
Volume 47 Number 4 August 2009 ISSN 0022-4405
JOURNAL OF
SCHOOL
PSYCHOLOGY
Editor: Edward J. Daly III
ELSEVIER

ELSEVIER

Journal of School Psychology 47 (2009) 215–243

Journal of School Psychology

Brief experimental analysis of early reading interventions ☆

Anna-Lind Petursdottir *, Kristen McMaster, Jennifer J. McComas, Tracy Bradfield, Viveca Braganza, Julie Koch-McDonald, Roxana Rodriguez, Haley Scharf

University of Minnesota, United States

Received 20 November 2008; received in revised form 15 February 2009; accepted 23 February 2009

Abstract

The purpose of this study was to investigate how brief experimental analyses (BEAs) could be used to identify effective interventions for Kindergartners (2 girls and 2 boys, 5 years and 7–10 months old) with low performance and/or growth slope in letter sound fluency (LSF). Interventions were tested within a multielement design with brief mini-reversals until an intervention yielding at least 20% improvement on a specific subskill measure or a curriculum-based measure of LSF was identified. BEA-identified interventions were implemented one-on-one for 5 to 9 weeks. A multiple-baseline design across participants showed large intervention effects (average adjusted $d = 2.4$) on general outcome measures, supporting treatment validity of BEAs. Findings extend the BEA literature to younger participants, early reading interventions, and early reading measures.

Keywords: Reading; Early intervention; Experimental analysis; Response to intervention; Assessment; Letter sounds

☆ This research was conducted by the first author in partial fulfillment of the doctoral degree in educational psychology at the University of Minnesota. We thank the participants of the study as well as their parents and teachers. This research was supported in part by Grant #R305G040104 from the Institute for Education Sciences in the U.S. Department of Education to Vanderbilt University, the University of Minnesota, and the University of Texas — Pan American. The paper does not necessarily reflect the position or policy of the funding agency, and no official endorsement should be inferred.
* Corresponding author. Present address: School of Education, University of Iceland, Stakkahlid, 105 Reykjavik, Iceland. Tel.: +354 552 5335; fax: +354 525 5599.

E-mail address: annalind@hi.is (A.-L. Petursdottir).
ACTION EDITOR: Edward Daly.

0022-4405/$ - see front matter
doi:10.1016/j.jsp.2009.02.003

Brief experimental analysis of early reading interventions

In modern society, few skills are as essential for success in life as reading. Despite its importance, many children are not reaching adequate levels of reading proficiency in school. For example, in 2007, 33% of U.S. fourth graders were reading below the basic achievement level — or not reading well enough to achieve grade-level work (National Center for Education Statistics, 2007). Thus, a large proportion of elementary students are at risk for academic difficulties that could considerably limit their opportunities in life. Fortunately, effective early intervention can prevent reading failure, at least for a large proportion of struggling readers (Cavanaugh, Kim, Wanzek, & Vaughn, 2004; Coyne, Kame'enui, Simmons, & Harn, 2004).

Whereas a variety of interventions have been found to be effective for many at-risk readers, even the most successful interventions are not sufficient for a small proportion of students (e.g., Fuchs et al., 2001b). Around 12% of students have not responded well to evidence-based reading instruction in general education classrooms (Fuchs & Fuchs, 2005; McMaster, Fuchs, Fuchs, & Compton, 2005); 30% of at-risk students (Al Otaiba & Fuchs, 2002; Vellutino et al., 1996) and around 50% of children with disabilities have not responded well to early reading intervention (e.g., Fuchs et al., 2002).

Response to intervention

Response to Intervention (RTI) has been presented as an alternative to identification and instruction of students with learning difficulties (Fuchs & Fuchs, 1998). The defining features of RTI include identifying struggling learners as early as possible, monitoring their progress, providing increasingly intensive interventions when progress is not sufficient, and observing students' response to those interventions. Students who do not respond to increasingly intensive interventions may eventually become eligible for special education services. Supporters of RTI expect that it will provide a larger number of struggling readers with early intervention, better match instruction to individual needs, identify students with learning disabilities (LD) more accurately, and reduce the use of possibly stigmatizing labels (Fuchs, Mock, Morgan, & Young, 2003). Federal legislation supports these views. According to IDEA 2004 (U.S. Department of Education, 2004), local education agencies are no longer required to use an aptitude-achievement discrepancy to identify students with LD, but may use RTI as part of the evaluation process.

Standard treatment protocol versus problem solving approaches

Two main approaches to RTI have been discussed in the literature: the standard treatment protocol approach and the problem solving approach (Fuchs et al., 2003). In the standard treatment protocol approach, the same empirically validated intervention is used for all children with similar needs, making its implementation (including teacher training and fidelity assessment) easier than an individualized problem solving approach. The standard treatment protocol approach thus has the capacity to bring generally effective interventions to a large group of struggling readers. A study by Vaughn, Linan-Thompson, and Hickman (2003) exemplifies this approach. At-risk second graders received daily

35 min supplemental reading instruction, via a small-group standard treatment protocol. Students were assessed every 10 weeks; each time about 25% of the students reached exit criteria. However, after 30 weeks of intervention, 25% of the students still had not made sufficient progress to meet exit criteria. Thus, although standard treatment protocols may help the majority of at-risk students, they may not meet the instructional needs of all struggling readers.

The problem solving approach entails an inductive process that generally involves five steps (Deno, 1989): (1) the problem is identified; (2) the discrepancy between actual and expected performance is assessed; (3) possible solutions are explored and an intervention is selected; (4) intervention is assessed through progress monitoring, and if ineffective, other possible solutions are explored; and (5) when achievement goals are met, the intervention is terminated. A major assumption of this approach is that no one child characteristic reliably predicts which intervention will work best, and thus successful solutions must be induced by evaluating student responsiveness in a dynamic process involving teachers and specialists.

A recent report summarizing the current status of RTI implementation in the United States (Berkeley, Bender, Peaster, & Saunders, 2009) indicates that, whereas some states report using RTI models that are exclusively based on standard treatment protocols or problem-solving approaches, many states (20%) use a hybrid approach to RTI. Indeed, some researchers argue that RTI embodies problem-solving even when major components adhere to standard protocols, and that the two approaches are compatible (e.g., Batsche et al., 2005; Christ, Burns, & Ysseldyke, 2005). For example, a school-based instructional team might initiate a problem-solving approach by identifying problems and assessing student performance (Steps 1 and 2 above). Then, the student might be placed in a standard treatment protocol that presents an efficient, "best bet" intervention, because it has already been established through research to be likely to produce successful outcomes (Fuchs et al., 2003). If the intervention is deemed insufficient based on progress-monitoring data, the team might decide that a change is needed. They would then further analyze the problem to determine whether adjustments to the current intervention, or a different intervention, are/is needed (Step 3). They would continue to monitor progress to gauge the impact of these changes (Step 4).

In a problem solving model, the process of identifying an individualized intervention and assessing its effects requires systematic and efficient methods so that lost instructional time due to haphazardly selected, ineffective interventions is minimized. One relatively quick way to test which intervention might yield a positive student response is to use brief experimental analysis (BEA; Daly, Witt, Martens, & Dool, 1997). In a BEA, a single-subject design is used to briefly test instructional conditions affecting academic behavior (Daly et al., 1997). Depending on the number of interventions tested, the design employed, and the student's endurance, a BEA can be completed in matter of minutes or hours. This relatively brief process presents an efficient way to identify adjustments to current interventions, possibly avoiding the need to move to more intensive (and costly) interventions (Daly, Martens, Barnett, Witt, & Olson, 2007).

In a multielement design, each intervention is administered two or more times in semi-random order (e.g. Eckert, Ardoin, Daly, & Martens, 2002) and the intervention yielding the highest level of academic responding across sessions is deemed most promising. In a variant of brief multielement design with mini-reversals (Cooper et al., 1992), interventions are administered until a significant increase in the target behavior is observed, the baseline condition is re-implemented, followed again by the promising intervention, to establish a functional relationship between intervention and student performance. BEAs have most commonly been applied to oral reading (e.g., Daly & Martens, 1994; Daly, Martens, Hamler, Dool, & Eckert, 1999; Daly, Persampieri, McCurdy, & Gortmaker, 2005; Duhon et al., 2004; Eckert et al., 2002; Jones & Wickstrom, 2002; Noell, Freeland, Witt, & Gansle, 2001; Wagner, McComas, Bollman, & Holton, 2006), and have generally yielded differentiation in responding across interventions, identifying one as more effective than others for individual students.

Using brief experimental analysis to identify effective individualized interventions

Testing hypotheses

A key component to quickly and efficiently testing a student's response to interventions using BEA is to test hypotheses regarding the reasons for a student's lack of progress, such as: lack of motivation, lack of practice, need for additional assistance, need for an alternative approach/different materials, or need for easier material (Daly et al., 1997). These factors involve aspects of the learning environment that can be manipulated by the teacher and have been shown to be related to academic performance of students (see Daly et al., 1997). Using a BEA, the interventionist tests likely hypotheses by evaluating the effects of different instructional strategies (or combinations thereof) on a student's performance, until improvement is observed. To identify the most efficient intervention, the interventions can be tested in a hierarchy of those requiring least-to-most adult involvement (Daly et al., 1997).

Identifying interventions

Extending the use of BEAs to address the needs of struggling beginning readers within an RTI framework requires evidence-based, early reading interventions to serve as a foundation for individualization of instruction. Evidence-based instruction is critical because it helps to rule out the possibility that instruction is generally ineffective (see Vaughn & Fuchs, 2003; Vellutino et al., 1996). If instruction is generally effective, but an individual student shows little response to that instruction, it is reasonable to assume that the individual student may need more intensive or individualized intervention (Vellutino et al., 1996).

A substantial body of research emphasizes including explicit instruction in phonemic awareness, alphabetic principle, phonics, fluency, vocabulary, and comprehension instruction in early reading instruction (c.f., National Reading Panel, 2000). Thus, in extending BEAs to struggling beginning readers, it seems critical that interventions incorporate at least some of these elements, and that they are taught in an explicit, systematic way. In the present study, Kindergarten Peer-Assisted Learning Strategies (K-PALS, Fuchs et al., 2001b), a supplemental, classwide peer-tutoring program, served as the evidence-based instructional program. K-PALS was ideal for this study because it (a) incorporates many of the elements listed above, (b) has a substantial research base supporting its use with a wide range of beginning readers, and (c) provides a good example of a standard treatment protocol that schools might choose to implement with a broad range of students. See Fuchs et al. (2001b) for more detailed discussion of K-PALS effects.

In addition to evidence-based instruction, BEA requires additional strategies that can be employed to test hypotheses regarding a student's lack of progress, such as those listed above. For example, if "lack of motivation" is hypothesized to explain a student's low response to intervention, a strategy, ideally research-based, that focuses on improving motivation might be tested. Using the motivation example, research supports using goal setting, often paired with incentives, to improve student motivation (Morgan & Sideridis, 2006). Thus, adding a goal-setting with incentive component to an existing reading intervention might be appropriate to test using BEA (Duhon et al., 2004; Eckert et al., 2002; Noell et al., 1998, 2001).

Assessing intervention effects

In addition to strong interventions, BEAs require sensitive and independent measures to assess immediate and generalized effects of these interventions. Assessing immediate effects is important in determining whether an intervention seems promising enough to warrant extended implementation. Assessing generalized effects is critical for ensuring that the intervention will have a positive impact on long-term outcomes.

