

**SUSAN B. NEUMAN** is a professor in educational studies at the University of Michigan. She is Director of the Center for the Improvement of Early Reading Ability (CIERA). Her research interests include early literacy, parent involvement, and technology. She may be reached at the University of Michigan, 610 E. University, Room 1600, SEB, Ann Arbor, MI 48109, USA, or [sbneuman@umich.edu](mailto:sbneuman@umich.edu).

**DONNA CELANO** is a research associate in the Curriculum, Instruction, and Technology in Education Department at Temple University. She specializes in mass communications research, early literacy, and family involvement. She may be contacted at Temple University, 470 Ritter Hall, Philadelphia, PA 19122, USA, or by e-mail [dcelano@vm.temple.edu](mailto:dcelano@vm.temple.edu).

**JILL FITZGERALD** is professor of literacy studies at the University of North Carolina at Chapel Hill. She is currently an Associate Editor for *Reading and Writing Quarterly* and recently won the American Educational Research Association Outstanding Review of Research Award. She may be reached at the University of North Carolina at Chapel Hill, Peabody CB 3500, Chapel Hill, NC 27599-3500, USA, or [jfitzger@email.unc.edu](mailto:jfitzger@email.unc.edu).

**GEORGE G. HRUBY** is a doctoral student in Reading Education at the University of Georgia. He is also the current President and Program Chair of the Brain and Education Special Interest Group of the American Educational Research Association. His research interests center on adolescent literacy development within situativist, ethoecological, and socionaturalist theoretical frameworks. He can be contacted at the Department of Reading Education, 309 Aderhold Hall, University of Georgia, Athens, GA 30602-7125, USA, or by e-mail at [geohruby@mindspring.com](mailto:geohruby@mindspring.com).

**LEE GALDA** received her Ph.D. from New York University and is a professor of children's literature at the University of Minnesota. Her research interests focus on literature written for children, the ways in which children experience that literature, and how they respond to it. She teaches postbaccalaureate and master's courses in children's literature and language arts and doctoral courses on research in response and literature for children. She can be contacted at the Department of Curriculum and Instruction, 350 Peik Hall, University of Minnesota, Minneapolis, MN 55455, USA or by e-mail [galda001@tc.umn.edu](mailto:galda001@tc.umn.edu).

**RICHARD BEACH** is Wallace Professor of English Education at the University of Minnesota. He conducts research on response to literature, composition, and inquiry instruction. He may be reached at 359 Peik Hall, 159 Pillsbury Dr., S.E., Minneapolis, MN 55455, USA, or by e-mail at [rbeach@tc.umn.edu](mailto:rbeach@tc.umn.edu).

**THE 1999/2000 UNIVERSITY OF GEORGIA DOCTORAL SEMINAR IN READING EDUCATION** consisted of Steffanie Bowles, Barbara A. Bradley, Ronette Burnett, Elizabeth Carr Edwards, George Font, Michelle Francis, Alison H. Heron, Alicia A. McCartney, M. Kristiina Montero, Miri Park, Cathy R. Payne, Leslie Rush, Claire Henderson Smith, Katherine A. Dougherty Stahl, Don Tauferner, Patricia Waldrip, Jun-Chae Yoon.

**STEVEN A. STAHL** is a professor of Reading Education at the University of Georgia. He received his doctorate from the Harvard Graduate School of Education, under Jeanne S. Chall. Stahl has written extensively on issues in beginning reading, including the role of phonics. His research interests are centered around issues in reading instruction in the primary grades and the relationship between vocabulary knowledge and reading comprehension. He can be contacted at Reading Education, 309 Aderhold Hall, University of Georgia, Athens, GA 30602, USA.

**MICHELLE COMMEYRAS** is an associate professor of reading education at the University of Georgia. She received her Ph.D. in Education from the University of Illinois at Champaign-Urbana. She is interested in critical thinking and feminist theorizing as they relate to literacy education. She can be contacted at 309 Aderhold Hall, University of Georgia, Athens, GA 30602-7125, USA or by e-mail at [Michelle@coe.uga.edu](mailto:Michelle@coe.uga.edu).

# Can minimally trained college student volunteers help young at-risk children to read better?

Jill Fitzgerald

*University of North Carolina at Chapel Hill, USA*

The main purpose of the present study was to investigate at-risk first- and second-grade students' reading growth as they were tutored by minimally trained college students. Also examined were the nature of the growth and the effectiveness of minimally trained college student volunteers on the children's reading growth. The college students in this study were volunteer work-study students who were participating in the recent America Reads initiative sponsored by U.S. President Clinton. The children were first and second graders who were deemed at risk of failure by their classroom teachers and reading specialists in the schools.

In 1996 President Clinton established the America Reads Challenge to mobilize a million work-study college students as reading tutors for children in kindergarten through third grade. The initiative was begun with the understanding that children who do not learn to read well in the early grades usually continue to do poorly in subsequent grades (Clay, 1985; Juel, 1988; Stanovich, 1986), making the early years critical and pivotal in children's academic careers.

The advent of the America Reads initiative immediately provoked controversy in academic and practitioner literacy communities. The majority of prior efforts to provide extra literacy support for children in the early grades have been in-school programs that have involved one-on-one teaching by extensively trained professionals, with Reading Recovery as the prototype of such interventions (Hiebert, 1994; Shanahan, 1998). The popularity of

such programs has been fostered by the generally positive results obtained with Reading Recovery (DeFord, Lyons, & Pinnell, 1991). At center stage in the controversy then is the question of whether undergraduate work-study students with no or very little training in teaching reading could significantly impact young children's reading abilities (Wasik, 1998).

Two domains of prior theory and research are important to the present study—early reading development and tutoring effectiveness. With regard to research-based reading development theory, two early phases are Literacy Roots, which typically occurs gradually from birth through kindergarten, and Initial Literacy, which typically occurs in first grade (Chall, 1996; Fitzgerald & Shanahan, *in press*). The critical kinds of knowledge developed during Literacy Roots are *metaknowledge*, in particular, knowing about functions and purposes of reading and writing; *content knowledge*, about a wide array of subjects; knowledge of particular *text attributes*, especially phonological knowledge and graphemic knowledge; and *procedural knowledge*, or learning what texts are and how to manipulate them (Fitzgerald & Shanahan, *in press*). Research has documented that literate knowledge in this period develops concomitantly in reading, writing, listening, and speaking (Teale & Sulzby, 1986). One kind of learning stands out among the others as extremely important during the Literacy Roots phase; that is, development of phonological awareness, or the ability to hear and manipulate separate words, parts in words, and indi-

### Can minimally trained college student volunteers help young at-risk children to read better?

The study addressed at-risk first and second grade students' reading growth as they were tutored by minimally trained college students. The college students were volunteer work-study students participating in the recent national America Reads initiative. In all, 144 children received some amount of tutoring. Thirty-nine tutors used a four-part instructional lesson with the children. For main analyses, 64 children who received the full complement of tutoring sessions were compared to 19 who received fewer sessions. The main conclusions were: (a) Comparisons using a within-program control group showed that, on

average, children made statistically significant gains in instructional reading level that could be attributed to the tutoring. The average gain for children receiving the full term of tutoring was 1.19 grade levels during six months of tutoring. (b) The greatest impact of tutoring was in affecting children's ability to read words. Among the children who received the full term of tutoring, most of their instructional reading level growth occurred during the second half of the program. (c) Patterns of growth in instructional reading level were different for low- and high-gains groups of children.

### ¿Pueden estudiantes universitarios voluntarios, mínimamente entrenados, ayudar a leer mejor a niños en riesgo?

Este estudio se ocupó del progreso en lectura de estudiantes en riesgo de primero y segundo grado que tenían como tutores a estudiantes universitarios mínimamente entrenados. Los estudiantes universitarios eran trabajadores voluntarios que participaban de la reciente iniciativa nacional América Lee (America Reads). En total, 144 niños recibieron cierta asistencia. Treinta y nueve tutores usaron con los niños un tipo de lección de intervención didáctica consistente en cuatro partes. Para los análisis principales, 64 niños que recibieron las sesiones completas de tutoría fueron comparados con 19 que recibieron menos sesiones. Las principales conclusiones fueron: (a) Las comparaciones usando un grupo de control intra-

programa mostraron que, en promedio, los niños hicieron progresos estadísticamente significativos en el nivel de lectura que podrían atribuirse a la tutoría. El progreso promedio de los niños que recibieron la tutoría completa fue de 1.19 niveles de grado durante seis meses de tutoría. (b) El mayor impacto de las tutorías se produjo en la habilidad de los niños para leer palabras. Entre los niños que recibieron la tutoría completa, la mayor parte del progreso en lectura ocurrió durante la segunda mitad del programa. (c) Los patrones de crecimiento en nivel de lectura fueron diferentes para los grupos de niños de bajo y alto progreso.

### Können freiwillige, minimal vorbereitete college-studenten jungen kindern mit lernrisiken zu besserem lesen verhelfen?

Die Studie adressiert Lesesteigerungen von risikobetroffenen Schülern der ersten und zweiten Klasse, während sie von minimal vorbereiteten College-Studenten Nachhilfeunterricht erhielten. Die College-Studenten bestanden aus im Praktikum-Studium befindliche freiwillig mitarbeitende Studenten, die an der jüngsten amerikanischen nationalen Leseinitiative *Amerika Reads* teilnahmen. Insgesamt erhielten 144 Kinder eine gewisse Menge an Nachhilfe. Neununddreißig Tutoren nutzten einen aus vier Teilabschnitten bestehenden Unterricht mit den Kindern. Für die Hauptanalysen wurden vierundsechzig Kinder, die den vollen Verlauf an unterstützendem Nachhilfeunterricht vermittelt erhielten, mit 19 anderen verglichen, die weniger Unterricht erhielten. Die Hauptschlußfolgerungen ergaben: (a) Vergleiche unter Anwendung eines Internprogramms mit einer

Kontrollgruppe zeigten, daß die Kinder im allgemeinen statistisch bedeutende Fortschritte in der Leseanweisungsstufe machten, welches dem Nachhilfeunterricht zugeschrieben werden könnte. Der durchschnittliche Zuwachs bei Kindern, die über den vollen Zeitraum Nachhilfe erhielten, war 1.19 Zensurengrade im Laufe von sechs Monaten Nachhilfeunterricht. (b) Die größte Wirkung der Nachhilfe bestand in der Beeinflussung der Kinder in ihrer Fähigkeit, Worte zu lesen. Unter jenen Kindern, die über den vollen Zeitraum Nachhilfe erhielten, fand der größte Zuwachs aus der Leseanweisungsstufe während der zweiten Hälfte des Programmes statt. (c) Die Zuwachsmuster in den Leseanweisungsstufen waren bei Kindern in Schwach- und Hochleistungsgruppen unterschiedlich.

