

## Response to Intervention:

### *Empirically Based Special Service Decisions From Single-Case Designs of Increasing and Decreasing Intensity*

David W. Barnett, *University of Cincinnati*

Edward J. Daly III, *University of Nebraska-Lincoln*

Kevin M. Jones and F. Edward Lentz, Jr., *University of Cincinnati*

There have been several proposals to the effect that special service decisions could be based, at least in part, on the construct of response to intervention and not necessarily on traditional child measures. Single-case designs that focus on intervention response *and* intensity are reviewed as a potential evaluation framework for interdisciplinary teams to help answer special services resource questions. Increasing-intensity designs are based on sequential intervention trials ordered on a continuum that builds in intensity. Decreasing-intensity designs start with more comprehensive or multicomponent interventions and intervention facets that are systematically withdrawn so that interventions become more natural and ecologically sustainable. The advantages and challenges associated with these designs for use in special education eligibility decisions are discussed as models for child evaluation in schools.

The traditional use of one-point-in-time assessments of child characteristics to make decisions about special education services has been the source of a long-standing controversy (Reschly & Ysseldyke, 2002). Measures used for these decisions have included intelligence and achievement tests, adaptive behavior scales, and alternative measures of these and associated constructs, which number in the thousands. However, traditional decision processes have failed to focus on problem solution within the ecology of presenting problems in favor of perceived within-child deficits. Alternatively, response to intervention has evolved in status to be part of the criteria suggested by the President's Commission on Excellence in Special Education (U.S. Department of Education, Office of Special Education and Rehabilitation Services [OSERS], 2002) for determining services for students with specific learning disabilities (SLD) or other high-incidence disabilities. Response to intervention builds on concepts found in the Individuals with Disabilities Education Act (IDEA) as well as the No Child Left Behind Act, which, among many science-based program components, requires that students undergo effective instruction and progress monitoring before entering special education, to provide a starting place for educational accountability. Several proposals for operationalizing response to intervention have been made (Fuchs & Fuchs, 1998; Speece & Case, 2001; Vellutino et al., 1996). The field can expect more efforts like these and, for a time at least, different models to be tested (e.g., Pennypacker & Hench, 1998). Therefore, it is premature to advocate any single model. Although the issues are complex and multifaceted, a key aspect of the development of any response-to-intervention model is

the need for high-quality evaluation designs for decision making. The core features of single-case design may prove especially useful for evaluating interventions along a continuum of intensity that underlies response to intervention.

The purpose of this review was to examine single-case designs, within the context of classroom and curricular activities, as decision aides for special service questions. As mandated by IDEA, team judgments about service needs are the final arbiters, but aspects of single-case designs can help organize the construct of response to intervention by creating data sets that help guide decisions about children's services. To accomplish this purpose, we outline how evaluation designs may be incorporated into response to intervention (RTI) models. The primary advantages of these designs are that school-based teams can use scientifically supported methods for making special education decisions, potentially addressing the failures of traditional decision making and the concerns of the President's Commission on Excellence in Special Education. To this end, we describe the requirements of a response-to-intervention model linked to single-case designs to further develop valid decision-making frameworks.

### The Need to Change Decision-Making Practices

Traditional special education decision making is plagued by a number of serious problems, including the static nature of assessment that guides classification decisions, the lack of demonstrated technical adequacy (reliability and validity of

decisions) for both single and combined measures in making classification decisions, and the failure of the process to lead to defensible and useful categories (Barnett, Lentz, & Macmann, 2000). The milestone decisions in special education (classification, Individualized Education Program [IEP] development, progress monitoring, reevaluation, reintegration) are not typically made using a common, valid data set connected across decisions (e.g., Hardman, McDonnell, & Welch, 1997; OSERS, 2002). Although these are all compelling reasons to reconsider how classification decisions are made, the most important issue is the failure of traditional methods to be directly linked to effective, ongoing intervention planning and, thus, to positive outcomes for children (Gresham & Witt, 1997). This failure to produce functional outcomes is the most serious indictment of a process of classification decision making that should be judged ultimately by how well it leads to improved academic and social trajectories for at-risk students and students with disabilities.

