the degree of impact that volunteer reading programs have on the reading achievement of students in the primary grades nor the degree to which the training of volunteers influences level of impact are clear from previous research.

A closer examination of SMART compared to the three other volunteer reading programs that used a simi-

lar experimental design in their evaluations provides some preliminary answers. When possible in these analyses, we rely on effect size comparisons on comparable measures of reading achievement.

The Howard Street Tutoring Program. The effect size on measures of word recognition for SMART and the Howard Street Tutoring Program was nearly identical, .44 and .42 respectively. Reading researchers have long considered word recognition to be the linchpin for successful reading (Adams, 1990; Foorman et al., 1998; Stanovich, 1986).

The School Volunteer Development Project. The overall effect size of the School Volunteer Development Project was .50 on a measure of overall achievement, the Metropolitan Achievement Test (Wasik, 1998). In SMART, the mean effect size across all reading measures was .44. The difference, though quite small, suggests a slightly stronger effect for the School Volunteer Program. It is not clear how many hours students were tutored in the School Volunteer Development Project, but sessions were 30 minutes long and were conducted four to five times per week. Thus, the total time in tutoring sessions over one year was likely to have been roughly the same as the SMART condition.

The Intergenerational Tutoring Program. Preliminary analysis of the Intergenerational Tutoring Program became available last year (American Academy of Arts and Sciences & Boston Partners in Education, 1999). The preliminary analysis is based on 140 students, assigned randomly to either the experimental tutoring group or the no-treatment comparison group. Across a number of outcome reading measures, the report indicates there was a statistically significant difference between the groups only on a measure of letter identification. On measures of word reading, phonemic awareness, and reading of text, there were no statistically significant differences between the groups. Further analysis needs to be conducted to determine impact, but it appears the effect of SMART is greater than the effect achieved in this program. Because the preliminary report on the effect of the Intergenerational Tutoring Program does not present data on the comparison group, calculating an effect size was not possible.

Implications

At the beginning of the study, we asked many of our colleagues with expertise in early reading instruction to make predictions about the outcome SMART would have on reading achievement. Like us, they were unsure what to predict. Some expressed concern that a program as loosely structured as SMART might not be able to affect the reading achievement of students who were clear-

ly among the teachers' greatest concerns. After all, they reasoned, volunteers received only minimal training, and the wide latitude they were given in organizing the tutoring sessions could result in a pattern of tutoring decision making not particularly helpful to students struggling to learn to read. There was also concern that SMART in no way specifically supported the development of phonemic awareness skills.

However, other colleagues raised the possibility that the very looseness and flexibility of the program could be a strength. Without feeling burdened by extensive procedural expectations and routines, and without the need to attend more than one formal training session a year, the adult volunteers (many of whom were very successful in their respective occupations and active members of their communities) would rely on their own resources and insights to figure out how to best tutor children. Regardless of the specific positions we and our colleagues took in predicting reading outcomes, we agreed there would be positive benefits of regular one-to-one interactions between children and adults beyond the scope of reading achievement. We also agreed that such benefits could not be achieved easily in typical general education classrooms.

Comparison to other experimentally evaluated programs

What explains the impact of SMART; given similar effects compared to programs that provide more extensive training to their volunteers? Part of the explanation may be that, compared to the other three programs that were evaluated with the use of an experimental design, SMART provided tutoring to students over 2 years rather than 1. Although SMART lasted for 2 years, SMART students participated at a less intensive level during that time, with the result that they actually received either a comparable or lesser amount of time with an adult tutor as compared to the other programs.

For instance, students in SMART were tutored for 2 years and students in Howard Street were tutored for 1 year. In SMART, students received an average of 73 sessions over 2 years, in two 30-minute sessions per week. Howard Street sessions were 1 hour long, and 50 sessions were provided during the course of the year. Thus the total average time in tutoring sessions for each child was 36.5 hours for SMART (spread over 2 years) and 50 hours for Howard Street over 1 year.