## 最小限の訓練を受けた学生ボランティアにより、危機に瀕した子供たちは読めるようになるのだろうか？

この研究は、最小限の訓練を受けた学生が学習を補助することにより、危機に瀕した一年生と二年生の読解力が向上したことを扱っている。学生は最近の全国的なAmerica Reads国民発案に参加するボランティアの作業研究の学生である。全部で144人の子供が一定量の学習補助を受けた。39人の学生が4部門からなる教材を用いた。補助学習を全て受けた64人の子供たちが一部を受けた19人の子供たちと比較されている。結論は、(a) プログラム内の統制群を用いた比較により、平均して、子供たちは補

助学習に割り当てられる教材のレベルで統計的に有意な進歩を遂げたことが分かる。補助学習の全期間を終了した子供の平均得点は半年の補助学習期間で1.19グレードレベルである。(b) 補助学習の最大の効果は子供たちの単語を読む力への影響である。全期間を終了した子供たちに関して、教科書読解レベルの向上はほとんどがプログラム後半において生じている。(c) 教科書読解レベルの向上のパターンは得点の高いグループの子供と低いグループの子供では差が見られる。

## Des étudiants bénévoles de collège universitaire ayant reçu une formation minimale sont-ils en mesure d'aider des enfants à risque à mieux lire ?

L'étude porte sur les progrès en lecture d'élèves de première et seconde année lors d'un tutorat d'étudiants de collège universitaire ayant reçu une formation minimale. Les étudiants étaient des étudiants qui s'étaient portés volontaires pour un travail pratique dans le cadre de la récente initiative « L'Amérique lit ». Cent quarante quatre enfants au total ont reçu une certaine quantité de tutorat. Trente neuf tuteurs ont conduit une leçon en quatre parties avec les enfants. Les analyses principales comparent 64 enfants ayant suivi la totalité des sessions de tutorat et 19 enfants ayant suivi moins de sessions. Les conclusions principales sont : a) Les comparaisons utilisant un pro-

gramme interne de groupe contrôle montrent que, en moyenne, les enfants ont fait des progrès statistiquement significatifs en lecture que l'on peut attribuer au tutorat. Les enfants ayant suivi le programme complet de tutorat ont progressé de 1,9 années après six mois de tutorat. b) Le tutorat a eu un impact maximum sur la capacité des enfants à lire des mots. Parmi les enfants ayant suivi tout le programme de tutorat, la majeure partie des progrès ont eu lieu au cours de la seconde moitié du programme. c) Le type de progression du niveau de lecture n'est pas le même selon que les enfants ont beaucoup ou peu progressé.

## Может ли минимально обученный студент колледжа помочь научиться читать детям группы риска

исследовались успехи учеников первого и второго классов, принадлежащих к группе риска, которые учились читать под руководством минимально обученных студентов колледжа. Эти студенты добровольно включились в новую общенациональную инициативу «Америка читает». Помощь в том или ином объеме получили 144 ребенка. 39 студентов-тьюторов проводили обучение по 4-частной схеме. Для основного анализа сравнивались 64 ребенка, прошедших полный курс обучения с тьюторами, и 19 детей, которые прошли его в усеченном виде. Основные выводы:

(a) В целом, сравнение с внутрипрограммной контрольной группой демонстрирует статистически значимые достижения в уровне чтения, которые можно отнести за счет обучения с тьюторами. За 6 месяцев работы средний балл прошедших полный курс повысился на 1,19. (б) Наибольший эффект достигается в умении читать слова. Среди детей, прошедших полный курс, основные улучшения в навыках чтения приходятся на вторую половину программы. (в) Динамика роста навыков чтения в группах, достигших низких и высоких результатов, была различна.

vidual sounds in words (Adams, 1990). However, although phonological awareness is critical to literacy development, research indicates that it can develop as a result of reading and writing growth (Adams, 1990; Tunmer & Nesdale, 1985), and that it therefore is not necessarily a precursor to literacy development.

In Chall's words, the essence of the next phase, Initial Literacy, is "learning the arbitrary set of letters and associating these with the corresponding parts of spoken words" (1996, pp. 15-16). *Metaknowledge* continues to be a key, but of a different sort in this stage than in Literacy Roots. Here, students are learning increasingly that readers and writers work together, that is, that each works on the premises of the other (Nystrand, 1989). Further, they learn about self-monitoring of their own meaning and word making. For example, monitoring strategies such as asking, "Does this make sense?" are essential for word reading and production (Clay, 1993). *Content knowledge* continues to grow. Key knowledge about *text attributes* is immeasurably important in this period. During this time, students are continuing to acquire grapheme awareness, and they are learning orthographic or morphological patterns in words. Children also are learning that (a) their knowledge of syntactically acceptable word orders learned previously through oral language applies in reading and writing, and (b) their understandings of acceptable word order can be used to help make sound attempts at new words as they read and write. Finally, they begin to develop *procedural knowledge* such as using strategic searches and learning to select and produce graphic and syntactic cues to read and write words and sentences (Clay, 1993; Schwartz, 1997).

Perhaps most important to the present study with regard to developmental theory about learning to read is that research has affirmed the centrality of word recognition in early reading (Ehri, 1991). Although understanding and response to what is read are always the predominant goals of reading, it is clear that the major work of early reading revolves around recognizing and pronouncing words. During Literacy Roots and Initial Reading, three phases of word recognition occur. In the beginning, word learning is predominantly logographic (Ehri, 1991). That is, children start to remember words using "strictly visual characteristics rather than letter-sound correspondences" (Ehri, 1991, p. 387). Next, children enter an alphabetic phase of phonologically recoding "spellings into pronunciations according to grapheme-phoneme correspondence rules" (Ehri, 1991, p. 396). This typically may begin toward the end of kindergarten or in first grade. Finally, children enter an orthographic phase when they attempt "instant analysis of words into orthographic units without phonological conversion" (Frith, 1985, p. 306).

Tutoring in various subject areas can be an effective strategy for improving student achievement, though the extent of effectiveness varies from program to program and across individuals (Shanahan, 1998). Similarly, it appears that across various subjects, tutors with more training and experience have a greater impact (e.g., Cohen, Kulik, & Kulik, 1982; Shanahan, 1998; Wasik, 1998; Wasik & Slavin, 1993). However, it is clear that not all children benefit from tutoring, with research results suggesting that tutoring is not especially effective with many low-achieving readers (e.g., Mathes & Fuchs, 1994; Shanahan, 1998).

There have been few direct studies of tutor training for reading instruction (Shanahan, 1998; Wasik, 1998). At least two recent studies do suggest that either *long-term training* of college students as tutors (Juel, 1996) or *intense and directed supervision* of community volunteers by graduates in reading education (Invernizzi, Rosemary, Juel, & Richards, 1997) can have an impact on children's reading progress.

In sum, considerable research has accumulated to support a clearly articulated theory of early reading development. Important keys to early reading progress involve development of phonological awareness and word recognition abilities. However, to date little is known about the extent to which reading tutors who have minimal training can significantly affect children's early reading progress. The complexities involved in helping children to acquire the necessary phonological awareness and word recognition strategies might lead to the expectation that only teachers with considerable understanding of theory of early reading and of the intricate web of word recognition mechanisms could suitably nurture early reading. Yet, volunteer tutoring programs are currently plentiful in the U.S. America Reads is a costly project with high visibility, but many literacy educators have argued that having work-study students tutor children is not the most effective use of federal monies. Instead, many advocate that the funds might have more impact on children's reading if they were spent in other ways, such as for training reading teachers in school systems.

It seems imperative to research various aspects of volunteer tutoring in reading programs. The present study was designed as a modest step in this direction, an assessment of the effects of one program, involving minimally trained college work-study students, on at-risk children's reading. To my knowledge, this study is among the first involving the America Reads initiative (Gambrell & Dromsky, 1999; Morrow & Woo, 1999).

The present study also is significant in that it provides an example of a unique alternative method for conducting research on field-based programs where traditional control groups may not be possible. This was the case in the present situation because the schools de-

clined the possibility of incorporating control groups into the design. Concern was expressed that teachers of the children in the control group might "look bad," even if they were performing well. Such concern is not unusual in educational programs. In a classic experimental-control comparison, the control group receives none of the experimental treatment. In essence, this is a test of how effective a treatment is compared with none of the treatment. The alternative control group design used in the present study builds on a method demonstrated in research on a volunteer program done by Invernizzi et al. (1997), involving a within-program control group format, where a group of children is compared with a group of similar children in the same program who received identical instruction, but less of it. The alternative design used by Invernizzi et al. (1997) is essentially a test of how effective a treatment is compared to a lesser amount of it, rather than none of it. This alternative within-program control group design is potentially a more stringent test of a treatment than the classic form of comparing something to nothing. If statistically significant outcomes arise, they clearly can be directly attributed to what is happening in the program in conjunction with the amount of treatment that is provided.

## Method

### Sites and participants

*Schools.* Children from four elementary schools and 40 different classrooms were tutored. A representative from each of four University of North Carolina at Chapel Hill Partnership Development School districts chose an elementary school. In each case except one, the school was chosen because it had either the lowest or second-lowest average reading scores in the district on the prior year end-of-grade assessments required by the state. In one case, the school was chosen because it was the closest to the university (approximately 20 minutes away). This school was a designated Title I school, as were the other three.

*Tutor and supervisor selection and demographics.* Work-study students who were interested in becoming tutors or supervisors under the America Reads initiative were screened through initial interviews with a coordinator responsible for program operations and with Literacy Studies faculty. The coordinator was a Reading Recovery teacher on paid reassignment from a local school district. The purpose of screening interviews was to ensure that all tutors and supervisors were enthusiastic, had a desire to work with young children, were willing to exert the necessary time and effort, would be cooperative, and had no prior criminal record. During the screening the basic

tutor and supervisor responsibilities were explained, required appearance and demeanor at the schools were explained, and questions were asked about the students' reasons for wanting to tutor or supervise and about their past experiences with children. Students were asked to sign a pledge regarding prior criminal history and to commit to complete attendance and promptness at training and tutoring sessions. Only one tutor interviewee was turned down.

Forty work-study students were selected as tutors. All except one completed the year. That one dropped out after approximately 2 weeks of training. Five graduate students were selected as supervisors, one of whom graduated midyear. Two additional supervisors were selected for the spring semester. A demographic questionnaire was administered to tutors and supervisors in September.

*Tutor gender and ethnicity.* Thirty-seven of the tutors were female, two were male. There was considerable ethnic diversity. Forty-six percent reported themselves as having African heritage, 40% as Anglo, 5% Native American, 3% Asian, 3% Latino, and 3% Indo.