The most commonly employed dependent measure in BEA studies thus far has been oral reading (correct words read per min), in part because oral reading measures have been shown to be sensitive to immediate intervention effects, and also because they provide reliable and valid indices of overall reading proficiency (Deno, 1985; Fuchs, Fuchs, Hosp, & Jenkins, 2001; Wayman Wallace, Wiley, Ticha, & Espin, 2007). However, for beginning readers, oral reading is not likely to be sensitive to growth over brief periods, and smaller measurement units, such as letter sounds or words, may be called for (Good, Simmons, & Kame'enui, 2001). For example, researchers have suggested that letter sound fluency (LSF; correct letter sounds identified in 1 min) in Kindergarten is related to reading performance in later grades (Fuchs & Fuchs, 2004; Howe, Scierka, Gibbons, Silberglitt, 2003). In the only study that has included LSF as a dependent measure in a BEA, Noell et al. (2001) assessed the effects of different interventions on 2nd- and 3rd-graders' responding on their current grade level and two easier levels of materials. LSF showed differentiation in responding between baseline and intervention for two participants, but exhibited a ceiling effect for the third participant during both the BEA and extended analysis of intervention. Further research is needed to determine whether LSF can be used in a BEA to successfully identify promising interventions for struggling beginning readers.

For a BEA to yield meaningful results (differentiation in responding to interventions compared to baseline), assessment materials need to meet three criteria (Daly et al., 1997). (1) assessment materials used across interventions need to be of equal difficulty; (2) BEA measures need to be sufficiently different from each other (Daly et al., 1997) so that interventions do not affect performance on assessments in subsequent conditions, and (3) assessment materials need to have high content overlap with instructional materials to be sensitive enough to effects of brief intervention trials (Daly, Martens, Kilmer, & Massie, 1996). To identify measures that meet these three criteria, we considered two approaches to assessing beginning reading: general outcome measurement (GOM; e.g., Fuchs & Deno, 1991) and specific subskill measurement.

GOM involves long-term goal measurement, in which testing procedures are constant over long time periods (e.g., 1 year), providing teachers with an efficient, reliable, and valid way to determine the effectiveness of instruction (Fuchs & Deno, 1991). LSF is an example of GOM in that it involves equivalent probes that can be repeatedly administered to monitor beginning readers' progress over time (Fuchs & Fuchs, 2004). However, whereas LSF seems to be a viable alternative to oral reading fluency for BEA of beginning reading, it does not fully meet the above BEA criteria. For example, although LSF probes are of equal difficulty (Criterion 1; all LSF probes contain the 26 letters of the alphabet presented in random order), alternate forms containing the same 26 letters are not sufficiently different from each other (Criterion 2), making interference across interventions likely. Also, for a student with limited knowledge of letter sounds, the acquisition of one or two letter sounds during a short BEA intervention trial would not result in a large change in LSF (e.g., the student might not reach the part containing the newly acquired letter[s] during the 1 min timing). Therefore, LSF may not have sufficient content overlap with instruction (Criterion 3) to be sensitive to effects of intervention administered for the extremely short durations typical of BEAs, and thus may not be well suited for assessing the effects of brief early reading interventions during a BEA.

Specific subskill measurement differs from GOM in that it involves breaking down global curriculum goals into subskills with specific criteria for mastery (Fuchs & Deno, 1991). The teacher assesses student performance on one set of subskills and then, when mastery occurs, moves on to the next set of subskills. Specific subskill measures of letter-sound recognition, each containing a subset of letter-sounds to be mastered, may be more appropriate for measuring the effects of brief interventions. Specific subskill measures with subsets of letters can be created for each student, with a fixed proportion of known to unknown letter-sounds, ensuring equal difficulty (Criterion 1). By including different unknown sounds in each subskill measure, sufficiently different measures can be created (Criterion 2), while including the same known letter-sounds in all subskill measures keeps the content overlap between alternate measures constant. Finally, by applying each BEA intervention directly to subskill measures with different subsets of letters, high/complete content overlap between instructional and assessment material can be ensured (Criterion 3). Thus, assessment with specific subskill measures is likely to be sensitive to change while keeping the content overlap between alternate forms to a minimum. Research is needed, however, to determine which types of measures are most suitable to assess the effects of early reading interventions during BEA.

The main purpose of this study was to explore how BEA could be used to identify effective interventions for Kindergartners for whom classwide evidence-based early instruction was not sufficient to improve beginning reading skills, and to assess the treatment validity of BEAs. Research questions include: (1) How can BEAs be used to identify effective reading interventions for Kindergartners who are not responding to classwide evidence-based early reading instruction? Specifically, what procedures are needed, and what types of measures are most suitable to assess brief intervention effects? (2) Does implementation of a BEA-identified intervention increase early reading skills of Kindergartners not responding to evidence-based early reading instruction, as measured with general outcome measures?

Method

Participants and setting

Participants were drawn from a large-scale, multi-site study of Kindergarten Peer-Assisted Learning Strategies (K-PALS; Fuchs, Berends, McMaster, Sáenz, & Yen, 2006). The larger study included 46 kindergarten classrooms across three states, and involved an investigation of factors needed to successfully bring evidence-based reading instruction to scale. The present study included students from three of the participating classrooms in two public schools in a large Midwestern metropolitan area. K-PALS was implemented as part of the daily reading curriculum for all students in each classroom.

Classrooms 1 and 2 were in a school with a student population of 350 students, of whom 19% were African-American, 21% were Asian-American, 52% were Caucasian, and 7% were Hispanic American with 47% of students receiving free or reduced priced lunch. Classrooms 1 and 2 had 25 and 21 students, respectively. Classroom 3 had 21 students and was in a school with a student population of 350 students, of whom 64% were African-American, 12% were Asian-American, 3% were Caucasian, and 2% were Hispanic American with 94% of students receiving free or reduced priced lunch.

Selecting at-risk students

A Rapid Letter Sound pretest (see Measures section) was administered to all participants in the larger study in November of the study year. The lowest-scoring 20% of students in each classroom were considered at risk for reading difficulties (see Kaminski & Good, 2002) and were invited to participate in this study, with the exclusion of students receiving supplemental reading instruction or diagnosed with a disability to avoid possible confounds. Written parental consent and student assent were obtained for all participants.

Low responders to classwide K-PALS

Participants received 20 to 30 min of classwide K-PALS sessions, one to four times per week, for 10 to 15 weeks in late fall through early winter of the study year. Five weeks into the K-PALS program, we began monitoring students' beginning reading progress weekly for 5 to 9 weeks with three GOMs (described under Dependent variables section). Students' LSF progress was compared to a goal line drawn from their pretest performance in November to a benchmark goal of 35 correct sounds per min at the end of Kindergarten (recommended by Fuchs & Fuchs, 2004). Four students, one each from Classrooms 1 and 2, and two from Classroom 3, were identified as low responders to classwide K-PALS based on visual analysis of their connected LSF data showing performance levels furthest below the goal line and/or growth slopes judged to be too flat (or negative) to enable them to reach the benchmark goal by the end of the school year (see students' baselines in Figs. 5–9).

Leo was a Caucasian boy who was 5 years, 10 months old at the beginning of the study. His teacher was concerned about his lack of progress and attention problems. Leo's father reported a familial history of dyslexia and was concerned that Leo might also have dyslexia. Leo was reluctant to do reading activities; he required special encouragement and/or incentives to participate and stay on task. Prior to his selection for participation in this study, his teacher had attempted two modifications to K-PALS: she assigned an adult to assist Leo and his K-PALS partner, then assigned easier material to him when this adult assistance appeared ineffective.

Amy was a Caucasian girl who was 5 years, 7 months old at the beginning of the study. Her teacher was concerned about her modest progress and that her performance level was substantially below her classmates. Amy seemed to have very little confidence when it came to reading and often responded, “I don’t know” rather than attempting items on the measures.

Teri was an African-American girl who was 5 years, 8 months old at the beginning of the study. Teri’s teacher was concerned about her declining performance and reported that Teri was easily frustrated and sometimes shed tears during reading activities.

Jim was an African-American boy who was 6 years, 2 months old at the beginning of the study. His teacher was concerned about his lack of progress and his disinterest in reading activities. He easily got frustrated during reading activities and actively avoided them.

All progress monitoring, BEA, and intervention sessions took place in a one-on-one setting, in a relatively quiet environment near students’ classrooms. This included a resource room across the hallway from Classrooms 1 and 2, and in the hallway outside of Classroom 3.

Measures

Screening

The Rapid Letter Sound measure (Fuchs et al., 2001b) was used to screen students for inclusion in this study. This measure contains all 26 lower case letters of the alphabet in random order in five rows. Students were instructed to say the letter sounds as quickly as they could in 1 min; the score was the number of sounds correctly identified. The most common sound for each letter was counted as correct. If the student finished within 1 min, the score was prorated by dividing the number of correct sounds by the number of seconds it took the student to finish, and multiplying by 60.

General outcome measures (GOMs)

GOMs were used to identify BEA participants from the pool of at-risk students, to monitor progress during BEAs, and to assess generalized effects of BEA-identified interventions. GOMs included letter-sound fluency (LSF) and word identification fluency (WIF; both developed for the larger K-PALS study) and the DIBELS Nonsense Word Fluency measure (NWF, Good & Kaminski, 2002).

LSF consisted of 20 alternate forms. For progress monitoring, two forms were individually administered weekly using the same standardized procedures as the Rapid Letter Sound measure (see in Fuchs et al., 2001b), with the mean serving as the score. During BEAs a single administration provided the score. NWF consisted of 10 alternate forms including 50 nonsense VC (vowel-consonant) or CVC words, arranged in 10 rows and was individually administered with standardized procedures (Good & Kaminski, 2002). The score was the number of correct letter sounds produced during 1 min. WIF consisted of 14 probes containing 23 words randomly selected from a pool of 138 high-frequency and decodable words in the K-PALS program. The WIF was administered by asking the student to read the words as quickly and correctly as possible for 1 min and to sound out any unknown words. When a student paused for 3 s, the examiner prompted the student to try the next one. To be counted as correct, a word had to be pronounced correctly with all of its letter sounds blended together. The number of correctly read words in 1 min was recorded, and the mean of two probes was the score.

Specific subskill measures

To assess the effects of brief interventions during BEAs, individualized subskill measures of letter sound and decoding fluency were created for each participant, controlling for content overlap between alternate forms and difficulty level. Each letter-sound subskill measure (LSM) contained 56 single letters with the same proportion of two known to two unknown sounds based on each participant's LSF progress monitoring baseline performance. Each decoding subskill measure (DSM) contained 32 decodable words (real and nonsense) made from the same letters as the corresponding LSM with each of the two known sounds occurring eight times and each of the unknown sounds occurring 20 times. The CVC, VC, and CV words were arranged into eight rows, with each word occurring four times in random order. Alternate versions of LSM and DSM contained different unknown letters but the same known letters (usually "a" and "m"), resulting in 29% content overlap. A complete description of the process for creating subskill measures can be obtained from the first author.

Administration of the LSM was essentially the same as LSF with the exception that if the student finished attempting all letter sounds within 1 min, he/she was prompted to start again from the beginning. Administration of the DSM was similar to that of NWF and WIF. The student was asked to read the real and nonsense words as quickly and carefully as possible, or say the sounds of individual letters. If the student paused for 3 sec, he/she was prompted to try the next one. If the student finished within 1 min, he/she was prompted to start again from the beginning. The score on both subskill measures was the number of correct letter sounds in 1 min.

Dependent variables

Correct sounds per minute (CS/M)

The primary dependent variable was a student's fluency in producing letter sounds (the most common phoneme) in isolation and in decodable real and nonsense words during 1 min timed probes using LSF, NWF, LSM and DSM. If the student finished within 1 min, the score was prorated by dividing the number of correct letter sounds by the number of seconds it took the student to finish, and multiplying by 60.

Correct words per minute (CW/M)

The number of words read correctly in 1 min on the WIF was recorded; the score prorated if the student finished within 1 min.

Interobserver agreement and fidelity

Interobserver agreement (IOA) was checked on 34% to 53% of sessions within each phase (live or from audiotapes) and calculated based on exact agreement on the correctness of each letter sound, letter name, or word, by dividing agreements with agreements plus disagreements, and multiplying by 100 to obtain a percentage. IOA checks conducted on 39% of data collected across BEA conditions averaged 99% (range, 77–100%). IOA checks conducted on 37% of all probes during baseline and intervention phases averaged 96% (range: 78%–100%).