SMART differed from the other volunteer programs in providing tutoring to students during both first and second grade. The Howard Street Tutoring Program and the School Volunteer Tutoring Project began tutoring students in second grade. The Intergenerational Tutoring Program provided tutoring to students in first grade only, and to date the reading outcomes have been mixed.

Rapid growth in reading occurs in both Grades 1 and 2, and most reading experts agree these are excellent grades for adult tutoring (Juel, 1994).

The well-known Book Buddies intervention (Invernizzi, Juel, & Rosemary, 1996) found good effects at the end of first grade, after just 1 year of tutoring. Juel (1994) suggested, however, that the long-term benefits of Book Buddies would likely be much stronger if tutoring was provided to students while they are in first and second grade. She stated that

> successful intervention in first [grade] may be enough to ensure word recognition skill, or at least to have this skill under way so that a follow-up in the second grade could cement it. Without such a follow-up, those children who do not read during the summer are in danger of losing some of their skill in word recognition. (p. 59)

Researchers have long been aware of the problem of summer loss in reading achievement for many low-income students or students with reading difficulties (Natriello, McDill, & Pallas, 1990). Growth curve analysis in this study indicated that for all three groups (i.e., SMART, comparison, and average ability), greatest growth occurred in Grade 1 and tapered off slightly in Grade 2. We believe there are at least two reasonable explanations for this pattern.

For purposes of the growth curve analysis, we attempted to correct for the approximately 3 months of summer between the end of Grade 1 and the beginning of Grade 2—a period of time during which we did not expect students to make growth in reading. The three data points in the growth curve analysis represented the beginning of Grade 1, the end of Grade 1, and the end of Grade 2. We estimated 7 months between fall and spring testing in first grade and 9 months between end-of-first and end-of-second grade testing. The 9 months may have undercorrected for the loss in reading proficiency over the summer. This could have resulted in the slight curvilinear trend noted for all three samples. A second reasonable explanation is that it is quite possible that growth for all three samples was lower in second grade than first grade on norm-referenced tests.

The ability to serve a large number of students

The flexible nature of SMART has played an important role in its rapid expansion. The founders of the program would like to implement SMART in as many low-income schools in the state as possible. Currently, SMART is in 16% of the elementary schools in the state, the great majority of which are located in low-income neighborhoods.

SMART is unique among volunteer reading programs in that it has used solid evaluation methods to

demonstrate a positive impact on reading, and has achieved a widespread impact in terms of the number of students and geographic areas served. For students who have serious reading problems, tutoring by an adult in just two 30-minute sessions per week—1 hour total per week—might seem insufficient to yield measurable reading benefits. However, from the student's own perspective, 1 hour per week may be quite sufficient. To sit down with an adult for that amount of time each week and focus solely on books and reading may well have a profound effect on a struggling student, especially a student who may receive little literacy-related support at home.

Training of volunteers

We believe one of the major attractions of the SMART program to volunteers (and to classroom teachers) is its simplicity. Volunteer tutors need not obtain knowledge of each classroom's reading program nor must teachers spend time explaining the reading program to tutors. SMART operates essentially independently of a given teacher's approach to reading instruction. On one hand, it is impressive that program impact is statistically significant even in the absence of substantive training and expectations of tutors. At the same time, it is unclear just exactly what the sessions consisted of. Clearly, a formal observational study of the range of methods SMART volunteers actually use during a tutoring session would be a logical next step in this line of research. When left largely to their own devices, what do adults do with struggling readers when the goal is reading improvement? It may be that a positive experience with a caring adult better characterizes the adult-student relationship than the use of specific reading instruction techniques. Perhaps the nature of the relationship leads tutored students to invest more effort in their interactions with the classroom teacher and thus benefit more from instruction.

In designing the present study, we did conduct informal observations in two schools to get a sense of the nature of the tutoring sessions and whether tutors seemed to be following the very general guidelines the program provided (Deathridge, 1993). Across a number of tutoring sessions in two schools, it was clear that volunteers took their roles as reading tutors seriously. They used an array of activities and approaches during the sessions, and in general, they had students practice reading independently, especially students in the second grade. Most volunteers demonstrated useful strategies for helping their students figure out what to do when they encountered difficult text. It also seemed clear that students felt supported during the tutoring sessions and looked forward to the time they would spend with their tutors. However, these qualitative observations of approximately

8 tutors are not sufficient to link tutoring methods to reading outcomes, or to generalize to the range of SMART volunteers tutoring children.