*Tutor college level and major.* Eight percent were freshman, 45% sophomores, 21% juniors, and 26% seniors. Tutors reported a wide variety of declared or desired majors, including education (21%), communications or journalism (18%), sciences (18%), psychology (15%), sociology (5%), a policy area (5%), English (5%), economics (3%), and history (3%), with 7% undecided. Future plans were also varied, with a full 30% of tutors saying they'd like to teach or be involved in education in some way. Twenty-five percent planned to go to graduate school, 17% said something like they "just wanted to get a job," 11% planned to go to a professional school, and 17% were undecided.

*Tutor prior experience.* Many of the tutors (62%) had previous experience with some kind of tutoring. Of those who had tutored before, only one had tutored in reading or writing. Of those who had tutored, nine (39%) had worked with elementary-grade children. Forty-six percent of them had tutored for a year or less, with the remaining having tutored from 2 to 5 years. Aside from tutoring per se, a full 87% of the 39 tutors said they had worked with young children in some form (55% in day-care or baby-sitting situations; 39% in volunteer settings; and 6% in their own family, coaching, or counseling at camp).

*Tutors' reasons for joining America Reads.* It is noteworthy that the majority (67%) of the tutors were interested in becoming a tutor mainly because they wanted to help children in some way. Another 23% were involved because they wanted to teach or help in literacy efforts; and 10% indicated their primary reason was that the pay was good for a work-study student (tutors received ap-

proximately US\$1.50 more per hour than other work-study students) or that they weren't sure of their reasons.

*Supervisor gender, ethnicity, and college major.* Five supervisors were female, two male. Three reported having African heritage, three Anglo, and one declined to report. Two were working on a master's in library science, two a master's or Ph.D. in education, two a master's in guidance or school psychology, and one was a Ph.D. student in comparative literature. Two planned to teach in the future, two planned to become library/media specialists, one planned to become a school counselor, and two declined to answer.

*Supervisors' prior tutoring experience.* Four of the supervisors indicated that they had tutored before in reading, writing, or multiple subjects, with elementary children or with multiple ages. Aside from tutoring, five indicated they had worked with young children, all in situations such as camp counseling or field work for college classes. Only one had ever been a supervisor before. This was in a field component for her college course work.

*Supervisors' reasons for joining America Reads.* Two individuals said they wanted to be supervisors mainly because the experience would be relevant to their field of study, two said they wanted supervisory experience, one said the program looked like a good one to help advance literacy, and the others did not respond.

*Child selection and demographics.* The coordinator met with principals and teachers at each site in August to explain the program and describe the child selection procedures. Teachers were asked to refer children who they thought were among the poorest readers in their classes, but who also were not receiving additional special services such as those provided by Reading Recovery or resource room teachers. First graders were requested, but in some situations first-grade teachers did not refer enough children relative to available tutors, so children from another grade received services.

In all, 144 children received some amount of tutoring, though some of these were children who entered and then left quickly because they moved. One-hundred-thirty-one children were in the program long enough to have received and returned parent permission forms for inclusion in the research study. One-hundred-nineteen permission slips to use data were returned for a 91% return rate. Two parents who returned slips indicated they did not give permission. Of the 117 students whose parents gave permission for their inclusion in analyses, 4 were kindergarten children, 77 were first graders, 32 were second graders, and 4 were third graders. The children were in 40 different classrooms. Nineteen were from school A, 27 from school B, 36 from C, and 35 from D. In 3 schools, 10 children were referred who in fact were not reading below grade as measured by our pretests. The

principals or teachers initially did not want any of these 10 children dropped from the program, though some agreed to replacements later in the year.

No child missed more than two sessions while in the America Reads program. However, some children moved away from the school, others moved in and were placed in the tutoring program, and in some schools, children in the tutoring program were shifted in or out of the program by the school principal. In three of the schools Reading Recovery teachers were working with children who scored the highest among the lowest-scoring 25% of first graders, and these students were not included in the study. A few children attended America Reads tutoring while they were on a waiting list to go to Reading Recovery. After they entered the Reading Recovery program, they typically did not return to the America Reads program. No more than three children who exited from Reading Recovery unsuccessfully were subsequently placed in our tutoring program.

Main analyses were conducted comparing 64 children (47 first and 17 second graders) who received a full complement of sessions (50 sessions, from September through March ) to a within-program group of 19 children (15 first and 4 second graders) who received 6 to 12 continuous weeks of tutoring.

Since the targeted instructional group was first and second graders, and since there were very few kindergarten students ( $N = 4$ ) and third graders ( $N = 3$ ), the two latter groups were excluded from any further analyses. All children for whom permission was not granted were also deleted as well as the 10 children who entered the program reading at or above grade level. Following are demographics for the remaining 99 first and second graders. These demographics were obtained from the school with parent permission.

There were 41 females and 58 males. One was Asian; 30 were African American; 60 were Anglo; 5 were Latino; and 3 indicated multiple ethnicities. Information about mother's highest level of education ( $N = 55$ ) and whether the child received free or reduced-cost lunch ( $N = 88$ ) was not available for all children. The available data suggested that although some children receiving services were from very low socioeconomic backgrounds, the majority probably were not. Seven percent of the children's mothers for whom data were available did not have a high school diploma; 25% had completed high school; 38% had some college education; 24% had a bachelor's degree; and 5% had a master's or doctoral degree. Only 28% of the children received free or reduced-cost lunch.

### **Tutoring sessions**

The Literacy Studies faculty at the University of North Carolina at Chapel Hill (consisting of myself and

two others, James W. Cunningham and Dixie Lee Spiegel) designed the tutoring session format based on assumptions gleaned from previous theory and research as detailed in the introduction to this article. Significant assumptions about children's learning were that (a) children learn to read by reading in meaningful contexts; (b) the ultimate goal of reading is gaining and making meaning with text; (c) the major work of beginning reading is getting words; (d) phonological awareness, or the ability to hear separate words, chunks in words, and separate sounds in words, is an important corollary to early literacy learning; (e) reading and writing develop simultaneously, and learning in one informs learning in the other; and (f) learning to read is facilitated through interaction with a more knowledgeable other.

The tutoring session format was modeled after aspects of Reading Recovery lessons and a format used by Invernizzi et al. (1997). The format was designed to incorporate techniques that have been documented to be effective in research studies. There were four parts as follows: repeated reading of familiar text, word study, writing for sounds, and reading a new book. The first part of the session, rereading, has been shown to facilitate fluency (Samuels, Schermer, & Reinking, 1992), aid automaticity in word recognition (Samuels, 1979), and help comprehension (Dowhower, 1987; Rasinski, 1990).

The word study portion consisted of various activities and games designed to (a) explicitly teach sight words, (b) explicitly teach several word-getting strategies (using context to get words, structural analysis, and phonics), and (c) provide practice in learning sight words and word-getting strategies. Prior research suggests that *explicitly* helping children to analyze words and learn strategies may be especially beneficial to at-risk readers (e.g., Lysynchuk, Pressley, d'Ailly, Smith, & Cake, 1989; Pearson, 1984, 1985).

In the writing for sounds component, children were asked to write a sentence of their choosing. During this writing, a tutor could use various prompting activities designed to assist the child in developing phonemic awareness and knowledge of letters associated with sounds. Prior research suggests that the segmenting of speech and matching of letters to sounds required of young children as they begin to write short sentences is an excellent activity for developing phonemic awareness (Clay, 1985; cf. Invernizzi et al., 1997).

Last, a new book was presented to the child through guided reading. This part of the format was designed as a way for children to progress through increasingly difficult books using scaffolded instruction. Scaffolding provided by a knowledgeable other in material that has a few, but not too many, unknown words is a beneficial way to *move children's learning* (cf. Vygotsky, 1978).

Each child received two 40-minute tutoring sessions per week. These sessions were intended to supplement (not replace) classroom reading instruction. Tutors remained with the same child throughout a given child's stay in the program. Tutors were required to maintain brief lesson plans. Books for the sessions were selected from a set of leveled books often used in Reading Recovery (e.g., Wright Group books). Supervisors visited tutoring sessions at least once a week and also held 1-hour group meetings to address questions with their tutors once a month.

### **Tutor and supervisor training**

The university work study office representative interpreted the America Reads guidelines to indicate that a minimal amount of tutor training could be done. Consequently, we held a total of 33 hours of paid training across the fall and spring semesters. Supervisors attended the same sessions as tutors and were provided no additional training. Approximately 10 of the 33 hours were conducted exclusively by the coordinator and dealt primarily with logistical and practical considerations, such as how to talk with teachers and other considerations important to being a university representative working in a school. All tutors and supervisors were required to attend all sessions. Sessions were videotaped, and individuals who missed a session were required to watch the tape.

The three Literacy Studies faculty conducted the remaining training sessions. Twelve hours were held before tutoring began, and the remaining were spaced over the rest of the two semesters. The main components of the sessions were (a) an introduction to principles of emergent literacy; (b) an initial overview of the four parts of the lesson, which included a tape of a Reading Recovery teacher doing a lesson with a first grader; (c) discussion of each of the four parts; and (d) examples of how each of the four parts might be done. Methods of conducting the training sessions included using videotapes, lecture, small-group discussion, and demonstrations. Each session was conducted by pulling from material the three Literacy Studies faculty use in their undergraduate course on teaching reading and writing and then modifying the session for the tutors. Modifications included making sure that all jargon was explained and pared down to essentials, primarily emphasizing the techniques we wanted the tutors to use.

Some examples used in instructing the tutors and supervisors about each of the four lesson parts follow. First, for the session on rereading, the faculty representative talked about what it was, how it functioned in the large picture of what early readers should learn, why it was important, and what research revealed about its importance. Faculty then demonstrated a rereading by ask-



ing a student from the class to participate. At the end of the session, tutors met in small groups with their supervisors to select books for rereading for their children and to ask and answer questions.

Second, for the session on word study, four main ways of recognizing words were elicited from the tutors and supervisors by using a short story written with hieroglyphic-like graphemes. The participants read the story as the university professor showed them new words for each page. After the reading, the professor elicited from the class the various strategies and ways that they recognized and remembered the words in the story. She then grouped these ways into four basic mechanisms—sight words, use of context, use of structural analysis, and phonics. Each of the four word recognition techniques was then studied in turn. Important characteristics of each were discussed (e.g., which words to use, how many to select, when to teach which strategies), and a minilesson for teaching each was demonstrated by asking a small group of students in the class to participate. Handouts were given for minilessons for each of the four techniques. At the end of the sessions tutors met in small groups with their supervisors to plan for implementation of selected minilessons on strategies or techniques of their choice in their word study sections for lessons in the upcoming week. Finally, the whole group met to ask and answer questions.

Third, in a session on writing for sounds, the faculty representative first discussed the importance of using writing for developing certain reading skills. Then he demonstrated several examples of activities tutors might use with their children, for instance, how tutors might show a child several pictures and ask the child to select one to talk about. The child would then say a sentence about the picture to write. He showed how, if the child had difficulty with a word he or she should be able to write, the tutor could prompt the child to hear certain sounds in the word and think about the correct letter(s) to match the sound. At the end of the session, tutors met in small groups with their supervisors to plan implementation of at least one of the writing-for-sounds techniques demonstrated.