Fidelity of classwide K-PALS implementation was assessed with a checklist of specific student and teacher behaviors that should occur during K-PALS (e.g., Fuchs et al., 2006) and averaged 87.3% for both teacher and student implementation combined (range, 85% to 90%) in classroom 1, 75% (range, 73% to 77%) in classroom 2, and 80% (range, 75% to 85%) in classroom 3. Fidelity of 44% of BEA conditions, assessed with checklists corresponding to scripts of assessment and instructional strategies used with each participant, averaged 98% (range, 87% to 100%). Fidelity of 28% of BEA-identified interventions, assessed with checklists corresponding to the intervention scripts, averaged 94% (range, 91 to 98%).

Procedure

Screening

In November, trained staff administered the Rapid Letter Sound measure (Fuchs et al., 2001b) to participants one-on-one in a quiet environment as part of a test battery for the larger study (see Fuchs et al., 2006).

K-PALS

All students in participants' classrooms worked in reciprocal Coach-Reader pairs on early reading K-PALS tasks (Fuchs et al., 2001b; Fuchs & Fuchs, 2005) after activities had been modeled for the class by the teacher. Activities entailed practice of letter sound correspondence, reading common sight words, decodable words, sentences and short books. For a complete description of K-PALS activities, see Fuchs et al., (2001a).

Progress monitoring

Trained staff collected data with LSF, NWF, and WIF for 5 to 9 weeks during baseline and 5 to 10 weeks to assess effects of BEA-identified interventions. On all GOMs, if students finished within 1 min, scores were prorated by dividing the number of correct responses by the number of seconds it took the student to finish, and multiplying by 60.

BEAs

Instructional strategies were tested in a hierarchy from those requiring least to most adult involvement, to identify an effective and efficient intervention. Selection of strategies also took into account teacher reports indicating that certain interventions were more or less likely to be effective. Each test started with a baseline condition where a student's performance was assessed with the subskill measures (LSM and DSM) and the GOM (LSF). Then, an instructional strategy was applied for 10 min (5 min to letters and 5 min to words) to the same set of subskill materials as used for the preceding baseline assessment. Immediately following, the student's response to the strategy was assessed with the corresponding LSM and DSM (same version as used for baseline and instruction) and a new form of the LSF. BEA sessions ranged from 6 to 25 min, depending on strategies involved. Typically, one baseline and one strategy session were conducted during each meeting.

Strategies tested under the BEA conditions included goal setting with incentive, modeling, and modeling plus goal setting with incentive. In the goal setting with incentive condition, students were offered a choice of rewards (e.g., small toys, pencils, candy) contingent on improving their baseline scores on subskill measures or LSF by 20%. In the modeling condition, the examiner modeled correct responses to each letter and word, guided the student's response, and prompted the student to respond independently (e.g., Simmons & Kame'enui, 1989). For each word, the examiner modeled segmenting (saying each sound of the word slowly) and blending (combining the sounds into the word), guided the student to do each step with her, and then prompted the student to perform each step independently ('model-lead-test'). In the modeling plus goal setting with incentive condition, the examiner explained that after practicing the material together, the student would do the tasks independently and have the opportunity to choose a reward by improving the preceding baseline score by 20%.

Instructional strategies were tested one by one until at least a 20% increase in performance from the immediately preceding baseline condition on either the LSM or LSF was observed. The 20% criterion was based on Carnine, Silbert and Kame'enui's (1990) recommendation that a 'Repeated Reading' intervention should increase reading fluency by 40% to be considered effective, with the rationale that half of that increase observed during a short BEA condition could be considered promising (Noell et al., 2001). Then, a mini-reversal was attempted by repeating the baseline condition (A) followed by another test of the instructional strategy (B). If responding in the ABAB phases was clearly differentiated, with lower responding observed in both baseline conditions and better responding in both intervention conditions, a functional relationship between the instructional strategy and dependent variable was assumed and the BEA terminated. If differentiation was unclear (i.e. due to overlap between conditions), baseline and intervention conditions were implemented again, to replicate and confirm the effect. The number of required replications varied between participants, resulting in different numbers of data points per condition. If a mini-reversal was not achieved, other instructional strategies in the hierarchy were tested. The strategy yielding consistent performance at least 20% higher than baseline was considered promising for a particular student and selected for intervention.

Training of tutors

Graduate students in special education and school psychology received individual training from the first author for up to 2 h in K-PALS and BEA-identified interventions through modeling, role play, practice, and feedback until reaching at least 90% accuracy assessed with checklists corresponding to the intervention scripts.

Implementing BEA-identified interventions with K-PALS materials

Promising instructional strategies identified in BEAs were applied to K-PALS materials. Three tutors took turns working with participants, and each tutor worked with two to four students. Interventions were applied to the latest and/or next K-PALS lesson targeted in the classroom during 12- to 42-min (depending on strategies, with interventions involving modeling taking more time) one-on-one sessions 3.5 times per week on average (ranging from one to four sessions per week due to participants' absences, field trips, and other typical school interruptions). Participants' performance was recorded on tutoring log forms to keep track of attendance and progress over time. BEA-identified interventions were provided for 5 to 10 weeks, until students reached the LSF goal line or showed a steep trend of progress.

Design and data analyses

A brief multielement design with mini-reversals (Cooper et al., 1992) was used to test the effects of one or more instructional strategies on early reading skills during the BEAs. A multiple-baseline design across participants was used to assess the generalized effects of the BEA-identified interventions implemented over time on early reading skills.

Growth slopes were calculated by running linear regression for each GOM using weeks as the independent variable and GOM scores as the dependent variable (see Vellutino et al., 1996). The slope reflected the rate of change along the regression line across weeks. The standard error of the slope estimate (SE_b) was calculated by dividing the standard error of the estimate (SEE) by the product of the standard deviation of x (time in weeks, S_x) and the square root of the sample size ($\sqrt{n}$). In other words, the formula used for calculating the standard error of slope estimate was: $SEE/(S_x * \sqrt{n})$ (Cohen & Cohen, 1983).

Two indices of effect size of BEA-identified interventions were calculated for each GOM. First, the *No Assumption effect size* (NAES; Busk & Serlin, 1992), recommended for statistical analysis of single-subject data (Olive & Smith, 2005; Manolov & Solanas, 2008), was calculated by dividing the difference between each participant's baseline and intervention phase means by the standard deviation of the baseline. Second, an adjusted NAES was calculated based on Rosenthal's (1994) formula, $(X_{\text{intervention}} - X_{\text{baseline}})/(SD_{\text{pooled}}/\sqrt{(2\,(1-r))})$, taking into account the autocorrelation between repeated measures, to put it on the same scale as Cohen's d (Riley-Tilman & Burns, 2009) and thus allow comparisons to findings across other research studies. The difference between the means of the last three measures of baseline and intervention phase was divided by the second half of the formula in which SD_{pooled} is the standard deviation of the three last measures of baseline and intervention phases pooled and r is the correlation between the last three measures of baseline and intervention phases. The last three measures in each phase were used to calculate the adjusted NAES (see Swanson & Sachse-Lee, 2000).

Results

Results are presented for BEAs first, followed by progress monitoring data evaluating the generalized effects of the BEA-identified intervention for each participant.

Using BEAs to identify promising early reading interventions

Results of participants' BEAs are presented in Figs. 1–4, with each top graph showing LSM performance, each middle graph showing DSM performance, and each bottom graph showing LSF performance.

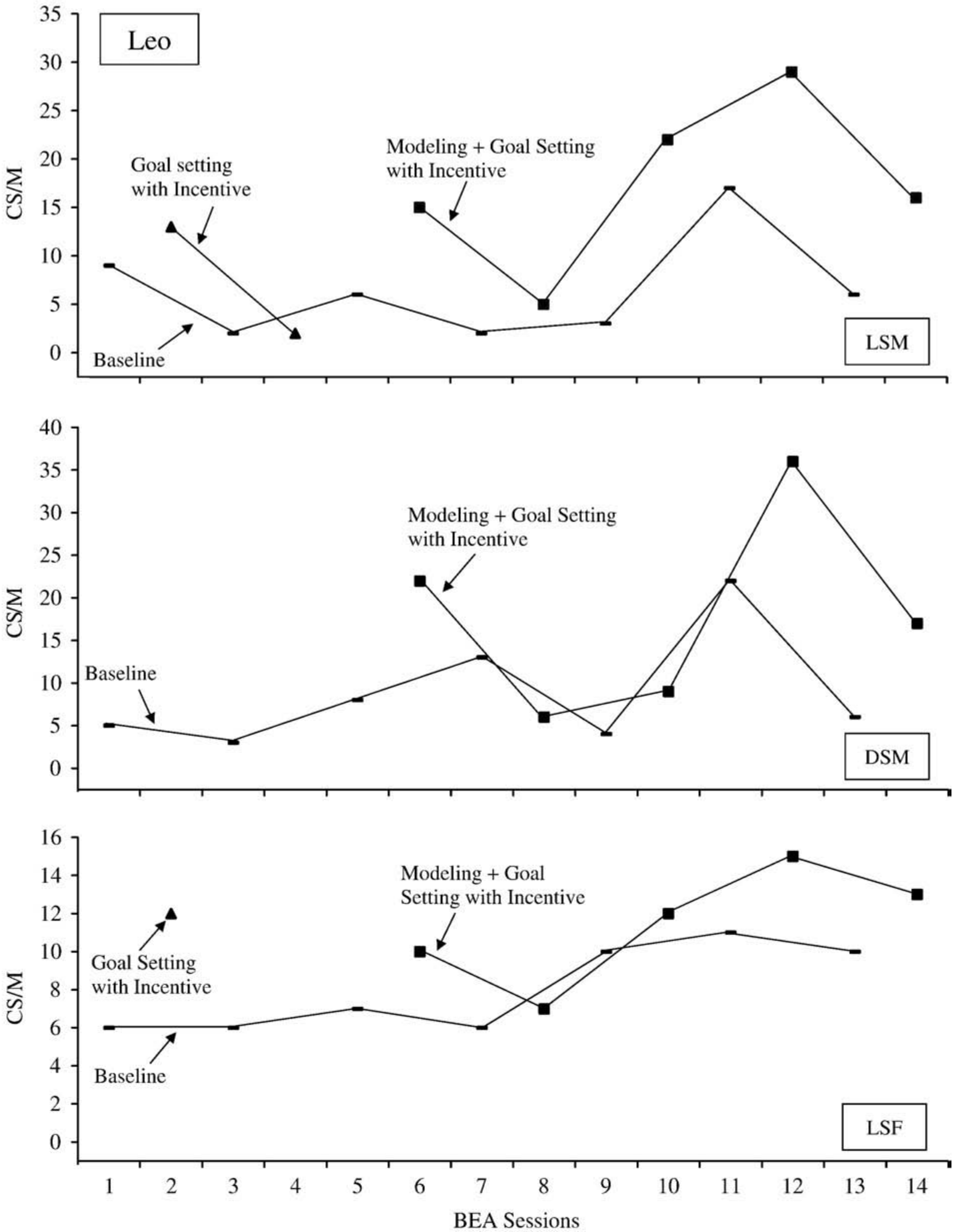


Fig. 1. Leo's correct sounds per min (CS/M) on the letter-sound subskill measure (LSM), decoding subskill measure (DSM), and LSF across different BEA conditions.