SMART is clearly less structured than most other volunteer reading programs. In terms of volunteer training, SMART provides less initial and ongoing training than other programs. Shanahan (1998) recently concluded that although most of the research on tutoring describes programs with intensive tutor training, extensive training may not always be necessary. Our results tend to support his hypothesis.

Most educators believe that careful training of tutors is an important component of volunteer tutoring programs. Roller (1998), for example, who directed a volunteer tutoring program at the University of Iowa for America Reads, said that

> Reading tutors need to know a great deal....They need to know what tutoring looks like, they need to know how skilled reading operates, and they need to know how reading and writing develop....Reading tutors need to know the letter-sound relations that characterize the English writing system and the high-frequency words that make up much of the English text. (p. 50)

Actually, few direct studies have been conducted that investigate the impact on reading achievement of different types of volunteer training (Shanahan, 1998). Shanahan found only one study that evaluated the impact of tutor training on the learning gains of the students tutored. Most relevant to the training of volunteer tutors was a study conducted by Leach and Siddall (1990), in which greater progress in reading accuracy and comprehension occurred for a group of young children when parent tutors were provided with $1\frac{1}{2}$ hours of training. Note that this amount of training is also quite minimal.

The findings from our study suggest that accelerated reading outcomes can be achieved by volunteers with minimal formal training, using their own judgment and instincts on how to support literacy development. Although desirable, intensive tutor training may not always be available or feasible, and considering the high turnover of adult volunteers that invariably occurs, intensive initial training may not be a good use of fiscal resources. The fact that minimally trained adult tutors can enhance meaningful growth in reading has important implications in designing programs such as those related to the America Reads initiative.

REFERENCES
ADAMS, M.J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: The MIT Press.
AMERICAN ACADEMY OF ARTS AND SCIENCES & BOSTON PARTNERS IN EDUCATION. (1999). *The Intergenerational Tutoring Grogram.* Boston: Author.
BAKER, S., STOOLMILLER, M., & GERSTEN, R. (2000). *Predictors of stu-*

dents' ability to benefit from adult tutoring in first and second grade. Manuscript in preparation.

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

BROWN, M.C. (1997). *Arthur babysits.* New York: Little, Brown.

BROWN, M.C. (1998). *Arthur accused.* New York: Little, Brown.

BRYK, A.S., & RAUDENBUSH, S.W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

CALIFORNIA DEPARTMENT OF EDUCATION. (1998). *Reading/language arts framework for California public schools: Kindergarten through grade twelve.* Sacramento, CA: Author.

CARLE, E. (1984). *The very hungry caterpillar.* New York: Putnam.

CARLE, E. (1996). *The grouchy ladybug.* New York: HarperCollins.

CHALL, J.S. (1967). *Learning to read: The great debate.* New York: McGraw-Hill.

CLINTON, W.J. (1996, August 30). Campaign train tour speech. Chicago, IL.

COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

COOK, T.D., & CAMPBELL, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand-McNally.

COOPER, H., & HEDGES, L.V. (Eds.). (1994). *The handbook of research synthesis.* New York: Russell Sage Foundation.

DEATHRIDGE, M. (1993). *Observations of SMART tutoring* (Tech. Rep. No. 93-01). Eugene, OR: Eugene Research Institute.

FREEMAN, D. (1980). *A pocket for Corduroy.* London: Puffin.

FOORMAN, B.R., FRANCIS, D.J., FLETCHER, J.M., SCHATSCHNEIDER, C., & MEHTA, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*(1), 37–55.

FUCHS, L.S., FUCHS, D., & MAXWELL, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9,* 20–28.

INVERNIZZI, M., JUEL, C., & ROSEMARY, C.A. (1996). A community volunteer tutorial that works. *The Reading Teacher, 50,* 304–311.