Finally, during the session on reading a new book, a shared reading lesson and a language experience lesson were discussed and demonstrated by asking students in the class to participate. Handouts were given for samples of shared reading and language experience lessons with first graders. Students met in small groups with their supervisors to plan at least one implementation of either a shared reading lesson or a language experience lesson for the following week. The whole group then met to ask and answer questions.

## Data collection

Children's reading level, various reading abilities, and attitude toward reading were measured. Information was also collected from tutors and supervisors. The data about children's reading and information from tutors and supervisors were collected at three time points, at the inception of the program (September), in the middle (December), and at the end (March).

In a 2-hour session, tutors were trained in how to administer all of the assessments. Refresher training took place prior to the midyear assessments and the final assessments. Immediately after each assessment, all protocols were examined by the supervisors and the coordinator for issues such as clarity in miscue marking and whether the tutor had asked the child to advance high enough in the materials. Then I scored and coded all running records, and a trained graduate assistant scored and coded everything else.

Reliability estimates for all measures were obtained by having a second person independently score and categorize a random sample of 20% of the protocols. This person was a clinical research instructor who had recently completed a Ph.D. in special education with a secondary emphasis in literacy. Independent scoring took place after both the graduate assistant and the clinical research instructor were trained about the scoring and categorizing.

## Child measures

Several measures were used with the children, and all were individually administered and repeated at all three time points, except an attitude toward reading measure, which was given only at the beginning and the end. Instructional reading level, including book level; ability to read words in isolation; knowledge of letter names; knowledge of sounds for letters in isolation; knowledge of letters for sounds in context; and attitude toward reading were measured.

*Instructional reading level* was assessed with the use of passages from the Bader Reading and Language Inventory (Bader & Wiesendanger, 1994), but following procedures outlined by Clay (1993) for running records. The Bader Reading and Language Inventory is an informal reading inventory. It was chosen in part because grade-level equivalents of passages of this published inventory are well established and because, unlike some other inventories, early reading levels are represented in both preprimer and primer passages, providing greater sensitivity to initial reading levels. The child read orally from increasingly difficult passages while the tutor marked the child's miscues (deviations from text) on a second copy of the passage. Then I coded the miscues, and percentage of words read accurately was calculated to help determine instructional level. Inability to success-

fully pass a preprimer-level passage was assigned a score of "0." A score of ".25" indicated that the child successfully passed the preprimer, but no higher passage. On average, in a normally functioning population, a first grader can read at the preprimer level by the end of November, or about one fourth of the way through the school year. A score of ".50" indicated that the child successfully passed the primer level, but no higher passage. On average, a first grader can read at the primer level by about halfway through the school year. A score of "1" indicated that the child successfully passed the first-grade passage, generally achieved, on average, by the end of the year. A "2" meant the child successfully passed the second-grade passage. Reliability for identifying miscues was .94, and for coding categories, reliability ranged from .88 to .98.

*Book level* was the Reading Recovery book level being used during tutoring sessions. At the beginning, middle, and end of tutoring, tutors reported these book levels. Levels 3 through 8 are approximately preprimer level; 9 through 12 are approximately primer level; 13 through 16 are typical levels for end of first grade, and 17–20 are typical for beginning of second grade (q.v., Modern Curriculum Press, 1996, as cited in Invernizzi et al., 1997).

*Ability to read words in isolation* was measured by asking children to read words from graded word lists on the San Diego Quick Assessment (LaPray & Ross, 1986). Scores were determined with the use of the test directions, except that as just described for the previous measure, the final score was "0" if the preprimer list was not passed, .25 if the preprimer list was passed, and so on. The reliability estimate for scoring this test was .97.

*Knowledge of letter names* was measured using part of Clay's (lowercase) (1993) Letter Identification subtest of the Observational Survey. Children looked at individual letters and said the name. The score was given as percent correct. Reliability was .98.

*Knowledge of sounds for letters in isolation* was measured by use of a second part of Clay's (1993) Letter Identification subtest of the Observational Survey. The child was shown letters and asked to say a sound the letter could make. The score was given as percent correct. Reliability was .94.

*Knowledge of letters for sounds in context* was measured using Clay's (1993) writing dictation task. The child wrote a short sentence as the teacher repeatedly said the sentence slowly. The score was given as the percent of sounds in words correctly represented. Reliability was .94.

*Attitude toward reading* was assessed using seven Likert item statements such as "How do you feel when it's time for reading in school" (see Appendix). The child looked at each question as the examiner read it aloud and then chose and circled one face from a continuum of faces with varying degrees of smiles to frowns. The score

was the mean across the items and could range from 1 to 5. Scoring reliability was 1.00.

### **Tutors' and supervisors' questionnaires**

At entry and at the end of the year, tutors and supervisors filled out questionnaires that provided demographic information (detailed previously) as well as participants' views on what they would (did) enjoy about their role in the program, what would be (was) most difficult, what strategies would be (were) most important in teaching reading, what materials would be (were) most useful, confidence in their role, and the main reason for confidence in their role. At the end, they were also asked about weaknesses and strengths of tutor training.

Questionnaire responses were coded. Where categories of responses were created, intercoder agreement reliabilities ranged from .70 to 1.00.

### **Focus group interviews**

Focus group interviews were conducted at the beginning, middle, and end of the year with 8 to 10 randomly selected tutors. All of the supervisors participated in parallel focus groups at the beginning and end of the year. The interviews gave participants the opportunity to talk in depth about problems they were having, positive events, all aspects of the program, and their training. The interviews were tape-recorded and later transcribed.

## *Analyses and results*

### **Preface: Attendance and treatment veracity**

Tutor attendance at training and tutoring sessions was outstanding. Occasionally a tutor was sick, but all missed sessions were made up at a later date. During assignment to the tutoring program, no child missed more than two sessions.

Supervisors regularly reported that all tutors were using the four-part lesson format, and a reading of selected tutorial transcripts of sessions also supported the belief that the tutors made a good faith effort to carry out the sessions in the prescribed way.

### **General indicators of children's reading levels at various time points**

It seems important to provide a general picture of the overall progress of the children in the America Reads program in advance of assessing specific research questions. This provides a context or backdrop against which the research questions may be addressed. Fifty-eight percent of the children in the full-term program ( $N = 64$ ) made gains of a half-year or more on instructional reading level (a gain of .50 or better). This is about the gain one

would expect in a normally functioning population across a 5- to 6-month period. It is important to keep in mind that the posttests were administered in March, which was the sixth month of a 9- to 10-month school year. Thirty-seven percent made outstanding gains, from 1 year to 4 years' growth over the course of the 6 months. Only 16% showed no gain, while 27% showed an approximate 3-month gain over the 6 months (a gain of .25).

Of the 99 children who met the criteria for inclusion in data analyses, 33% exited the tutoring program reading at or above grade level at the time of their exit. Note that this figure includes the children who received short-term instruction. Of those who were in the program less than the full term, only 24% exited reading at or above grade level.

Table 1 shows figures across time for various reading indicators for the first and second graders with permission for data inclusion. These figures should be taken *only* as general indicators of reading levels for children who were in the program at a particular time point. It is important to remember that each time point does not include the same group of children, and further, some children at later time points had only recently entered the program. Therefore, one cannot accurately assess actual growth from these figures.

The variable, Book Level, shows the Reading Recovery level of books tutors said they were using with children in their tutoring sessions. Children in the program later in the year were reading at higher levels than those in the program at the beginning of the year, and on average, the children in the program at posttest time were reading in primer-level books. Note also that while some children were reading in high levels, some were reading in very beginning levels. A similar picture emerged for reading instructional level except that when reading novel material for the posttest assessment, at the

end of the year children were, on average, actually reading at first-grade level, that is, above the level expected in March for a normally functioning population. However, note again the considerable variability, with some children's reading instructional levels very low and others' very high. Finally, children's attitudes toward reading were generally high at both the beginning and ending of the year.

### Was the tutoring program effective?

*Preliminary assessments.* To test for program effects on achievement level, children who received full-term tutoring were compared to children who received less than full-term tutoring (cf. Invernizzi et al., 1997). In this comparison children who received less tutoring served as controls for those who received more tutoring. All aspects of the treatment were the same for both groups except for length of treatment.

Children who entered the program at the beginning of the fall semester and stayed in the program throughout the year were designated as receiving high-level treatment. These children ( $N = 64$ ; 47 first graders and 17 second graders) received 25 weeks of tutoring. To ensure distinct groups, those receiving medium amounts of instruction were eliminated. Children ( $N = 19$ ; 15 first graders and 4 second graders) who received 6 to 12 weeks of tutoring were designated as receiving low-level treatment. In the low-level treatment group, 2 children (1 first grader, 1 second grader) received 6 weeks of tutoring, 6 (5 first and 1 second) received 8 weeks, 3 (first) received 10 weeks, and 8 (6 first, 2 second) received 12 weeks. No more than three of the low-level treatment group children had previously received Reading Recovery assistance. Of the 15 tutors who taught children in the low-level treatment group, 12 also taught children in the high-level treatment group. There were six tutors in the

**Table 1** Means, standard deviations, and ranges for selected reading measures at pre-, mid-, and posttest for all first and second graders with permissions for data inclusion

	Mean			SD			Range		
	Pre(N)	Mid(N)	Post(N)	Pre	Mid	Post	Pre	Mid	Post
Book level <sup>1</sup>	2.50(76)	6.46(87)	10.48(69)	2.06	3.26	4.88	1-10	1-16	1-22
Instructional reading level	.06(92)	.46(93)	1.17(75)	.13	.97	1.33	0-5	0-4	0-4
Attitude toward reading <sup>2</sup>	4.23(88)		4.11(74)	.68		.85	2-5		1-5

<sup>1</sup> Using Reading Recovery levels. Levels 3 through 8 are approximately preprimer, 9 through 12 are approximately primer level, and 13 through 16 are typical levels for end of first grade (cf. publisher Modern Curriculum Press, 1996).

<sup>2</sup> Possible score range, 1-5.

Note. Numbers of children vary due to differing numbers of children in the program at various times and due to missing data.

high-level group who had previous tutoring experience with elementary grade children, ranging from 1 to 5 years. Three of the same six tutors also tutored children in the low-level group, and there were no other tutors in the low-level group with prior tutoring experience with elementary-grade children.

To rule out other potential factors that might suggest that any arising statistically significant differences between groups would not indeed be due to the actual tutoring, several analyses were done. First, number of sessions was independent of the pretest measures of instructional reading level, letter name knowledge, knowledge of sounds for letters in isolation, knowledge of letters for sounds in context, and ability to read words in isolation (correlations =  $-.01$  to  $-.19$ ,  $p = .96$  to  $.09$ ). These results confirm the suitability of using length of tutoring as a basis for defining a within-program control group.