Leo

Results of Leo's BEA are presented in Fig. 1. Given teacher's and data collector's reports of Leo requiring incentives to participate in reading activities, it was decided to only test strategies involving incentives with Leo, thus skipping modeling alone. Leo's

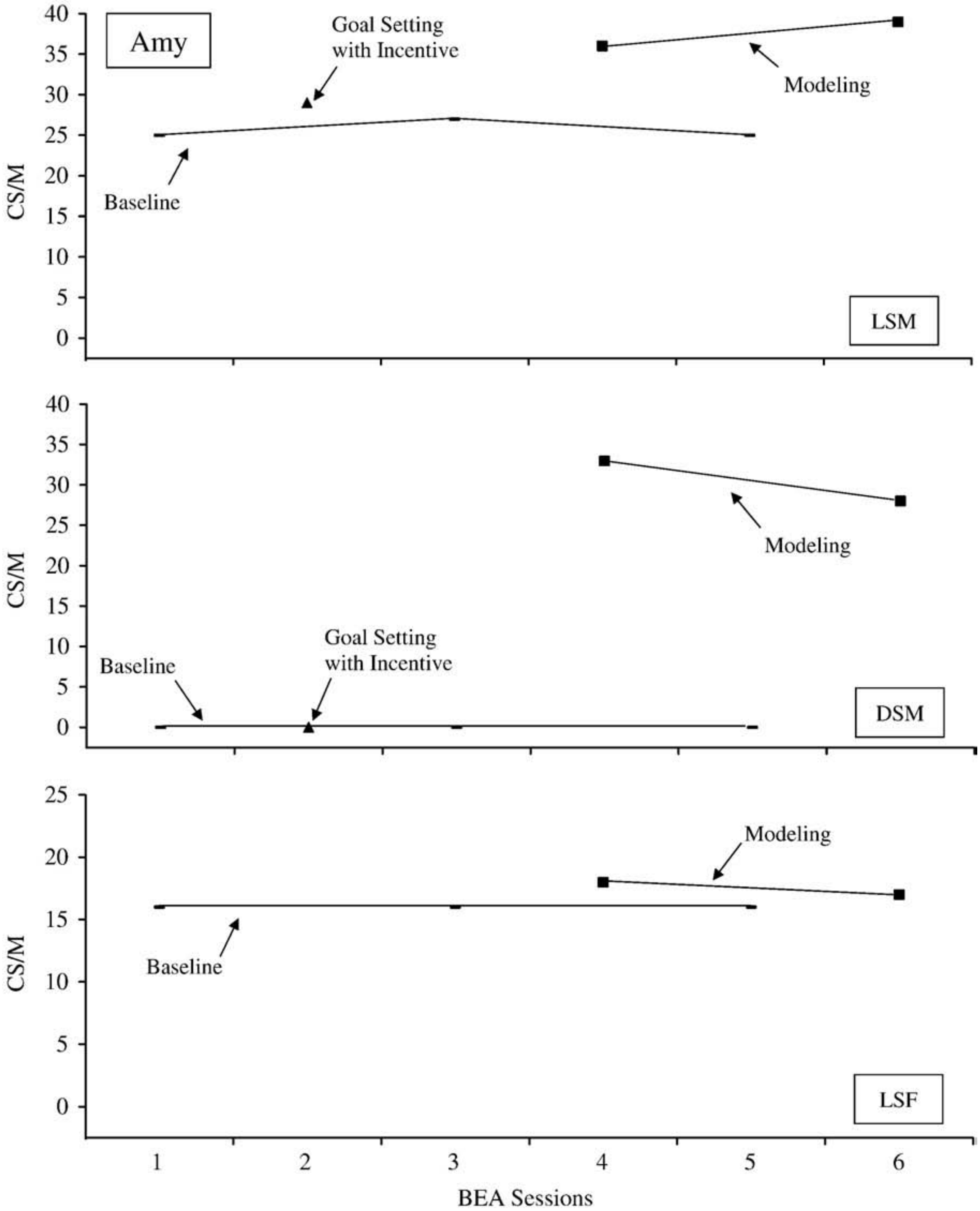


Fig. 2. Amy's correct sounds per min (CS/M) on the letter-sound subskill measure (LSM), decoding subskill measure (DSM), and LSF across different BEA conditions.

performance was consistently higher during modeling plus goal setting with incentive sessions than during preceding baseline sessions as measured with the LSM, DSM and LSF. However, the LSM appeared to be more sensitive to intervention effects as reflected in larger average increases over baseline (10.6 CS/M or 234%) than observed with LSF (2.6 CS/M or 29%), and clearer differentiation between data series than both the DSM and LSF. Based on the overall BEA findings, modeling plus goal setting with incentive was selected as the intervention for Leo.

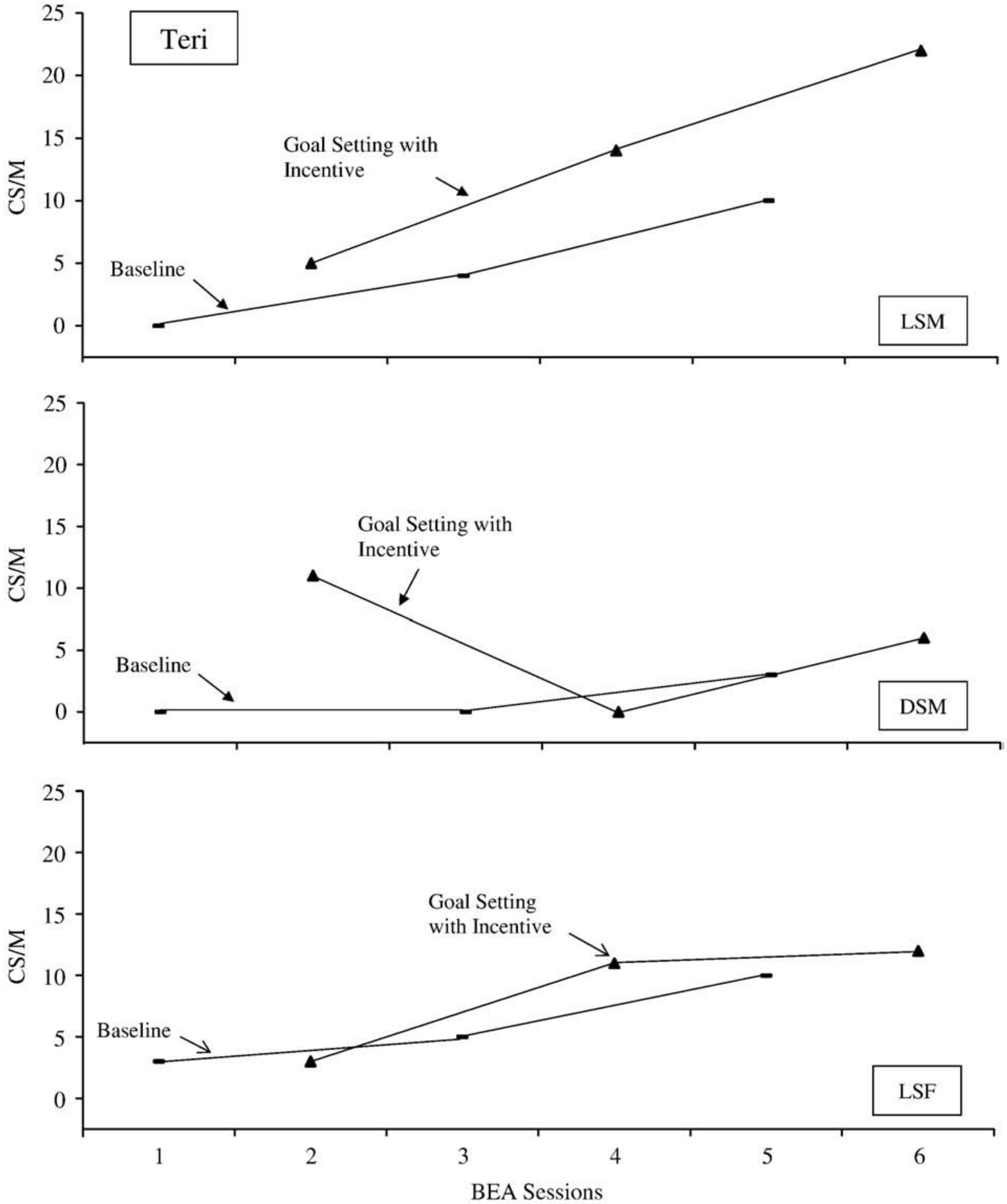


Fig. 3. Teri's correct sounds per min (CS/M) on the letter-sound subskill measure (LSM), decoding subskill measure (DSM), and LSF across different BEA conditions.

Amy

Results of Amy's BEA are presented in Fig. 2. Amy's performance was consistently higher during the Modeling sessions than during baseline as measured with the LSM and DSM. The DSM appeared to be more sensitive to intervention effects as reflected in larger average increases over baseline (30.5 CS/M) and clearer differentiation between conditions than observed with the LSM (9 CS/M or 46%). Virtually no differentiation between

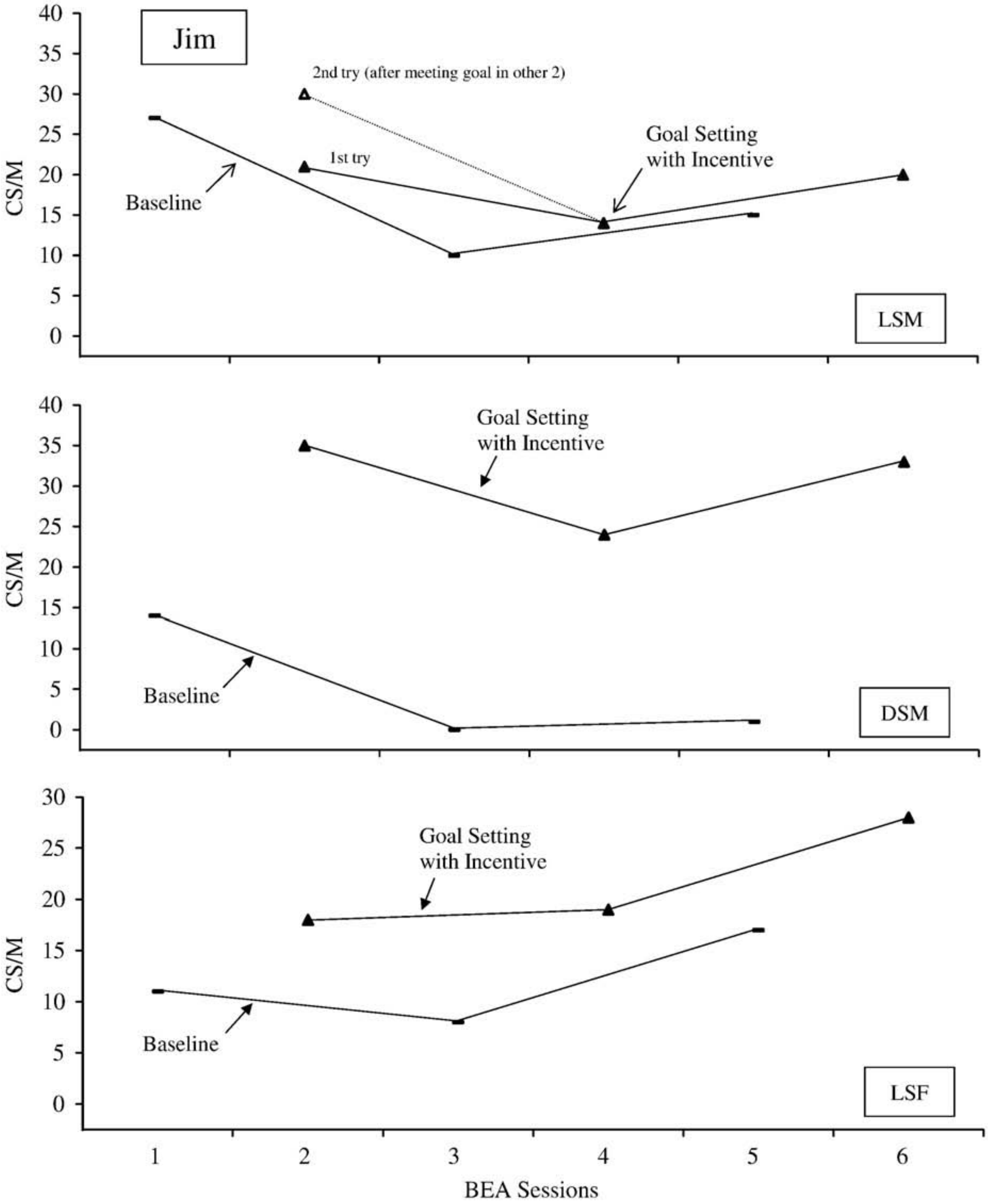


Fig. 4. Jim's correct sounds per min (CS/M) on the letter-sound subskill measure (LSM), decoding subskill measure (DSM), and LSF across different BEA conditions.

conditions was observed with LSF. Based on the overall BEA findings, modeling was selected as the intervention for Amy.