JUEL, C. (1994). At-risk university students tutoring at-risk elementary school children. In E.H. Hiebert & B.M. Taylor (Eds.), *Getting reading right from the start* (pp. 39–61). Boston: Allyn & Bacon.

JUEL, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly, 31,* 268–289.

KAMINSKI, R.A., & GOOD, R.H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25,* 215–227.

LEACH, D.J., & SIDALL, S.W. (1990). Parental involvement in the teaching of reading: A comparison of hearing reading, paired reading, pause, prompt, praise, and direct instruction methods. *British Journal of Educational Psychology, 60,* 349–355.

MARSTON, D. (1989). Curriculum-based measurement: What is it and why do it? In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18–78). New York: Guilford.

MATHSOFT. (1998). *Splus 4 guide to statistics.* Seattle, WA: Author.

MORRIS, D., SHAW, B., & PERNEY, J. (1990). Helping low readers in grades 2 and 3: An after-school volunteer tutoring program. *Elementary School Journal, 91,* 133–150.

MUTHEN, L.K., & MUTHEN, B.O. (1998). *Mplus user's guide.* Los Angeles, CA: Muthen & Muthen.

NATRIELLO, G., MCDILL, E.L., & PALLAS, A.M. (1990). *Schooling disadvantaged children.* New York: Teachers College Press.

O'CONNOR, R.E., NOTARI-SYVERSON, A., & VADASY, P.F. (1996). Ladders to literacy: The effects of teacher-led phonological activities for kindergarten children with and without learning disabilities. *Exceptional Children, 63,* 117–130.

OREGON CHILDREN'S FOUNDATION. (1992). *SMART volunteer handbook.* Portland, OR: Author.

OREGON CHILDREN'S FOUNDATION. (1998). *SMART volunteer handbook.* Portland, OR: Author.

POTTER, M.L., & WAMRE, H.M. (1990). Curriculum-based measurement and developmental reading models: Opportunities for cross-validation. *Exceptional Children, 57,* 16–25.

ROLLER, C.M. (1998). *So . . . what's a tutor to do?* Newark, DE: International Reading Association.

SHANAHAN, T. (1998). On the effectiveness and limitations of tutoring in reading. In P.D. Pearson & A. Iran-Nejad (Eds.), *Review of research in education* (pp. 217–234). Washington, DC: American Educational Research Association.

SHINN, M.R. (1989). *Curriculum-based measurement: Assessing special children.* New York: The Guilford Press.

SHINN, M.R. (Ed.). (1998). *Advanced applications of curriculum-based measurement.* New York: The Guilford Press.

SHINN, M.R., GOOD, R.H., KNUTSON, N., TILLY, W.D., & COLLINS, V. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21,* 459–479.

SHINN, M., TINDAL, G.A., & STEIN, S. (1988). Curriculum-based measurement and the identification of mildly handicapped students: A research review. *Professional School Psychology, 3*(1), 69–85.

SNOW, C.E., BURNS, M.S., & GRIFFIN, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

STANOVICH, K.E. (1986). Cognitive processes and the reading problems of learning-disabled children: Evaluating the assumption of specificity. In J.K. Torgesen & B.Y.L. Wong (Eds.), *Psychological and educational perspectives on LD* (pp. 87–131). Orlando, FL: Academic Press.

STOOLMILLER, M. (1995). Using latent growth curve models to study developmental processes. In J.M. Gottman & G. Sackett (Eds.), *The analysis of change* (pp. 105–138). Hillsdale, NJ: Erlbaum.

TEXAS READING INITIATIVE. (1997). *Beginning reading instruction: Components and features of a research-based reading program.* Austin, TX: Author.

TORGESEN, J.K., MORGAN, S.T., & DAVIS, C. (1992). Effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology, 84,* 364–370.

U.S. DEPARTMENT OF EDUCATION. (1979). *School volunteer development project.* Proposal submitted to the Program Effectiveness Panel of the National Diffusion Network. Washington, DC: Author.