Second, it was possible that the low-sessions group was, by chance, composed of greater numbers of children associated with other risk factors such as poverty. If so, then statistically significant effects would not necessarily be attributable to programmatic treatment. To explore this alternative explanation, correlations were done between treatment group and several other variables: child's ethnicity; child's gender; whether the child received free, reduced-cost, or fully paid lunch; mother's highest level of education; whether the child received additional special

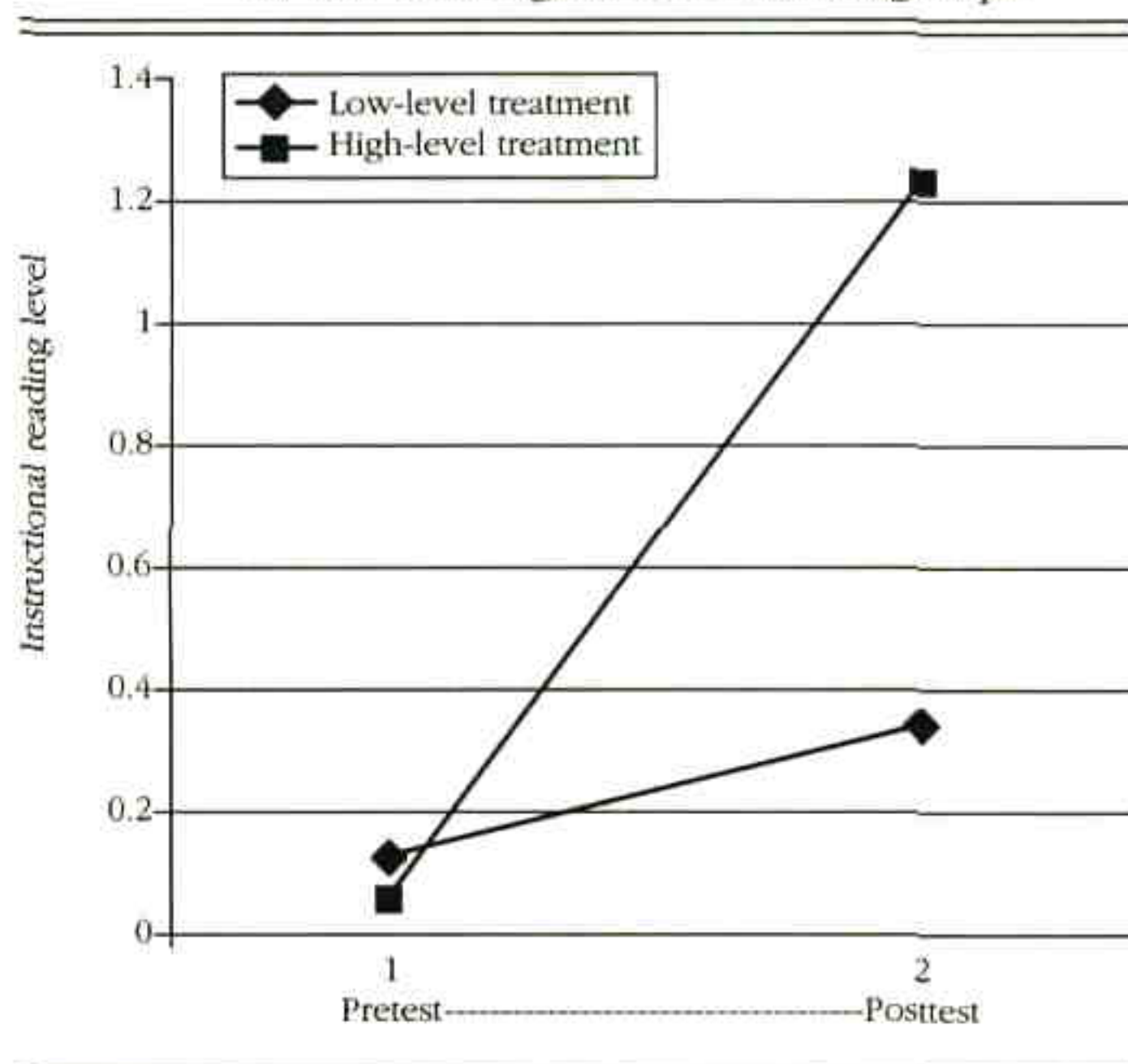
services; classroom teacher; and tutor. Only the correlation of gender with treatment group was statistically significant,  $\chi^2(N = 83) = .22$ ,  $p = .05$ . There were more boys in the group that received more tutoring. This result would not detract from the legitimacy of statistically significant effects, however, in that generally boys are more at risk for reading difficulty than are girls. Consequently, having more boys in the longer treatment group might be expected to work against greater growth, rather than working to facilitate it. Therefore, with the use of the selected variables, it appeared that the low-level treatment group was not composed of more at-risk children.

Finally, attribution of potential important differences between groups to an end-of-year growth spurt was ruled out. That is, the question was raised as to whether simply being in the program at the end of the year might reflect some sort of typical developmental growth spurt as opposed to an actual effect of tutoring. To address this issue the growth was looked at for five children in the low-level treatment group who were in the program at the end of the year. All entered reading at a  $.00$  or  $.25$  level, and none of these children showed an end-of-year growth spurt; their gain scores in instructional reading level ranged from  $.00$  to  $.25$ . Consequently, it did not appear that simply being in the program at the end of the year, regardless of amount of previous tutoring, could account for statistically significant outcomes.

In sum, the within-program control group made an excellent comparison group. The main statistical analyses were repeated measures analyses of variance. Each repeated measures model is described in detail in the following sections. The  $F$  ratios were calculated with the use of the multivariate method rather than using the potentially more powerful averaged univariate or mixed model method. That is, the  $F$  ratios generated for the models made no assumptions about the variance-covariance matrices as all tests of the variance-covariance matrices required for the mixed model method were significant.

*Program effects on achievement level.* Two sets of repeated measures analyses of variance were done to examine program effects on achievement level. First, a  $2 \times 2$  repeated measures analysis of variance was conducted with the use of instructional reading level as the outcome, with treatment level (low-level and high-level) as a between-subjects factor and time of test (pre- and posttest at the time of the child's entry and exit) as a within-subjects factor. There was a statistically significant treatment effect,  $MVF(1, 76) = 4.72$ ,  $p = .03$ . Overall, high-level treatment children outperformed low-level treatment children in instructional reading level. Table 2 shows the pretest and posttest means and standard deviations and mean gain for each group for pretest and posttest. There was a statistically significant time of test

**Figure 1** Pre- and posttest reading instructional level for low- and high-level treatment groups



**Table 2** Number of complete cases, means and standard deviations for reading measures at pre- and posttest by treatment groups

Measure	Number of cases	Pretest		Posttest		Gain
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Instructional reading level						
Low-level treatment	15	.08	.12	.37	.76	.29
High-level treatment	63	.04	.12	1.22	1.37	1.19
Total	78	.05	.12	1.06	1.31	1.01
Alphabet knowledge						
Low-level treatment	16	91.44	8.63	97.38	3.76	5.94
High-level treatment	57	92.51	9.31	98.75	2.24	6.24
Total	73	92.27	9.12	98.45	2.68	6.18
Sounds in isolation						
Low-level treatment	16	80.50	13.94	92.69	6.79	12.19
High-level treatment	57	75.65	24.82	86.60	20.03	10.95
Total	73	76.71	22.89	87.93	18.12	11.22
Sounds in context						
Low-level treatment	16	66.06	26.01	87.38	12.76	21.32
High-level treatment	57	56.84	31.06	92.49	8.90	35.65
Total	73	58.86	30.10	91.37	10.00	32.51
Reading words in isolation						
Low-level treatment	16	.16	.20	.36	.32	.20
High-level treatment	57	.12	.20	.82	.57	.70
Total	73	.13	.20	.72	.56	.59

Note: *N*s differ due to occasions where there were missing data for selected variables.

effect,  $MVF(1, 76) = 16.72, p < .01$ . Across both groups, on average, posttest instructional levels were higher than pretest reading instructional levels. Most important, there was a statistically significant interaction between treatment level and time of test,  $MVF(1, 76) = 6.27, p < .01$ . The children receiving more tutoring gained more (1.19) than the group with fewer sessions (.29). The effect size for this gain difference (using the high-level group mean gain minus the low-level group mean gain, divided by the low-group gain standard deviation) was large (1.29). The interaction is shown in Figure 1. This is particularly noteworthy since pretest instructional level reading scores were negatively correlated with number of tutoring sessions. That is, children who had lower pretest instructional levels and more tutoring outperformed those who had higher pretest scores but fewer sessions.

Second, a follow-up  $2 \times 2$  doubly multivariate repeated measures analysis of variance was done to explore in greater depth possible factors that might account for the instructional-level growth differences between low and high treatment groups. In this analysis, the four outcome measures were letter name knowledge, knowledge of sounds for letters in isolation, knowledge of letters for sounds in context, and reading words in isolation. Treatment level (low level and high level) was a between-