Teri

The results of Teri's BEA are presented in Fig. 3. Teri's performance was consistently higher during goal setting with incentive sessions than during baseline sessions on the LSM. The LSM yielded the clearest results, as reflected in larger average increases over baseline (9 CS/M) than observed with the LSF measure (2.7 CS/M or 47%), and greater differentiation between conditions. Results on the DSM were less clear due to overlap between the conditions. Based on the overall BEA findings, goal setting with incentive was selected as the intervention for Teri.

Jim

Results of Jim's BEA are presented in Fig. 4. Jim's performance was consistently higher during *Goal setting with incentive* than during baseline on the DSM and LSF. The DSM yielded the clearest results, as reflected in larger average increases over baseline (25.7 CS/M) than observed with LSF (9.7 CS/M or 89%) and greater differentiation between conditions. Based on the BEA findings, *goal setting with incentive* was selected as the intervention for Jim.

Table 1
Participants' growth slopes (unit increases per week), mean scores during baseline (BL) and BEA-identified interventions (INT), and effect sizes of interventions for different measures.

		LSF	NWF	WIF
Leo	BL slope (SE_b)	−0.3 (0.7)	−1.1 (0.3)	−0.3 (0.1)
	INT slope (SE_b)	1.8 (0.5)	1.9 (0.4)	0.9 (0.1)
	BL Mean (SD)	8.3 (2.0)	3.6 (2.0)	0.6 (0.4)
	INT Mean (SD)	23.8 (6.9)	14.1 (7.0)	4.4 (2.9)
	NAES	7.8	5.4	9.2
	Adjusted NAES	2.7	1.8	3.4
Amy	BL slope (SE_b)	1.9 (0.3)	−0.3 (0.5)	0.5 (0.2)
	INT slope (SE_b)	3.5 (1.1)	7.7 (3.8)	1.6 (0.3)
	BL Mean (SD)	12.0 (4.7)	1.8 (2.6)	0.9 (1.4)
	INT Mean (SD)	27.3 (6.8)	14.8 (15.9)	5.0 (2.6)
	NAES	3.2	5.1	2.9
	Adjusted NAES	3.1	1.5	0.4
Teri	BL slope (SE_b)	−0.7 (0.5)	0.1 (0.4)	−0.1 (0.1)
	INT slope (SE_b)	0.3 (0.5)	3.8 (1.8)	0.3 (0.5)
	BL Mean (SD)	5.2 (3.4)	0.9 (2.5)	0.4 (0.5)
	INT Mean (SD)	23.1 (2.9)	15.5 (15.2)	6.3 (3.3)
	NAES	5.2	5.9	11.5
	Adjusted NAES	3.5	1.8	2.7
Jim	BL slope (SE_b)	0.8 (0.3)	1.3 (0.6)	0.0 (0.1)
	INT slope (SE_b)	0.3 (0.7)	4.6 (0.6)	0.2 (0.1)
	BL Mean (SD)	10.3 (3.3)	7.2 (5.8)	0.3 (0.4)
	INT Mean (SD)	27.2 (4.8)	14.5 (10.9)	1.3 (1.2)
	NAES	5.1	1.3	2.2
	Adjusted NAES	3.3	2.1	2.8

LSF = curriculum-based measure of letter sound fluency, NWF = nonsense word fluency, WIF = word identification fluency. Unit of measurement for LSF and NWF is CS/M. Unit of measurement for WIF is CW/M.
SD = standard deviation, SE_b = standard error of the slope estimate, NAES = no assumption effect size, calculated by dividing the difference between each participant's baseline and intervention phase means by the standard deviation of the baseline. Adjusted NAES was calculated with Rosenthal's (1994) formula, $(X_{intervention} - X_{baseline}) / (SD_{pooled} / \sqrt{(2\ (1-r))})$, based on the three last measures in BL and INT phases, see Method section for more detailed description.

Generalized effects of BEA-identified interventions

Generalized effects of BEA-identified interventions applied to K-PALS materials were assessed with LSF, WIF, and NWF GOMs. Each student's growth slope, average scores during baseline and interventions, and effect sizes are presented in Table 1. According to Cohen's (1988) widely accepted guidelines, adjusted effect sizes around 0.20 were considered *small*, 0.50 *medium*, and 0.80 and greater *large* (see also Howell, 2002).

Leo

Fig. 5 shows Leo's progress before and after the BEA-identified intervention was implemented 3.5 times on average per week. During baseline, Leo did not show improvement on any of the measures, despite evidence-based K-PALS instruction with

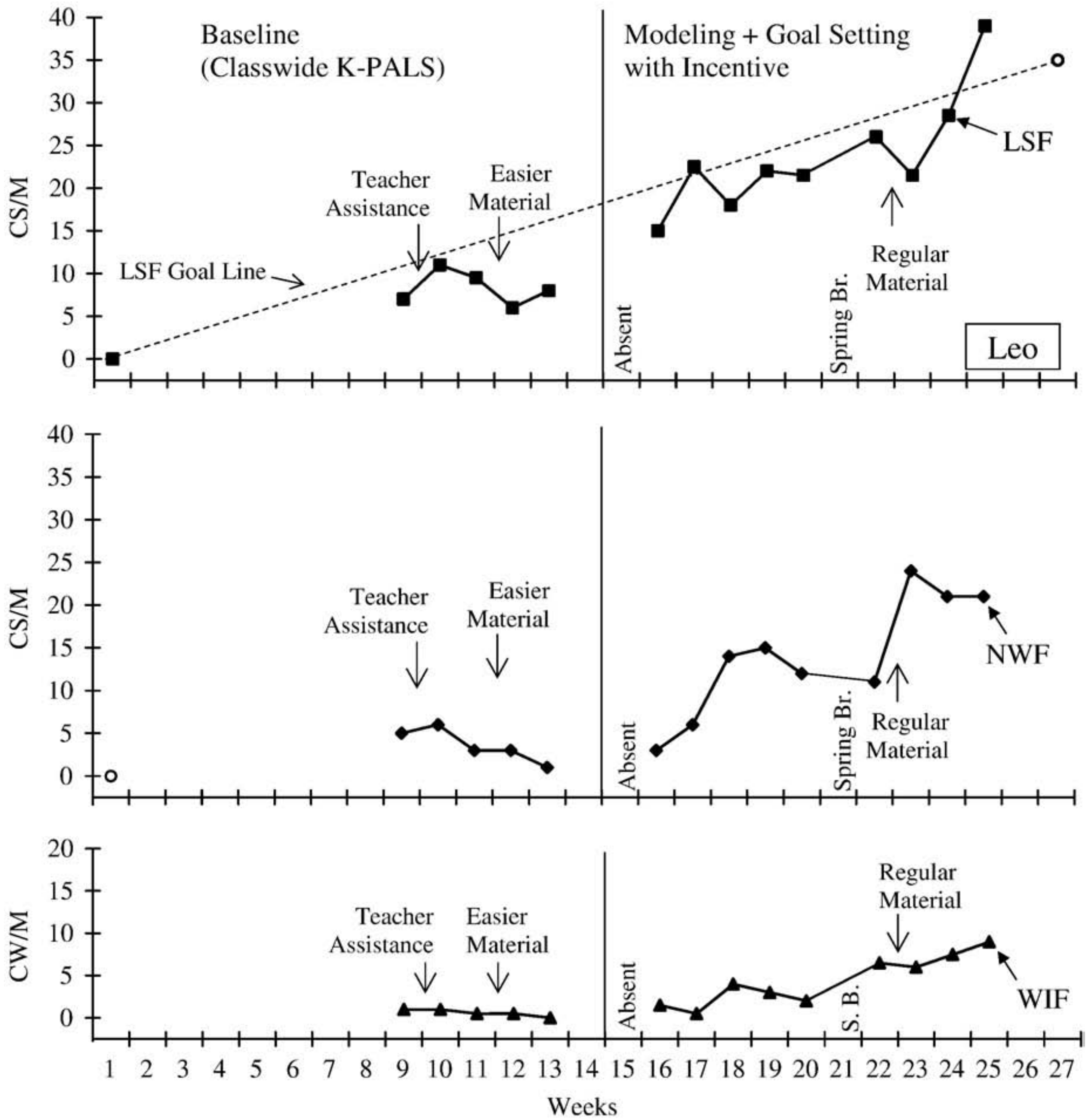


Fig. 5. Leo's performance on general outcome measures during baseline and BEA-identified intervention: correct letter sounds per min (CS/M) on LSF and NWF and correct words/min (CW/M) on WIF.

additional teacher assistance and easier materials. When the BEA-identified intervention (modeling and goal setting with incentive) was implemented with K-PALS materials, large effects on Leo's performance on all GOMs were observed. All measures showed a positive trend, with slopes ranging 0.9 to 1.9 unit increases per week. Greatest effects were observed on LSF and WIF, with adjusted NAES of 2.7 and 3.4 respectively.

Amy

Fig. 6 shows Amy's progress before and after the BEA-identified intervention was implemented an average of 3.8 times per week. During baseline, Amy improved on the GOMs, but her level of performance remained substantially below that of her classmates.

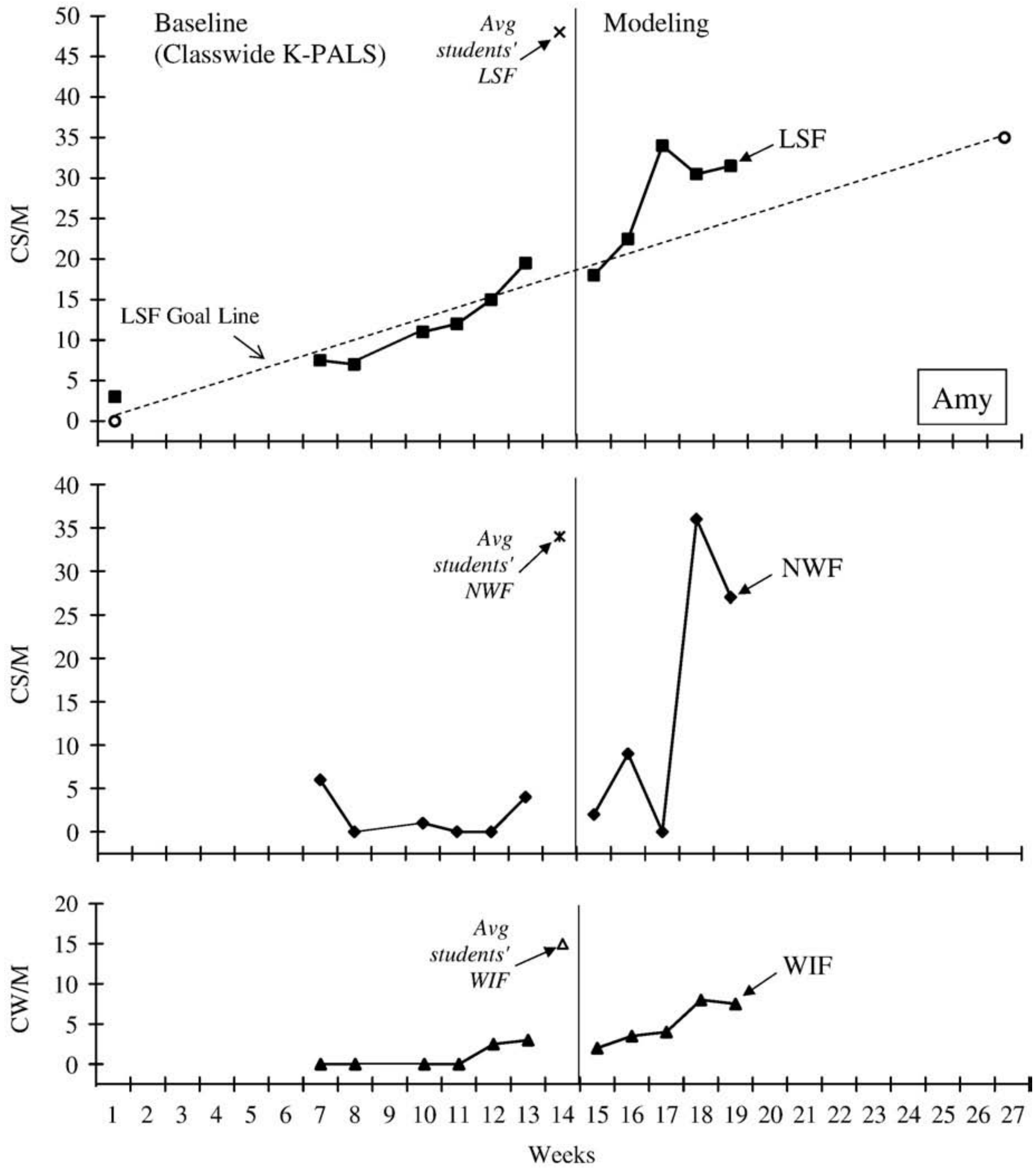


Fig. 6. Amy's performance on general outcome measures during baseline and BEA-identified intervention: correct sounds per min (CS/M) on LSF and NWF and correct words/min (CW/M) on WIF.