U.S. DEPARTMENT OF EDUCATION, NATIONAL DIFFUSION NETWORK DIVISION. (1981). *Educational programs that work.* (8th ed.). San Francisco: Far West Laboratory for Educational Research and Development.

WASIK, B.A. (1998). Volunteer tutoring programs in reading: A review. *Reading Research Quarterly, 33,* 266–292.

WASIK, B.A., & SLAVIN, R.E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28,* 178–200.

STUDENT SELECTION

Teacher: _____ Grade & room: _____

Please consider these criteria for your students when making your selections for SMART:

> Lowest 25% of the class in reading skills
> Limited books in the home
> In need of one-on-one relationship with a caring adult
> Students not being served in other programs

Prioritize your list of students according to who could benefit the most from SMART. List up to 10 students in the space below, and give a brief explanation for your selection (e.g., improve reading skills, needs books, needs one-on-one time, etc.). Note any times that are particularly good or bad for your students to attend SMART.

SMART will try to serve as many students on your list as possible.

Student (grade)	Reason for selection	List the best times for your class to attend (half-hour slots)
1.		1.
2.		
3.		2.
4.		
5.		3.
6.		
7.		**Times not available**
8.		
9.		
10.		

STUDENT SELECTION

Teacher: _____   Grade & room: _____

Please select 4–6 additional students in your class who you believe are average readers. These students should not be the best readers in your class, but they should be solid readers who you believe will continue to make good progress in learning to read. For example, if you were to have low, middle, and high reading groups, these would be solid readers in the middle reading group.

Student (grade)
1.
2.
3.
4.
5.
6.

What the examiner says is in ***bold/italics***.

Rapid Letter Naming

I am going to show you some letters and I want you to tell me the names of the letters as quickly as you can. Start here (point to the first letter) ***and go across the page*** (point across the first row). ***Try to name each letter. If you don't know a letter, I'll tell it to you. Are you ready? Begin.***

- The child is given 1 point for each **letter or letter-sound** she or he states correctly.
- Mark a slash through each letter (or sound) said incorrectly. Put a bracket after the last letter the child says after 60 seconds.
- If the child takes more than 3 seconds to say a letter, tell the student the letter, and point to the next letter so the student continues.

Examiner copy:

D	N	b	H	f	i	m	O	A	R
s	E	W	y	L	T	c	X	g	K
B	F	o	j	a	S	p	r	U	e
M	z	K	C	t	q	n	J	P	x
u	G	Q	I	w	Z	I	v	Y	d
V	h								

Total correct _____ Seconds _____ (if less than 60)

Rapid Letter Naming
Student copy

D N b H f i m O A R

s E W y L T c X g K

B F o j a S p r U e

M z K C t q n J P x

u G Q I w Z I v Y d

V h

(continued)

Segment words into three phonemes

This time I will say a word, and you will tell me the sounds in the word. My turn. I can say the sounds in Mike. M—i—ke (pause 1 second between each sound). *Your turn. Say the sounds in Mike.*

- If the child is correct, go to item # 1 (soap).
- If child gets the first sound right say: *Yes the first sound in Mike is /M/. I'll say all the sounds in Mike. M—i—ke. Say all the sounds in Mike.*
- If the child is incorrect say: *I'll say the sounds in Mike. M—i—ke. Say the sounds in Mike.*
- For children who get the first sound correct or no sounds correct on the first practice item, administer the next two practice items: shop *Sh—o—p;* cat *c—a—t*
- Begin timing the student with the first test item.
- For the test items, write the parts the child says, or a + if all 3 parts are correct.
- If the child gets an item incorrect, tell the student the correct answer, but do not have the student repeat the correct answer.
- Score 1 point for each correct sound.
- Circle the number indicating how far the child got in one minute, but give all items!!!

1. soap _____	6. big _____
2. mom _____	7. fall _____
3. food _____	8. dad _____
4. gum _____	9. mud _____
5. ten _____	10. dog _____

Total segments correct   (out of 30)   Words segmented correctly   (out of 10)

Reading Fluency Test

Place the unnumbered copy in front of the student. Place the numbered copy in front of you—but shielded so the student cannot see what you record. Say these specific directions to the student for each passage:

When I say "begin," start reading aloud at the top of this page. Read across the page (demonstrate by pointing). Try to read each word. If you come to a word you don't know, I'll tell it to you. Be sure to do your best reading. Are there any questions?"