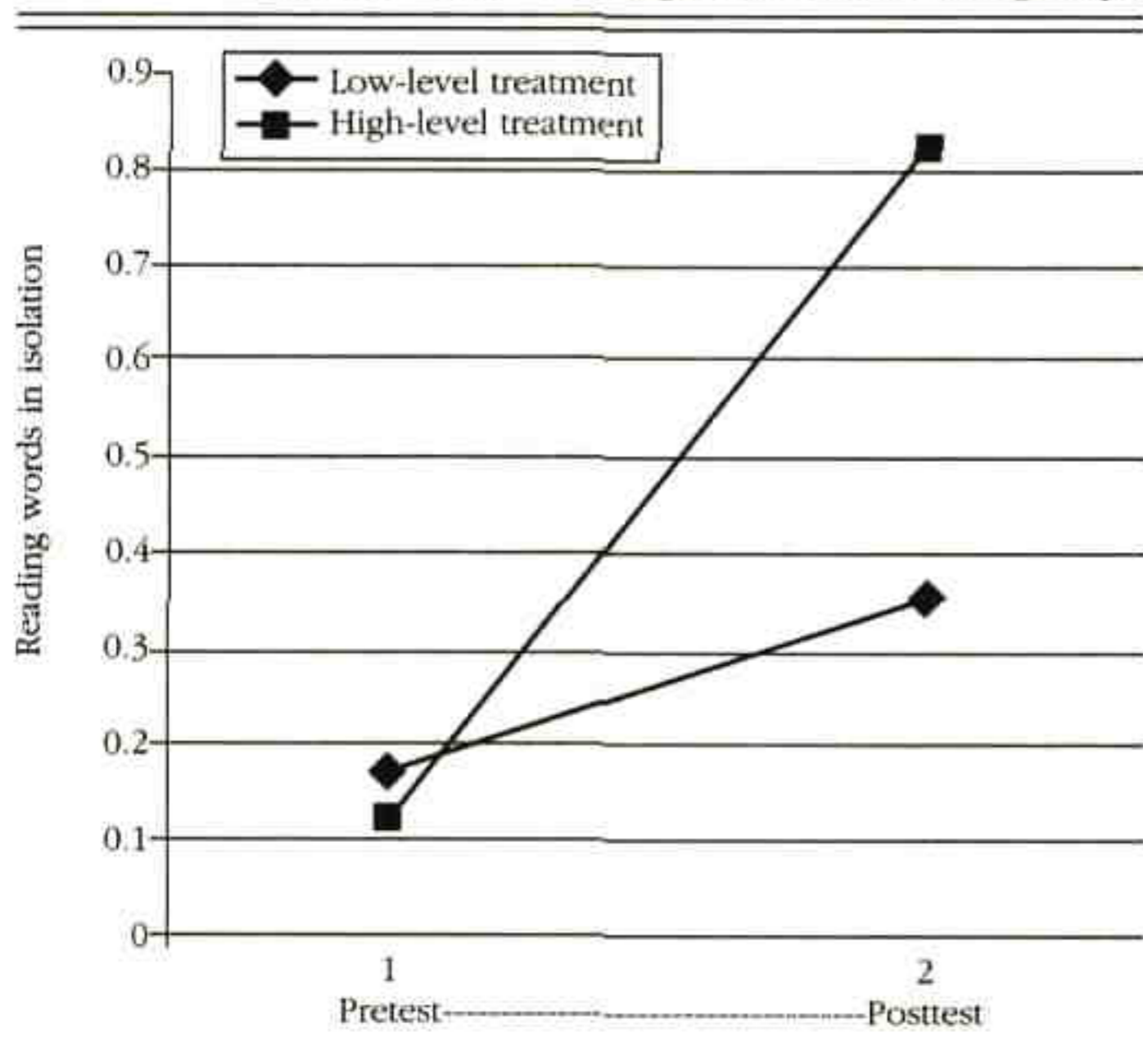
subjects factor, and time of test (pre- and posttest at the time of the child's entry and exit) was a within-subjects factor. All initial multivariate *F*s showed statistically significant effects when all four outcome measures were considered at once. There was a statistically significant treatment effect,  $MVF(4, 68) = 3.20, p = .02$ . In addition, there was a statistically significant time of test effect,  $MVF(4, 68) = 31.93, p < .01$ . There was also a statistically significant interaction between treatment and time of test,  $MVF(4, 68) = 5.43, p < .01$ .

The follow-up univariate results shed more light on the multivariate results. (See Table 2 for means and standard deviations.) There were statistically significant time of test effects for each of the four variables. Overall, children's posttest scores were higher than pretest scores for knowledge of letter names, knowledge of sounds for letters in isolation, knowledge of letters for sounds in context, and word recognition.

The remaining statistically significant effects showed that the greatest impact of tutoring was in affecting children's ability to read words. There was a treatment effect for reading words in isolation,  $F(1, 71) = 5.09, p = .03$ . On average, for reading words in isolation, those who received longer treatment had higher word reading abilities overall. Time of test was statistically significant for read-

ing words in isolation,  $F(1, 71) = 53.66, p < .01$ . Posttest scores were higher than pretest scores. Most important, there was a statistically significant interaction of treatment level with time of test,  $F(1, 71) = 16.42, p < .01$  for reading words in isolation. Children who received tutoring for a longer period made greater gains in reading words in isolation (.70) than did the other children (.20). (See Figure 2.) The effect size for the difference in gains was again large (3.13). This is a particularly noteworthy result because the pretest scores for reading words in isolation were negatively correlated with number of tutoring sessions. Thus, children who had lower pretest word recognition scores and a higher number of sessions outperformed children who had higher pretest word recognition scores but a lower number of sessions.

**Figure 2** Pre- and posttest reading words in isolation scores for low- and high-level treatment groups



The only other statistically significant univariate effects were for time of test for each of the three other outcomes—letter name knowledge, knowledge of sounds for letters in isolation, and knowledge of letters for sounds in context. In each case posttest scores were higher than pretest ones, for letter knowledge  $F(1, 71) = 21.78, p < .01$ ; for sounds in isolation  $F(1, 71) = 8.55, p = .01$ ; and for sound to letters in context  $F(1, 71) = 52.98, p < .01$ .

*What was progress like over time for the long-term group as a whole?* To assess the pattern of progress over time for the group of children who received the program for the entire 6 months, two repeated measures analyses were done for that group, with the use of their scores for the beginning, middle, and end of the program. In the first analysis, the within-subjects variable was time (with three time points) and the outcome measure was instructional reading level. The time effect was statistically significant,  $MV F(2, 61) = 24.64, p < .01$ , with planned contrasts showing statistically significant growth during the first half of the program,  $F(1, 62) = 9.83, p = .00$  and again during the second half of the program,  $F(1, 62) = 48.03, p < .01$ . Inspection of the means (see Table 3) showed that the greatest growth occurred in the second half of the program.

In the second analysis, a follow-up doubly repeated measures analysis of variance was done to probe growth patterns in specific reading variables. Time was again the within-subjects variable (beginning, middle, and end), and the four measures were alphabet knowledge, sounds in isolation, sounds in context, and reading words in isolation. There was a statistically significant multivariate time effect,  $MV F(8, 46) = 34.24, p < .01$ . There were statistically significant univariate results for time for each of the four measures as well. For alphabet knowledge the effect was  $F(2, 106) = 17.68, p < .01$ ; for sounds in isolation,  $F(2, 106) = 26.29, p < .01$ ; for sounds in context,  $F(2, 106) = 74.26, p < .01$ ; and for reading words in isolation,  $F(2, 106) = 92.84, p < .01$ . Planned contrasts showed that for each measure there was statistically sig-

**Table 3** Number of complete cases, means, and standard deviations for reading measures at three time points for children receiving high-level treatment

Measure	Number of cases	Pretest		Midtest		Posttest	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Instructional reading level	63	.04	.12	.35	.82	1.21	1.34
Alphabet knowledge	54	92.37	9.49	96.31	4.18	98.69	2.28
Sounds in isolation	54	76.22	23.75	87.43	10.06	95.28	7.02
Sounds in context	54	56.85	30.46	84.39	18.51	92.37	9.12
Reading words in isolation	54	.11	.17	.32	.28	.80	.49

**Table 4** Number of complete cases, means, and standard deviations for reading measures at three time points for children who made low versus high gains on instructional reading level

Measure	Number of cases	Pretest		Midtest		Posttest	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Instructional reading level							
Low gains	27	.04	.09	.05	.10	.19	.16
High gains	22	.08	.16	.82	1.26	2.77	.97
Alphabet knowledge							
Low gains	22	89.36	13.28	93.73	4.66	97.82	2.68
High gains	20	94.40	5.31	98.10	10.48	99.60	1.23
Sounds in isolation							
Low gains	22	73.77	28.21	83.45	10.15	93.23	7.81
High gains	20	79.65	20.57	90.10	7.13	97.65	3.80
Sounds in context							
Low gains	22	47.00	30.96	73.91	23.72	87.27	12.14
High gains	20	70.15	29.02	94.80	7.90	96.35	2.54
Reading words in isolation							
Low gains	22	.09	.14	.17	.21	.49	.25
High gains	20	.18	.20	.50	.29	1.18	.54

nificant growth from each time period to the next, range of *F*s (1, 53) = 11.70 to 102.17, with all *p*'s < .01.

Inspection of the means (see Table 3) showed three different patterns of growth. For alphabet knowledge and sounds in isolation, the growth across the first half of the tutoring (from pretest to midtest) was about the same as the growth across the second half (from midtest to posttest). Note, however, that for both of these measures, average scores were quite high at the beginning. These children, on the whole, entered the program with fairly solid knowledge of letter names, and they could say a sound for a letter in isolation quite well. On the other hand, for sounds in context, there appeared to be more growth during the first half of the tutoring as compared to the last half. The pattern for reading words in isolation was reversed, with apparent greater growth during the last half of the tutoring than in the first half.

Taken together, these results suggest that during tutoring the predominant learning was initially in the area of phonological awareness, specifically, in the ability to hear segments in words and in the area of knowing which letters to attach to those sounds. The predominant learning toward the end of tutoring was in the ability to actually read words. These results appear to affirm conclusions from prior research that an optimal level of phonological awareness is a necessary contingency for sight word development (Adams, 1990).

*What did the progress of highest gain children look like as compared with the lowest gain children?* To ascertain potential differences in patterns between children who made high gains and those who did not, two repeat-

ed measures analysis of variance were done. The data set of the 83 first and second graders who were in the low-level and high-level treatment groups was used to form higher and lower gain groups. Those who made no gain or a gain of .25 on instructional reading level comprised the low-gain group (*N* = 27). Those who made a full year or more gain comprised the high-gain group (*N* = 22). Others who made moderate gains were deleted so as to form two distinct groups.

The between-subjects factor was gain group (low, high), the within-subjects factor was time (pretest, midtest, and posttest), and the outcome measure was instructional reading level. There was, of course (because of the selection criteria for the two groups), an overall group effect,  $F(1, 47) = 78.81, p < .01$ . The time effect was statistically significant,  $MVF(2, 46) = 123.04, p < .01$ ; for growth from pretest to midtest  $F(1, 47) = 11.23, p < .01$ ; for growth from midtest to posttest,  $F(1, 47) = 182.44, p < .01$ . Most important, there was a statistically significant interaction of group with time,  $F(2, 46) = 97.11, p < .01$ . The growth patterns from pretest to midtest and midtest to posttest were different between the low- and high-gains groups. Table 4 shows the means and standard deviations. The high-gains group made strong growth in the first half of the program and phenomenal growth in the second half of the program. In contrast, the low-gains group showed almost no growth throughout the year.

To assess the extent to which the high-gains group's growth might be attributable to a children in second versus first grade, I examined the percentage of first

graders versus second graders represented in the high-gains group, and the entry instructional reading levels for all 22 children in the high-gains group. Twelve of the 62 first graders (19%) and 10 of the 21 second graders (48%) were in the high-gains group. Thus, proportionately more of the second graders than first graders made gains of 1 year or more. All of the 12 first graders' and half of the second graders' entry instructional reading levels were .00. Only three of the second graders began at level .25, and two began at .50. On the one hand, the five second graders who began at preprimer (.25) and primer (.50) levels were likely to be at more advanced word recognition phases than the other children. It is also possible that even the remaining five second graders who began at level .00, by virtue of their added year of exposure to reading, had somewhat more knowledge about sounds and word patterns than what is reflected in their instructional reading level scores. On the other hand, the relatively small numbers of second graders represented in the entire data set, as well as in this subset, make it difficult to know what to make of the differences between first and second graders' high gains.