When the BEA-identified intervention (modeling) was implemented with K-PALS materials, all GOMs showed a steeper positive slope than during baseline, with unit increases ranging from 1.6 to 7.7 per week. The greatest effects were observed on LSF and NWF, with adjusted NAES of 3.1 and 1.5 respectively.

Teri

Fig. 7 shows Teri's progress before and after the BEA-identified intervention was implemented an average of 3.6 times per week. During baseline, Teri's performance on all GOMs remained low or showed a declining trend, despite K-PALS instruction. When the BEA-identified intervention (goal setting with incentive) was implemented, large effects on Teri's performance on all GOMs were observed, with adjusted NAES ranging from 1.8 to 3.5, and slopes ranging from 0.3 to 3.8 unit increases per week.

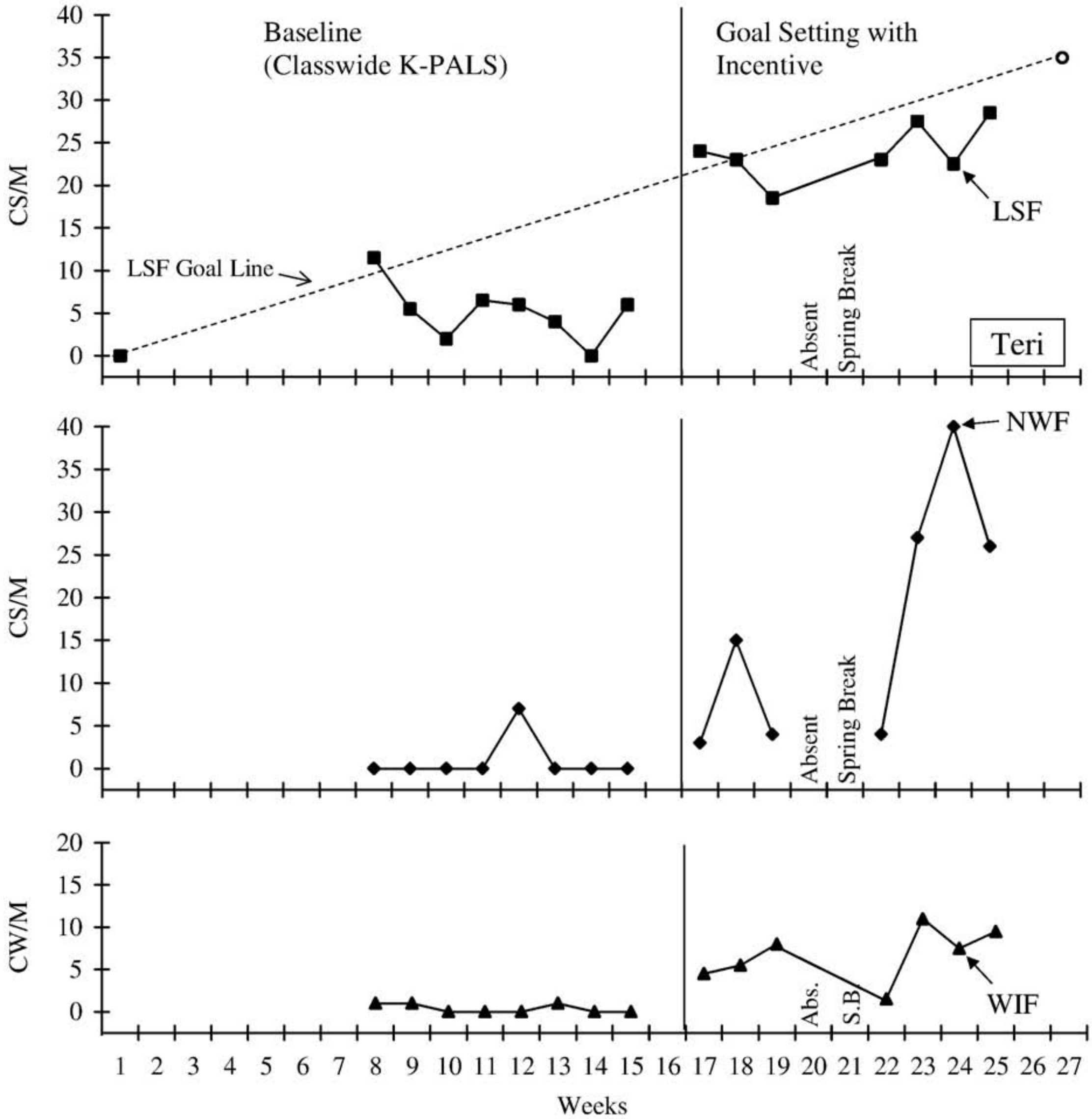


Fig. 7. Teri's performance on general outcome measures during baseline and BEA-identified intervention: correct sounds per min (CS/M) on LSF and NWF and correct words/min (CW/M) on WIF.

Jim

Fig. 8 shows Jim's progress before and after the BEA-identified intervention was implemented an average of three times per week. During baseline, Jim's performance on all GOMs remained low or showed only a modest positive trend, despite K-PALS instruction. When the BEA-identified strategy goal setting with incentive was implemented, large effects on Jim's performance on all GOMs were observed, with adjusted NAES ranging from 2.1 to 3.3, and slopes ranging from 0.2 to 4.6 unit increases per week.

Overall, effects of BEA-identified interventions were assessed within a multiple-baseline design across participants, with varying baselines to control for effects of extraneous factors, such as history, maturation, and testing. Fig. 9 shows participants' progress on each measure before and after BEA-identified interventions were implemented. Immediate changes in level of the primary dependent variable, CS/M (measured with LSF), were observed for all participants except Amy, who had already been making progress during baseline, and made a steep increase in progress soon after her BEA-identified intervention was applied to K-PALS activities.

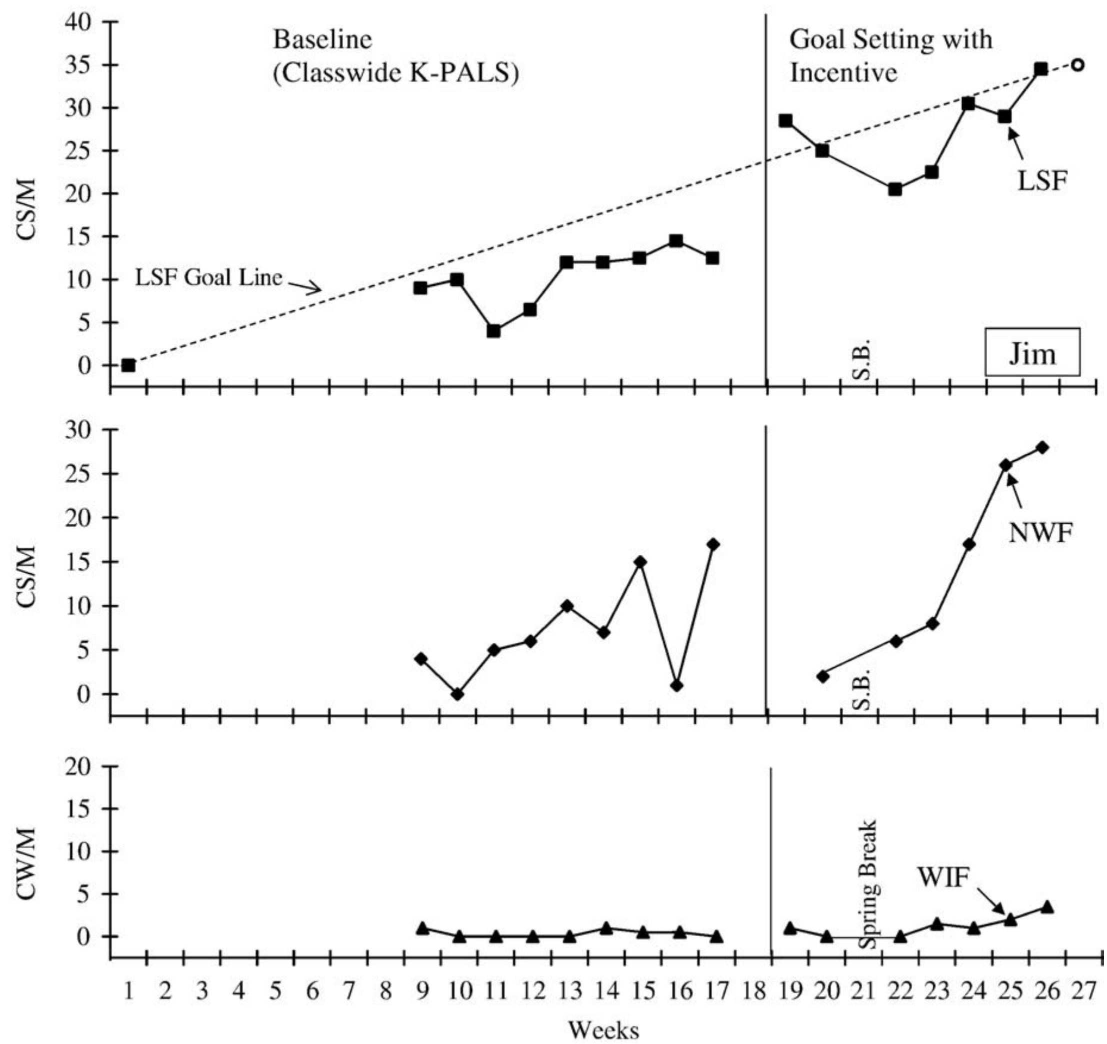
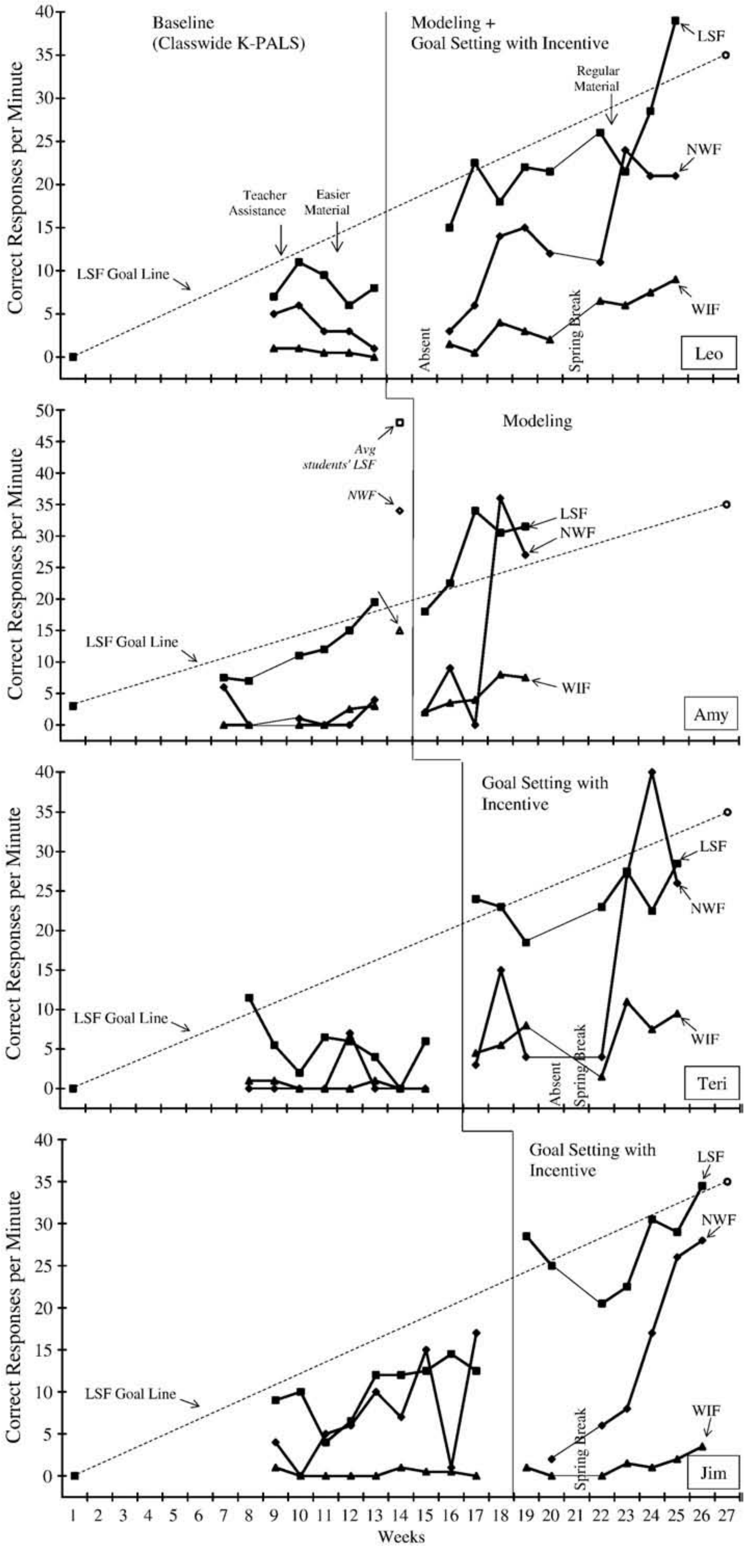