Say "Begin" and start your stopwatch when the student says the first word. If the student fails to say the first word of the passage after *three seconds*, tell the student the word and mark it as incorrect, then start your stopwatch. Follow along on your copy. Put a (/) through words read incorrectly. If a student stops or struggles with a word for *three seconds*, tell the student the word and mark it as incorrect. At the end of *one minute*, place a bracket ( ]) after the last word and say, *"Stop."*

(continued)

Scoring reading passages

Scoring the reading assessment is done by determining the number of words read correctly (WRC).

What is a "Word?"

Ex. 1.	cat	**TW = 1**
	read as:	
	"cat"	**WRC = 1**
Ex. 2.	I sat.	**TW = 2**
	read as:	
	"I sat."	**WRC = 2**

What is a "Correctly Read Word?"

Rule 1. *Correctly read words are pronounced correctly.* A word must be pronounced correctly given the context of the sentence.

Ex. 1.	The word "r-e-a-d" must be **pronounced** "reed" when presented in the **context** of:	
	He will *read* the book.	**WRC = 5**
	not as:	
	"He will *red* the book."	**WRC = 4**
Ex. 2.	The word "l-e-a-d" must be **pronounced** "led" when presented in the **context** of:	
	She picked up a *lead* pipe.	**WRC = 6**
	not as:	
	"She picked up a *leed* pipe."	**WRC = 5**

Rule 2. *Self-corrected words are counted as correct.* Words misread initially but corrected within 3 seconds are counted as correctly read. Write the abbreviation "SC" over the corrected word.

Ex. 1.	The river was *cold.*	**WRC = 4**
	read as:	
	"The river was *could*...(2 sec)...*cold.*"	**WRC = 4**
Ex. 2.	Matt cleaned the house *for* Mom.	**WRC = 6**
	read as:	
	"Matt cleaned the house *of*...(1 sec). cleaned the house *for* Mom."	**WRC = 6**

Rule 3. *Repeated words are counted as correct.* Words said over again correctly are ignored.

Ex. 1.	*Ted ran* swiftly.	**WRC = 3**
	read as:	
	"Ted ran...*Ted ran* swiftly."	**WRC = 3**
Ex. 2.	Sally saw *a* cat.	**WRC = 4**
	read as:	
	"Sally saw a...*a* cat."	**WRC = 4**

Rule 4. *Dialect or articulation.* Variations in pronunciation that are explainable by local language norms or individual speech difficulties are not errors.

Ex. 1.	They *washed* the car.	**WRC = 4**
	read as:	
	"They *warshed* the car."	**WRC = 4**
Ex. 2.	Let's go to the *park.*	**WRC = 5**
	read as :	
	"Let's go to the *pawk.*"	**WRC = 5**

Rule 5. *Inserted words are ignored.* When a student adds extra words, they are not counted as correct words or as reading errors.

Ex. 1.	Sue was happy.	**WRC = 3**
	read as:	
	"Sue was *very* happy."	**WRC = 3**
Ex. 2.	Kelly played the flute.	**WRC = 4**
	read as:	
	"Kelly played *a* the flute."	**WRC = 4**

What is an "Incorrectly Read Word?"

Rule 6. *Mispronounced or substituted words* are counted as incorrect.

Ex. 1.	The *dog* ate a bone.	**WRC = 5**
	read as:	
	"The *dig* ate a bone."	**WRC = 4**
Ex. 2.	Lynne has many *hats.*	**WRC = 4**
	read as:	
	"Lynne has many *hat.*"	**WRC = 3**
Ex. 3.	*He wanted* a new car.	**WRC = 5**
	read as:	
	"*She wants* a new car."	**WRC = 3**

Rule 7. *Omitted words* are counted as errors.