To probe specific causes that might underlie the instructional reading level growth, a doubly multivariate repeated measures analysis of variance was done, using the four measures of alphabet knowledge, sounds in isolation, sounds in context, and reading words in isolation. The between subjects factor was again comprised of low versus high gains on instructional reading level, and the within-subjects factor was time (pretest, midtest, and posttest). There was a statistically significant overall difference between the low- and high-gain groups,  $MVF(4, 37) = 7.10, p < .01$ , and in follow-up univariate effects for groups for each of the four measures except sounds in isolation, for alphabet knowledge  $F(1, 40) = 8.91, p = .01$ ; for sounds in isolation,  $F(1, 40) = 3.40, p < .01$ ; for sounds in context,  $F(1, 40), p < .01$ ; and for reading words in isolation,  $F(1, 40) = 29.31, p < .01$ . Looking at the means for these four variables in Table 4, we see that the high-gains group maintained an edge over the other group at every time point, except on the sounds in isolation measure. Time effect was statistically significant overall,  $MVF(8, 33) = 29.42, p < .01$ . The differences held across all four measures, and for each case, there was statistically significant growth from pretest to midtest and then from midtest to posttest,  $F_s(2, 80) = 13.01$  to  $87.86, p_s < .01$ ; contrast test  $F_s(1, 40)$  ranged from 7.05 to 58.72,  $p_s$  from .00 to .01. The low- versus high-gain group by time interaction was also statistically significant,  $MVF(8, 33) = 5.63, p < .01$ . Follow-up univariate results showed that the effect was mainly attributable to reading words in isolation,  $F(2, 80) = 15.58, p < .01$ , with both contrasts statistically significant,  $F_s$

(1, 40) = 15.47 and 16.18,  $p_s < .01$ . That is, the patterns in growth were the same for the two groups for all measures except reading words in isolation (see Table 4). Essentially, for letter knowledge, sounds in isolation, and sounds in context, there was a tendency for each group to make slightly to considerably greater gains in the first half of the year as compared to the second half of the year (although the high-gains group always started higher and ended higher, on average). However, for reading words in isolation, whereas both groups began at similar levels, the high-gains group made more growth in the first half of the year than did the low-gains group. In the second half of the program, the gap widened.

These results continue to support the position that reading words in isolation tends to be dependent upon a relatively high development of phonological awareness. The high-gains group tended to be higher on these measures throughout, and their word reading knowledge exploded at the end of the program. The low-gains group tended to be lower on measures that most directly involved phonological awareness, but improved throughout the year. An additional possible interpretation is that the low-gains group might have been traveling the same trajectory as the high-gains group, but that they were moving more slowly.

#### **Tutors' and supervisors' perceptions of aspects of the America Reads program**

We analyzed the focus group interviews using procedures outlined by Coffey and Atkinson (1996) and Glaser and Strauss (1967). Interviews were transcribed and two individuals carefully read them all, making notes in the margins and looking for themes. Thematic categories were created, and instances of support for each theme were compiled. In the next round of analysis, the transcripts were read for nonsupporting examples or competing explanations. Following is a summary of the most salient themes from the focus groups.

*Program benefits.* All tutors and supervisors in focus groups felt that the America Reads program was an excellent initiative. At the end of the program, they spoke of the great benefits for children. One tutor put it this way: "I think that being one-on-one with him, it helped him see that he really could do it well. And by the time we were finished, he was reading [level] 18s and 19s and 20s [approximately early second-grade level], when I started him off in the 2s and 3s. I met his father, and he said, 'I really appreciate what you are doing, because I can see it at home.' Even his teacher would come up and say, 'He's doing so much better in class.' That made me feel good, because I was not only helping him with his reading, but with his behavior and even in other parts of his education."



Tutors and supervisors also felt it was a wonderful opportunity for college students to make a difference and give back to the community and children. They felt they had learned a lot about the complexities of learning to read and about the great need for one-on-one tutoring for some children.

Supervisors spoke freely of the strong commitment tutors made. Some talked about tutors who had done a tremendous job, especially those who worked with children who had behavior problems, "turning around" many children. Many tutors, they said, had "gone the extra mile." Supervisors felt that the tutors' positive attitudes were a big factor in the children's progress.

*Frustrations and concerns.* In general, tutors and supervisors felt the tutoring "went well," although some frustrations were expressed about some children's behavior, negative attitudes about reading, and fatigue. As the year progressed, however, all tutors in focus groups felt their frustrations eased.

Some tutors sensed that principals and teachers went out of their way to make a special effort to help them to feel welcome and comfortable. In these schools, the tutors thought their children saw and felt this support and that they felt special because they had a reading buddy from the university. Some tutors did not feel as welcomed at their schools, and they tended to think the atmosphere surrounding their visits also affected the children they worked with.

The tutors' biggest concern was that their students would not make progress. As one said, "It would really hurt me if they didn't really get to the levels they need when we leave. That's what I'm most worried about. [I want] to try to make sure that I do make the impact that I'm out to make." All of the tutors in focus groups were very ambitious, wanting to make life-changing differences in their students. Many expressed concern about changing children's attitudes about reading, that is, about helping them to be positive about reading and learning.

*Tutor-supervisor relationships.* Most tutors appreciated their supervisors' help. All remarked that the supervisors were pleasant and easy to talk with. Most liked having their supervisor visit them on site to answer questions and help with behavioral problems or other difficulties. At the end of the program interviews some mentioned that having a supervisor sit with them during tutoring sessions was helpful. Tutors who had supervisors with past experience in teaching reading were extremely positive about their supervisors, more so than tutors who had supervisors with less experience.

Most supervisors felt they had a good relationship with their tutors. However, there were problems and personality conflicts in a few cases. Most of the problems noted in this area were related to logistical concerns,

such as turning in paperwork on time and arriving for the car pool on time. Required small-group supervisor-tutor meetings went well for some supervisors, but others said they were "ineffective."

Supervisors nearly unanimously said it was hard to become a supervisor because their role was not clearly defined for them from the outset and because they were about the same age as the tutors.

*Training sessions.* For the most part, tutors thought training sessions gave helpful information. They especially liked the sessions that they thought gave them ideas and information they could use with students, such as a training session on games to use for word study and another on how to take and interpret running records.

However, there were aspects of the training sessions that both tutors and supervisors thought should be changed. They thought initial training sessions were too long and repetitive, that there was too much information to absorb in a small amount of time, and that much of the early information did not make sense until they started working with the children.

*Suggested changes.* Tutors and supervisors alike suggested several changes for training, including shorter sessions spread out more over the year, starting with just one part of the lesson structure and activities and ideas for the first day, and incorporating the training sessions into the tutoring more. Supervisors also suggested a separate session for them at the outset of the year to focus on their leadership role.

There was consensus about other suggested changes as follows: (a) find a way to communicate better with teachers and parents, (b) give tutors a list of the

**Table 5** Approximate program operations costs

	In-kind hours	Dollar amount
University program operation funds		US\$63,000
Grant monies expended		300
Work-study monies to pay tutors		45,346
Work-study monies to pay supervisors		37,900
School of education monies to pay supervisor		7,200
TOTAL		\$153,746
Faculty hours for planning and training	120	
Secretarial time	40	
Principal and teacher time	50	
Steering committee time	112	
TOTAL	322	
Approximate dollar cost per student, excluding in-kind contribution (for 144 children)		\$1,068

kinds of knowledge about reading that their students would be expected to know by the end of the year, (c) ensure quiet places to tutor, (d) have a formal introduction to the principal and teachers at their schools, (e) make sure that tutoring sessions are not scheduled during play times, (f) clarify supervisors' roles from the outset both for them and for the tutors, (g) set consequences for tutors' tardiness and similar problems, and (h) help all teachers and principals to know that the tutors are anxious to help their children, and find ways for all teachers and principals to welcome and support them.

### Program costs

It is difficult to accurately portray program costs since many individuals voluntarily contributed time. However, a rough estimation may be made. Table 5 shows that approximately US\$153,746 was spent on the program, and approximately 322 additional in-kind hours were contributed by various university and public school personnel. The approximate dollar cost per student, excluding in-kind contributions, was US\$1,068.

## Conclusions and discussion

The following paragraphs detail the main conclusions of the study.

1. On average, the at-risk first- and second-grade children in this study who were tutored by minimally trained college student volunteers improved their reading achievement. Comparisons using a within-program control group showed that, on average, children made statistically significant gains in instructional reading level that could be attributed to the tutoring. The average gain for children receiving the full term of tutoring was 1.19 grade levels.

2. On average children made positive gains in letter knowledge, knowing sounds in isolation, and knowing letters for sounds in context, but the greatest impact of tutoring was in affecting children's ability to read words. Further, among the children who received the full term of tutoring, most of their instructional reading level growth occurred during the second half of the program. When considering specific underlying reading abilities, the predominant learning during the initial tutoring was in the area of phonemic awareness and correct letter attachment to sound. The predominant learning toward the end of tutoring was in the ability to actually read words.

3. Patterns of growth in instructional reading level were different for low- and high-gains groups of children. The high-gains children, on average, made strong growth in the first half of the program and phenomenal growth in the second half of the program. In contrast, the low-gains children's growth almost flatlined throughout the year. When considering underlying reading abilities,

again, growth in reading words was the factor that distinguished the two groups' growth patterns.

4. Finally, tutors and supervisors perceived that the program was beneficial to children and to themselves.

This study provides evidence that minimally trained college student volunteers can help at-risk readers. This is important because, to my knowledge, this is one of the first controlled studies to provide such evidence, particularly in relation to the America Reads initiative (cf. Gambrell & Dromsky, 1999; Morrow & Woo, 1999). Moreover, the 1.19 grade-level gain made by the children receiving the full complement of tutoring is the amount of gain expected for normally functioning children, on average, in a 9- to 10-month school year. Consequently, to achieve this average gain in 6 months with a group of at-risk readers is remarkable.

Further, the effect sizes were very high for low- and high-level treatment group gain differences for instructional reading level and for reading words in isolation (1.29 and 3.13, respectively). These were considerably higher than those reported in other tutorials using paraprofessionals or volunteers, or both, and were comparable to several tutorials using professionally trained teachers (Invernizzi et al., 1997; Slavin, Karweit, & Wasik, 1994).

The results also support the possibility of the effectiveness of a tutorial using a balanced lesson design that aims to develop beginning reading abilities by stressing repeated reading, word study, learning about sounds and sound-letter relationships through writing, and scaffolded learning in increasingly difficult materials.