Fig. 8. Jims's performance on general outcome measures during baseline and BEA-identified intervention: correct sounds per min (CS/M) on LSF and NWF and correct words/min (CW/M) on WIF.


Baseline
(Classwide K-PALS)
Modeling +
Goal Setting with Incentive
Teacher Assistance
Easier Material
Regular Material
LSF Goal Line
LSF
NWF
WIF
Absent
Spring Break
Leo
Modeling
Avg students' LSF
Amy
Goal Setting with Incentive
Teri
Jim
Correct Responses per Minute
Weeks

Discussion

Results indicate that BEAs can successfully identify effective interventions for Kindergartners for whom evidence-based reading instruction was not sufficient, although the clearest results were attained with different measures across participants. Assessment with GOMs over time showed that the BEA-identified interventions had a large effect on all participants' early reading skills, supporting the treatment validity of BEAs.

How can BEAs be used to identify effective early reading interventions?

Findings shed light on what kind of interventions are powerful enough to effect responding when administered extremely briefly during a BEA, which measures are sensitive enough to detect those effects, and how feasible early reading BEAs are.

Which interventions?

Different interventions yielded the largest improvements in performance across participants. *Modeling and goal setting with incentive* was found to be promising for Leo, yielding greater improvements in performance than *goal setting with incentive* alone. This finding is consistent with research showing a combination of instructional and motivational interventions to be successful (Eckert et al., 2002; Noell et al., 1998, 2001). *Modeling* alone was sufficient to improve Amy's performance, and worked better than *goal setting with incentive*, consistent with research showing instructional strategies alone being sufficient to increase some students' performance (Daly and Martens, 1994; Noell et al., 1998).

Goal setting with incentive for improved performance effectively raised the performance levels of Teri and Jim. This finding is consistent with findings from Bonfiglio, Daly, Martens, Lin, and Corsaut (2004) where offering contingent reinforcement for matching best previous performance was as effective in combination with instructional strategies. Similarly, Daly et al. (2005) found offering a reward for 30% improvement over last performance effective for one of the two participants (in fourth to fifth grade). However, other studies have found contingent reinforcement by itself to be insufficient to improve some students' performance (Daly, Martens, Dool, & Hintze, 1998; Daly et al., 1999; Jones & Wickstrom, 2002).

Findings of this study corroborate findings of most BEA studies, which have shown individual differences in intervention effects (e.g., Daly et al., 1998, 1999; Eckert et al., 2002; Noell et al., 1998, 2001), although a few have found the same intervention to be effective for all participants (Daly & Martens, 1994; VanAuken, Chafouleas, Bradley, & Martens, 2002). The observation that individual students respond differently to interventions supports using an individualized problem solving approach to address academic problems.

Which measures?

Because BEAs had not been used with Kindergartners before, it was not known what kind of measures would be best suited to reflect differential effects of early reading

Fig. 9. Performance on general outcome measures during baseline and BEA-identified interventions: correct letter sounds/min on LSF and NWF, correct words/min (CW/M) on WIF.

interventions during brief exposures in a BEA. Therefore, three measures were used to assess participants' responding, so that those yielding the clearest results could be identified. Across participants, different measures yielded the clearest results, as reflected in largest changes in level (percentage increases) and differentiation between baseline and intervention data series. The LSM yielded the clearest results for two participants and the DSM for the other two. As predicted, LSF was not as sensitive to intervention effects as the subskill measures, most likely because it had less content overlap with the instructional material than the subskill measures.

The one other study that has included a measure of LSF in a BEA indicates that the measure may not be well suited for BEAs with older, struggling readers. Noell and colleagues (2001) applied interventions to three levels of materials for second to third grade struggling readers: their assigned grade level passages and two lower level materials, down to word lists and letter sounds. Measures of LSF showed differentiation during BEAs for two students (second and third grader), but exhibited a ceiling effect for one second grader, because the score was not prorated, precluding a demonstrable increase in performance from baseline in some cases.

How feasible are early reading BEAs?

In this study, three of the four BEAs were completed within 50 min, but Leo's BEA took 3 h to finish because of considerable variability in responding, possibly related to attention problems and to unequal difficulty of subskill measures. In previous research, BEAs have been reported to last 40 to 90 min (Daly et al., 1998, 1999; Duhon et al., 2004; Jones & Wickstrom, 2002; Noell et al., 2001). However, even though BEAs require time and effort, that time does not go to waste because the majority of BEA conditions involve evidence-based interventions that have the potential of improving student's performance and the results could lead to more efficient use of intervention time. By testing interventions in a hierarchy from least-to-most adult involvement the most efficient intervention can be identified, making the most of valuable instructional time.

Do BEA-identified interventions increase kindergartners' early reading skills?

The progress monitoring data in this study support the treatment validity of BEAs for struggling readers in Kindergarten. After the BEA-identified intervention was implemented for each participant, an increase in level and/or slope of performance was observed on most GOMs. No assumption effect sizes (NAES) for the interventions ranged from 1.3 to 9.2 with an average of 5.39, which compares favorably to a mean NAES of 2.87 observed for the most effective interventions identified in BEAs in a recent meta-analysis (Burns & Wagner, 2008). Adjusted NAESs ranged from $d=0.4$ to 3.5, with an average of $d=2.42$, which compares favorably to mean effect sizes of $d=0.58$ to $d=0.79$ observed in meta-analyses of academic interventions for students at risk for reading failure (NICHD, 2000) or with learning difficulties (Swanson, 1999). In another meta-analysis on single-subject design intervention studies, Swanson and Sachse-Lee (2000) found an overall average adjusted effect size of $d=0.87$, with an average of $d=0.90$ for two intervention studies targeting general reading. Part of the explanation for the large effect sizes observed in this study could be different skills being targeted than in most interventions for students with learning difficulties. For example, letter sound fluency targeted in this study represents a finite and relatively easily mastered skill compared to more commonly targeted oral reading fluency or reading comprehension.

Intervention effects over time

The generalized effects of BEA-identified interventions were assessed with three GOMs within a multiple-baseline design across participants over 5 to 9 weeks. Thus, generalization of intervention effects was assessed across materials (with assessment material being different from instructional materials) and across time. In previous BEA research, the effects of BEA-identified interventions over time have been assessed in different ways. Usually the instructional material has been used for assessment, making it a test of continuing effectiveness of interventions, rather than an assessment of generalized intervention effects, which requires untaught material. The observed effectiveness of BEA-identified interventions across time in this study is consistent with previous research showing continuing effectiveness of BEA-identified interventions compared to various control conditions (Duhon et al., 2004; Jones & Wickstrom, 2002; VanAuken et al., 2002).

Generalization across materials

Our findings of generalized effects on students' early reading skills over time are consistent with previous studies showing generalization of intervention effects across materials in the area of reading fluency with older students (Daly et al., 2005; Jones & Wickstrom, 2002; Noell et al., 2001) but inconsistent with studies showing limited or no generalization of intervention effects across related materials (Daly et al., 1998, 1999). The differences in generalization across materials observed across studies could be due to differences in the amount of practice within the interventions, as well as longer exposure to interventions. In those studies where generalization was assessed only during BEAs, lack of generalization was observed (Daly et al., 1998, 1999) whereas those assessing generalization during extended analyses reported generalization to new (Daly et al., 2005; Noell et al., 2001) or related material (Jones & Wickstrom, 2002). Also, greater difficulty of passages used to assess generalization in the studies by Daly and colleagues (1998, 1999) might have hampered generalization. It is plausible that the generalized intervention effects observed in this study are related to the substantial amount of practice participants received and the nature of the targeted skills. In contrast to oral reading fluency, letter sound fluency is a relatively finite skill that has the potential of facilitating performance on new tasks, such as decoding words, improving performance on GOMs involving word identification fluency.

Limitations and future research

Although the overall findings of this study are encouraging, they are preliminary with only a small number of participants, and remain to be replicated directly (across participants) and systematically (e.g., across settings and reading programs). Future research should address the following limitations. First, upon administering the subskill measures, in a few instances the student actually did know the sound of an 'unknown' letter sound. Thus, the subskill measures were not always of equal difficulty across BEA conditions which increased error variability (even though measures where the student knew both ‘unknown’ sounds were removed). This variability could be minimized in future research by classifying sounds only as ‘unknown’ if the student did not get them correct during the last two assessments prior to the BEA. Further research should also explore whether BEAs for early reading interventions actually require the use of specific subskill measures. In this study, students’ responding on the LSF measure was differentiated across interventions for one participant, suggesting that in some cases LSF can be helpful in identifying effective interventions. It remains to be determined under what conditions GOMs, such as LSF or NWF, could be used to assess the effectiveness of early reading interventions within a BEA. Also, measures of phoneme segmenting and blending are amenable to experimental analysis (Daly, Chafouleas, Persampieri, Bonfiglio, & LaFleur, 2004; Daly, Johnson, & LeClair, 2009) and should be examined in future research.

Second, only a limited array of interventions (and in two cases, only one) was assessed in the BEAs with instructional strategies confined to a few general categories. Perhaps even greater effects could be achieved by attempting a wider variety of interventions and choosing the *most* effective one for implementation. In addition, findings are constrained to K-PALS, which served as the base for BEA-identified interventions (that were also implemented with greater fidelity than classwide K-PALS methods during baseline, which could have contributed to the positive outcomes). Researchers may wish to explore whether similar results would be attained with a different reading program. Relatedly, it is not clear from the current design whether the observed effects were in fact due to BEA-identified interventions or just due to the additional adult-led one-on-one K-PALS tutoring sessions, regardless of strategies employed. A group design is needed to answer that question, and to investigate whether conducting BEAs provides the most efficient approach to selecting effective individualized interventions.