Ex. 1.	Mario climbed the *oak* tree.	**WRC = 5**
	read as:	
	"Mario climbed the tree."	**WRC = 4**
Ex. 2.	The king fought *with an alligator* in the moat.	**WRC = 9**
	read as:	
	"The king fought in the moat."	**WRC = 6**
Ex. 3.	Sewing is my favorite hobby. *I enjoy sewing dresses and suits.* What is your favorite hobby?	**WRC = 16**
	read as:	
	"Sewing is my favorite hobby. What is your favorite hobby?"	**WRC = 10**

Rule 8. *Hesitations.* When a student hesitates or fails to correctly pronounce a word within *3 seconds*, the student is told the word and an error is scored.

Ex. 1.	Mark saw an elephant.	**WRC = 4**
	read as:	
	"Mark saw an ...(3 sec)"	**WRC = 3**
	or read as:	
	"Mark saw an elll-eee ...(3 sec)"	**WRC = 3**

Rule 9. *Reversals.* When a student transposes two or more words, those words not read in the correct order are errors.

Ex. 1	Charlie *ran quickly.*	**WRC = 3**
	read as:	
	"Charlie *quickly ran.*"	**WRC = 1**
Ex. 2	Shelly bought a *beautiful sweater.*	**WRC = 5**
	read as:	
	"Shelly bought a *sweater beautiful.*"	**WRC = 3**

(continued)

Rule 10. *Numbers written as numerals* are counted as words and must be read correctly within the context of the passage.

Ex. 1.	*May 5, 1989.*	**WRC = 3**
	should be read as:	
	"May *fifth, nineteen eighty-nine.*"	**WRC = 3**
	not as:	
	"May *five, one nine eight nine.*"	**WRC = 1**
Ex. 2.	He was in grade 3.	**WRC = 5**
	should be read as:	
	"He was in grade *three.*"	**WRC = 5**
	not as:	
	"He was in grade *third.*"	**WRC = 4**

Rule 11. *Hyphenated words.* Each morpheme separated by a hyphen(s) is counted as an individual word if it can stand alone.

Ex.	Fifty-seven	**WRC = 2**
	Daughter-in-law	**WRC = 3**

Rule 12. *Hyphenated words.* If one or more of the morphemes separated by a hyphen(s) cannot stand alone, the entire sequence is counted as one word.

Ex.	re-evaluate	**WRC = 1**
	Bar-be-que	**WRC = 1**

Rule 13. *Abbreviations* are counted as words, and must be read correctly within the context of the sentence.

Ex. 1.	Dr. Adams received a promotion.	**WRC = 5**
	should be read as:	
	"Doctor Adams received a promotion."	**WRC = 5**
	not as:	
	"D-R Adams received a promotion."	**WRC = 4**
Ex. 2.	Jan lives on Fifth *Ave.*	**WRC = 5**
	should be read as:	
	"Jan lives on Fifth *avenue*"	**WRC = 5**
	not as:	
	"Jan lives on Fifth *a-v-e*"	**WRC = 4**
Ex. 3.	Jan lives on Fifth *Ave.*	**WRC = 5**
	also should not be read as:	
	"Jan lives on Fifth *ave*"	**WRC = 4**

(continued)

Ex. 4.	John watched *T.V.*	**WRC = 3**
	can be read as:	
	"John watched *tee-vee*"	**WRC = 3**
	or as:	
	"John watched *television*."	**WRC = 3**
Ex. 5.	John watched *television.*	**WRC = 3**
	should be read as:	
	"John watched *television*."	**WRC = 3**
	not as:	
	"John watched *tee-vee*."	**WRC = 2**

Efficient scoring procedures:

1. If students appear to understand the instructions following the administration of the first passage, the examiner need only point to the first word at the top of subsequent passages saying "**Begin**." It is not necessary to reread the instructions each time.

2. Don't begin timing until the student says the first word. If necessary, supply the first word, put a slash through it, and begin timing.

3. If you completely lose track of where a student is reading, discontinue the reading and begin another passage.

4. If a student skips an entire row, put a line through it and continue the passage.

5. Score reading probes immediately after administration.