The pattern of children's progress over the year is consistent with results of prior research on emergent literacy, which suggest that phonological awareness is necessary for word learning (Adams, 1990). Also, the observed pattern for many children of learning more about sounds and letters during the first half of tutoring suggests compatibility with the belief that they were in an alphabetic phase (Ehri, 1991), and their increased actual word recognition abilities during the second half of tutoring suggests the possibility of movement into an orthographic phase (Ehri, 1991).

A related contribution of the present study to the literature is documentation of the importance of long-term tutoring to children's reading development. It is possible that, especially with at-risk early readers, reading development may occur in fits and starts. For example, it appeared in the present study that, on average, phonological awareness development and its concomitants required a rather lengthy early period of instructional support (in the first half of tutoring), but once instantiated, reading words as sight words and acquiring word-reading strategies blossomed (in the second half of tutoring). Consequently, continued tutoring seems critical.

With respect to the phenomenal second-half growth of the high-gains children, due to the relatively small numbers of second graders in the study it was not possible to ascertain the extent to which this growth might have revolved around some sort of optimal entry level related, for example, to being in a particular phase of early reading development. In future research it would be extremely helpful to include more children at the second-grade level in order to more directly assess grade-level effects on as many of the outcome variables and patterns in the data as possible.

Further, since low-gains children displayed a growth pattern different from the high-gains children, we might question whether these children were on the same trajectory, but slower paced, as the high-gains children or whether their developmental path was actually different. Though we cannot know with conviction, there was evidence that their yearlong growth resembled the first half-year growth of the high-gains children, supporting the same trajectory hypothesis.

Another important consideration is the cost of tutoring models such as the one used in the present study. The rough cost per student (US\$1,068), excluding approximately 322 hours of in-kind contributions, so far as it could be calculated, appeared in general to be less than other tutorials using highly trained specialists. For example, some have estimated the cost of Reading Recovery to be about US\$4,000 per child (Shanahan & Barr, 1995). Reading Recovery is an intensive one-on-one first-grade school reading program conducted by highly trained reading specialists. However, the cost per student in the present program was somewhat higher than that of the Book Buddies program (Invernizzi et al., 1997), a community volunteer program in which former graduate students in reading education assessed each child's needs, wrote daily lessons for tutors, and supervised them. The average cost per student across three cohorts in that program was US\$595. When considering the cost-effectiveness of reading tutorial programs, although the program in the present study was more expensive than some other volunteer tutorials have been shown to be, it is important to consider that the effect sizes achieved were more like those attained with programs using trained professionals, who represent a far greater expense to school systems.

It is important to remember that this is just one study, that there are limits to the study that must be considered as well, and that these limits suggest that the results should be interpreted cautiously, especially with regard to future policy decisions about similar initiatives. In one respect, the within-program comparison group provided a more rigorous comparison than some other types of control or comparison groups. In this design, all of the children received identical formats for tutoring. Only one

feature distinguished the two groups—amount of tutoring. Thus we know that statistically significant differences between the two groups are due to more tutoring of the sort that was offered. This is a design advantage over a typically used control or comparison in which children from mainstream classrooms are selected (often on a matched basis) for comparison. In a mainstream group comparison design, many other uncontrolled features may exist as competing explanations for statistically significant outcomes.

Also, finding statistically significant effects in a within-program comparison group may be especially important because it may be more difficult to find such significance when comparing a program to itself, so to speak. That is, if the tutoring is effective, then it may be more difficult to produce statistically significant effects when comparing more of it to less of it than when comparing it to a mainstream group of children who received no tutoring.

On the other hand, the design in this study did not reveal anything about the effectiveness of this particular kind of tutoring in comparison to another kind of tutoring, comparing minimally to maximally trained tutors, or comparing tutored groups to those with no tutoring. Further, we cannot know from this study which of the tutoring session components might most account for children's growth.

Other limitations to this study should also be taken into account when considering the meaning and practical significance of the tutorial model used in this study. For instance, long-term effects across more than one academic year are not known for the tutorial model used in this study. Also, the tutoring sessions were designed to address and ensure children's comprehension of what they were reading. However, the major work of beginning reading is getting words, and consequently, the measures used to assess children's progress reflected word learning and its concurrent or preceding factors rather than comprehension. Omission of a measure of reading comprehension may limit the generalizability of the work.

Finally, future research on tutorials such as the one used in the present study is needed to investigate several factors, including the following: (a) Which components of the tutorial contribute to its effectiveness? (b) What roles do social features of the tutorial play in children's progress? For example, does language that suggests a warm and caring tutor disposition tend to positively affect children's development more than on-task, take-charge directive language? (c) Does tutor experience play a role? For example, do children tutored by tutors who continue from year to year tend to do better than others? (d) To what extent does a supervisor's prior knowledge about literacy affect outcomes for children? (e) How do supervisor characteristics interact with tutor characteristics to enhance or detract

from children's progress? (f) What is the optimal amount of training and what are the essential characteristics of training that make a difference in student achievement?

## REFERENCES

- ADAMS, M.J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- BADER, L.A., & WIESENDANGER, K. (1994). *Bader reading and language inventory*. (2nd ed.). New York: Merrill.
- CHALL, J.S. (1996). *Stages of reading development*. Fort Worth, TX: Harcourt Brace.
- CLAY, M.M. (1985). *The early detection of reading difficulties*. Auckland, New Zealand: Heinemann.
- CLAY, M.M. (1993). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
- COFFEY, A., & ATKINSON, P. (1996). *Making sense of qualitative data*. Thousand Oaks, CA: Sage.
- COHEN, P.A., KULIK, J.A., & KULIK, C.C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, 19, 237-248.
- DEFORD, D.E., LYONS, C.A., & PINNELL, G.S. (Eds.). (1991). *Bridges to literacy: Learning from Reading Recovery*. Portsmouth, NH: Heinemann.
- DOWHOWER, S.L. (1987). Effects of repeated reading on second grade transitional readers' fluency and comprehension. *Reading Research Quarterly*, 22, 389-406.
- EHRI, L.C. (1991). Development of the ability to read words. In R. Barr, M.L. Kamil, P.B. Mosenthal, & P.D. Pearson, (Eds.), *Handbook of reading research*, Vol. II (pp. 383-417). New York: Longman.
- FITZGERALD, J., & SHANAHAN, T. (in press). Reading and writing relationships and their development. *Educational Psychologist*.
- FRITH, U. (1985). Beneath the surface of developmental dyslexia. In K.W. Patterson, J.C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301-330). London: Erlbaum.
- GAMBRELL, L.B., & DROMSKY, A.J. (May, 1999). *America Reads at the University of Maryland: Practices, evaluation, and implications*. Paper presented at the annual convention of the International Reading Association, San Diego, CA.
- GLASER, B.G., & STRAUSS, A.L. (1967). *The discovery of grounded theory*. Chicago: Aldine.
- HIEBERT, E.H. (1994). Reading Recovery in the United States: What difference does it make to an age cohort? *Educational Researcher*, 23(9), 15-25.
- INVERNIZZI, M., ROSEMARY, C., JUEL, C., & RICHARDS, H.C. (1997). At-risk readers and community volunteers: A 3-year perspective. *Scientific Studies of Reading*, 1, 277-300.
- JUEL, C. (1988). Learning to read and write: A longitudinal study of fifty-four children from first through fourth grade. *Journal of Educational Psychology*, 80, 437-447.
- JUEL, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly*, 31, 268-289.
- LAPRAY, M., & ROSS, R. (1986). The graded word list: Quick gauge of reading ability. In E.E. Ekwall (Ed.), *Ekwall reading inventory* (2nd ed., pp. 171-172). Boston: Allyn & Bacon.
- LYSYNCHUK, L.M., PRESSLEY, M., D'AILLY, H., SMITH, M., & CAKE, H. (1989). A methodological analysis of experimental studies of comprehension strategy instruction. *Reading Research Quarterly*, 24, 458-470.
- MATHES, P.G., & FUCHS, L.S. (1994). The efficacy of peer tutoring in reading for students with mild disabilities: A best-evidence synthesis. *School Psychology Review*, 23, 59-80.
- MORROW, L.M., & WOO, D.G. (1999, May). *The effect of a volunteer tutoring program on literacy achievement and attitudes of teachers, tutors, and children toward the program*. Paper presented at the annual convention of the International Reading Association, San Diego, CA.
- NYSTRAND, M. (1989). A social-interactive model of writing. *Written Communication*, 6, 66-85.
- PEARSON, P.D. (1984). Direct explicit teaching of comprehension. In G.G. Duffy, L.R. Roehler, & J. Mason (Eds.), *Comprehension instruction: Perspectives and suggestions* (pp. 223-233). New York: Longman.
- PEARSON, P.D. (1985). Changing the face of reading comprehension instruction. *The Reading Teacher*, 38, 724-738.
- RASINSKI, T. (1990). Effects of repeated reading and listening-while-reading on reading fluency. *The Journal of Educational Research*, 83, 147-150.
- SAMUELS, S.J. (1979). The method of repeated readings. *The Reading Teacher*, 32, 403-408.
- SAMUELS, S.J., SCHERMER, N., & REINKING, D. (1992). Reading fluency: Techniques for making decoding automatic. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about reading instruction* (pp. 124-144). Newark, DE: International Reading Association.
- SCHWARTZ, R.M. (1997). Self-monitoring in beginning reading. *The Reading Teacher*, 51, 40-48.
- SHANAHAN, T. (1998). On the effectiveness and limitations of tutoring in reading. In A. Iran-Nejad & D. Pearson (Eds.), *Review of research in education*, vol. 23 (pp. 217-234). Washington, DC: American Educational Research Association.
- SHANAHAN, T., & BARR, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly*, 30, 958-996.
- SLAVIN, R.E., KARWEIT, N.L., & WASIK, B.A. (1994). *Preventing reading failure: Research, policy, and practice*. Boston: Allyn & Bacon.
- STANOVICH, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-406.
- TEALE, W.H., & SULZBY, E. (Eds.). (1986). *Emergent literacy: Writing and reading*. Norwood, NJ: Ablex.
- TUNMER, W.E., & NESDALE, A.R. (1985). Phonemic segmentation skill and beginning reading. *Journal of Educational Psychology*, 77, 417-427.
- VYGOTSKY, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- WASIK, B.A. (1998). Volunteer tutoring programs in reading: A review. *Reading Research Quarterly*, 33, 266-291.
- 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, 179-200.

Received October 27, 1999  
Final revision received February 23, 2000  
Accepted February 28, 2000

## APPENDIX

### Items on the attitude toward reading measure

- 
- How do you feel about going to school?
- How do you feel when someone reads you a story?
- How do you feel when it's time for reading in school?
- How do you feel when you are asked to read out loud?
- How do you feel about how well you can read?
- How do you think you'll feel about reading when you're older?
- How would you feel if someone gave you a book for a present?
-