Third, interventions in this study were conducted one-on-one, which may not be realistic in the typical elementary school due to lack of resources. It is important to investigate whether BEA-identified interventions can be implemented through peer tutoring or in a group format with similar effectiveness, or whether BEAs can be used to directly test the effects of interventions in a group format.

Implications for practice

BEAs could play an important role as part of a problem solving approach within an RTI framework for addressing the needs of students with learning difficulties. BEAs provide a systematic, data-driven procedure to explore possible solutions and identify an intervention that is likely to be successful for an individual student. A strength of BEAs is that they focus on what educators can do to improve students’ performance (Daly et al., 1997), rather than on students’ limitations, deficits, or disorders. BEAs provide a direct link between assessment and intervention, and could strengthen traditional assessments by making them more solution-oriented, shifting the focus on what can be done to address a student’s academic problems (Martens, Eckert, Bradley, & Ardoin, 1999). This study demonstrates one relatively quick way that practitioners can test different interventions with struggling beginning readers to identify an effective approach that is likely to be successful in the long term and prevent reading failure.

References

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: a review of the literature. *Remedial and Special Education*, *23*, 300−316.

Batsche, G. M., Elliott, J., Graden, J., Grimes, J., Kovaleski, J. F., Prasse, D., et al. (2005). Response to intervention: Policy considerations and implementation Alexandria, VA: National Association of State Directors of Special Education, Inc.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of response to intervention: a snapshot of progress. *Journal of Learning Disabilities*, *42*, 85−95.

Bonfiglio, C. M., Daly, E. J., III, Martens, B. K., Lin, L. -H. R., & Corsaut, S. (2004). An experimental analysis of reading interventions: generalization across instructional strategies, time, and passages. *Journal of Applied Behavior Analysis*, *37*(1), 111−114.

Burns, M. K., & Wagner, D. (2008). Determining an effective intervention within a brief experimental analysis for reading: a meta-analytic review. *School Psychology Review*, *37*, 126−136.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: Applications in psychology and education* (pp. 187−212). Hillsdale, NJ: Lawrence Erlbaum Associates.

Carnine, D. W., Silbert, J., & Kame'enui, E. J. (1990). *Direct instruction reading*, (2nd Ed.) Columbus, OH: Merrill.

Cavanaugh, C. L., Kim, A., Wanzek, J., & Vaughn, S. (2004). Kindergarten reading interventions for at-risk students: twenty years of research. *Learning Disabilities: A Contemporary Journal*, *2*, 9−21.

Christ, T. J., Burns, M. K., & Ysseldyke, J. E. (2005). Conceptual confusion within response-to-intervention vernacular: clarifying meaningful differences. *Communiqué*, *34*(3), 1−8.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, (2nd ed.) New York: Academic Press.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlational analysis for the behavioral sciences*, (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, L. J., Wacker, D. P., Thursby, D., Plagmann, L. A., Harding, J., Millard, T., et al. (1992). Analysis of the effects of task preferences, task demands, and adult attention on child behavior in outpatient and classroom settings. *Journal of Applied Behavior Analysis*, *25*, 823−840.

Coyne, M. D., Kame'enui, E. J., Simmons, D. C., & Harn, B. A. (2004). Beginning reading intervention as inoculation or insulin: first-grade reading performance of strong responders to kindergarten intervention. *Journal of Learning Disabilities*, *37*, 90−104.

Daly, E. J., III, Chafouleas, S. M., Persampieri, M., Bonfiglio, C. M., & LaFleur, K. (2004). Teaching phoneme segmenting and blending as critical early literacy skills: an experimental analysis of minimal textual repertoires. *Journal of Behavioral Education*, *13*, 165−178.

Daly, E. J., III, Johnson, S., & LeClair, C. (2009). An experimental analysis of phoneme blending and segmenting skills. *Journal of Behavioral Education*, *18*, 5−19.

Daly, E. J., III, & Martens, B. K. (1994). A comparison of three interventions for increasing oral reading performance: application of the instructional hierarchy. *Journal of Applied Behavior Analysis*, *27*, 459−469.

Daly, E., Martens, B., Barnett, D., Witt, J., & Olson, S. (2007). Varying intervention delivery in response to intervention. *School Psychology Review*, *36*, 562−581.

Daly, E. J., III, Martens, B. K., Dool, E. J., & Hintze, J. M. (1998). Using brief functional analysis to select interventions for oral reading. *Journal of Behavioral Education*, *8*, 203−218.

Daly, E. J., III, Martens, B. K., Hamler, K. R., Dool, E. J., & Eckert, T. L. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *Journal of Applied Behavior Analysis*, *32*, 83−94.

Daly, E. J., III, Martens, B. K., Kilmer, A., & Massie, D. (1996). The effects of instructional match and content overlap on generalized reading performance. *Journal of Applied Behavior Analysis*, *29*, 507−518.

Daly, E. J., III, Persampieri, M., McCurdy, M., & Gortmaker, V. (2005). Generating reading interventions through experimental analysis of academic skills: demonstration and empirical evaluation. *School Psychology Review*, *34*, 395−414.

Daly, E. J., III, Witt, J. C., Martens, B. K., & Dool, E. J. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review*, *26*, 554−574.

Deno, S. L. (1985). Curriculum-based measurement: the emerging alternative. *Exceptional Children*, *52*(3), 219−232.

Deno, S. L. (1989). Curriculum-based measurement and special education services: a fundamental and direct relationship. In Mark R. Shinn (Ed.), *Curriculum-based measurement: Assessing special childrenThe Guilford school practitioner series*. (pp. 1−17) New York, NY: Guilford Press.

Duhon, G. J., Noell, G. H., Witt, J. C., Freeland, J. T., Dufrene, B. A., & Gilbertson, D. N. (2004). Identifying academic skills and performance deficits: the experimental analysis of brief assessments of academic skills. *School Psychology Review*, *33*, 429−443.

Eckert, T. L., Ardoin, S. P., Daly, E. J., III, & Martens, B. K. (2002). Improving oral reading fluency: a brief experimental analysis of combining an antecedent intervention with consequences. *Journal of Applied Behavior Analysis*, *35*, 271−281.

Fuchs, D., Berends, M., McMaster, K. L., Sáenz, L., & Yen, L. (2006, June). Scaling-up peer-assisted learning strategies: a longitudinal, multi-site randomized control trial to determine necessary on-site technical assistance to teacher implementers. *Poster presented at the Institute for Education Sciences Research Conference, Washington, DC*.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children*, *57*, 488−500.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: a unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice*, *13*, 204−219.

Fuchs, L. S., & Fuchs, D. (2004). Determining adequate yearly progress from Kindergarten through grade 6 with curriculum-based measurement. *Assessment for Effective Intervention*, *29*(4), 25−37.

Fuchs, D., & Fuchs, L. S. (2005). Peer-assisted learning strategies: promoting word recognition, fluency, and reading comprehension in young children. *The Journal of Special Education*, *39*, 34−44.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: a theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239−256.

Fuchs, D., Fuchs, L., Thompson, A., Al Otaiba, S., Yen, L., McMaster, K., et al. (2001). *Peer-Assisted Learning Strategies: Kindergarten Reading Teacher Manual.* Nashville, TN: Center on Accelerating Student Learning.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N. J., et al. (2001). Is reading important in reading-readiness programs? A randomized field trial with teachers as program implementers. *Journal of Educational Psychology*, *93*, 251−267.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N. J., et al. (2002). Exploring the importance of reading programs for kindergartners with disabilities in mainstream classrooms. *Exceptional Children*, *68*, 295−311.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice*, *18*(3), 157−171.

Good, R. H., & Kaminski, R. A. (2002). Nonsense word fluency. In R. H. Good & R. A. Kaminski (Eds.), *Dynamic Indicators of Basic Early Literacy Skills*, (6th ed.) Eugene, OR: Institute for the Development of Educational Achievement (Available: http://dibels.uoregon.edu/).

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading*, *5*(3), 257−288.

Howe, K. B., Scierka, B. J., Gibbons, K. A., & Silberglitt, B. (2003). A schoolwide organization system for raising reading achievement using general outcome measure and evidence-based instruction: one education district's experience. *Assessment for Effective Intervention*, *28*(3&4), 59−71.

Howell, D. C. (2002). *Statistical Methods for Psychology*, (5th ed.) Pacific Grove, CA: Duxbury.

Jones, K. M., & Wickstrom, K. F. (2002). Done in sixty seconds: further analysis of the brief assessment model for academic problems. *School Psychology Review*, *31*, 554−568.

Kaminski, R. A., & Good, R. H. (2002). Letter naming fluency. In R. H. Good & R. A. Kaminski (Eds.), *Dynamic Indicators of Basic Early Literacy Skills*, (6th ed.) Eugene, OR: Institute for the Development of Educational Achievement (Available Aug 10th 2005: http://dibels.uoregon.edu/).

Manolov, R., & Solanas, A. (2008). Comparing $N = 1$ effect size indices in presence of autocorrelation. *Behavior Modication*, *32*, 860−875.

Martens, B. K., Eckert, T. L., Bradley, T. A., & Ardoin, S. P. (1999). Identifying effective treatments from a brief experimental analysis: using single-case design elements to aid decision making. *School Psychology Quarterly*, *14*, 163−181.

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: an experimental field trial of identification and intervention methods. *Exceptional Children*, *71*, 445−463.

Morgan, P. L., & Sideridis, G. D. (2006). Contrasting the effectiveness of fluency interventions for students with or at risk for learning disabilities: a multilevel random coefficient modeling meta-analysis. *Learning Disabilities Research & Practice*, *21*, 191−210.

National Center for Education Statistics (2007). *The Nation's Report Card: Reading 2007.* (Retrieved December 26, 2007: http://nces.ed.gov/nationsreportcard/pubs/main2007/2007496.asp).

National Institute of Child Health and Human Development (2000). Report of the national reading panel. *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups (NIH Publication No. 00-4754)* Washington, DC: U.S. Government Printing Office (Retrieved Aug 11th 2005: http://www.nichd.nih.gov/publications/nrp/report.htm).

Noell, G. H., Freeland, J. T., Witt, J. C., & Gansle, K. A. (2001). Using brief assessments to identify effective interventions for individual students. *Journal of School Psychology*, *39*, 335−355.

Noell, G. H., Gansle, K. A., Witt, J. C., Whitmarsh, E. L., Freeland, J. T., LaFleur, L. H., et al. (1998). Effects of contingent reward and instruction on oral reading performance at differing levels of passage difficulty. *Journal of Applied Behavior Analysis*, *31*, 659−663.

Olive, M. L., & Smith, B. W. (2005). Effect size calculations and single subject designs. *Educational Psychology*, *25*, 313−324.

Riley-Tilman, T. C., & Burns, M. K. (2009). *Evaluating educational interventions: single-case design for measuring response to educational interventions.* New York: Guilford.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231−244). New York: Russel Sage.

Simmons, D. C., & Kame'enui, E. J. (1989). Direct instruction of decoding skills and strategies. *LD Forum*, *15*(1), 35−38.

Swanson, H. L. (1999). Instructional components that predict treatment outcomes for students with learning disabilities: support for a combined strategy and direct instruction model. *Learning Disabilities Research*, *14*(3), 129−140.

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject design intervention research for students with LD. *Journal of Learning Disabilities*, *33*(2), 114−136.

U.S. Department of Education (2004). Public Law 108-446 Individual with Disabilities Education Improvement Act. *One Hundred and Eighth Congress of the United States of America*.

VanAuken, T., Chafouleas, S. M., Bradley, T. A., & Martens, B. K. (2002). Using brief experimental analysis to select oral reading interventions: an investigation of treatment utility. *Journal of Behavioral Education*, *11*, 165−181.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: the promise and potential problems. *Learning Disabilities: Research & Practice*, *18*, 137−146.

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children*, *69*, 391−409.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S., Chen, R., Pratt, A., et al. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, *88*, 601−638.

Wagner, D., McComas, J. J., Bollman, K., & Holton, E. (2006). The use of functional analysis of academic response to intervention for oral reading. *Assessment for Effective Intervention*, *32*, 40−49.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *Journal of Special Education*, *41*, 85−120.