The report of the President's Commission on Excellence in Special Education (OSERS, 2002) provided extensive support for the above conclusions and made important recommendations for change. First, the report recommended the abandonment of the traditional classification process in favor of a decision-making process based on response to instruction for SLD (see also Fuchs & Fuchs, 1998, and Gresham, 2001). The literature suggests that this idea holds general promise for disability-related decisions (e.g., Barnett, Bell, et al., 1999; Gresham, 1991). Second, using scientifically validated, continuous progress monitoring (Fuchs & Fuchs, 1986) is strongly encouraged for making instructional decisions that lead to effective special services. Third, new models should not be based on "waiting for children to fail" before organized interventions are attempted (O'Shaughnessy, Lane, Gresham, & Beebe-Frankenberger, 2003). Finally, the President's Commission recommended the adoption of dynamic progress-monitoring methods for making decisions about continuing services (reevaluations).

The critical elements for implementation of these recommendations have been well established by supportive empirical research across several decades (as follows). What is needed is for key elements to be combined into a comprehensive decision-making model within an appropriate evaluation framework. An evaluation framework is justified, based on its ability to accurately document real changes on meaningful academic and social outcomes.

## The Contexts of Response to Intervention

Models based on response to intervention use the quality of student responses to research-based interventions as the basis for decisions about needed services. Any model guiding decisions should be comprehensive and meet all legal require-

ments, provide a standard process for making sequential decisions about student needs, emphasize the importance of using scientifically based interventions, and have judgments about validity focused on significant student outcomes occurring within the legal framework. Our premise is that a successful model for making special education decisions should be based on structured, data-based problem solving and flexible service delivery (Lentz, Allen, & Ehrhardt, 1996; Tilly, Reschly, & Grimes, 1999); monitor student progress on socially valid outcome measures (Wolf, 1978); and focus on what happens to students in natural classroom contexts.

Assessment within natural contexts may be accomplished through the use of direct assessment of student skills, problem behaviors, and functional contextual variables, to improve decisions made about intervention targets and related necessary environmental changes (Gresham, Watson, & Skinner, 2001; Lentz & Shapiro, 1986; McComas & Mace, 2000; O'Neill et al., 1997). The result of a useful assessment and analysis, in practice, leads to a plan to help close or ameliorate discrepancies between student performance and classroom expectancies; support for implementation of the intervention; and, the topic of this article, a design to organize and graph the data in order to inform intervention decisions.

Given this context, the evaluation methods we describe, increasing/decreasing-intensity single-case designs, use the concept of intervention intensity as the primary guiding factor for creation of a unified system covering a broad range of special education decisions, all judged by their impact on progress toward valued student outcomes. A common and continuous data set may be used across the team-based milestone decisions about special services, from structured intervention trials aimed at understanding problems before classification to decisions related to reintegrating students into general education.

## RTI as a Legacy and "Rule" for Special Education Decisions

Using a student's response to intervention as the basis for making special education decisions is not a new practice. Early models (e.g., Deno & Gross, 1973) defined many critical elements of a response-to-intervention model:

1. criteria for ensuring that students had critical deficits in basic skills for which special services were required and defined by the degree to which they were behind expected performance on socially important repeated measures (CBM probes, direct observations of behaviors) organized by time series (single case) designs,
2. goals for intervention efforts that would represent significant progress toward typical classroom expectancies, and

3. the need for special education based on the failure of a student to profit from structured preclassification efforts aimed at significantly reducing the deficit and demonstrating a trajectory toward a successful outcome.

These strategic elements provide a common thread across models that have been appearing in the literature (Fuchs & Fuchs, 1998; Speece & Case, 2001; Vellutino et al., 1996). They involve providing meaningful services prior to special education, employing systematic decision making, and demonstrating that special education would be necessary for further progress. However, several features require elaboration. A response-to-intervention model necessitates using decision-making methods that use graduated increases or decreases in intensity to demonstrate the initial and ongoing need for special services. *Intervention intensity* can be defined in various ways and will continue to evolve as a construct. Fundamentally, however, intervention intensity reflects qualities of time, effort, or resources that make intervention support in typical environments difficult as intensity increases, establishing a clear role for specialized services. Rather than leading to an intensity “cut score” used for classification, we view response to intervention as a data set based on individualized and careful analysis of student performance and needs for team decision making as required by IDEA.

The IEP development and monitoring process requires ongoing documentation of the intensity of services needed to sustain progress toward valid annual goals. Eventual reintegration from special to general education involves decreasing the intensity of an intervention and providing subsequent documentation to ensure that progress after special education could be sustained by general education teachers with less intense efforts. What we review are methods for systematically verifying that less intense interventions would be successful.

We examine basic design elements that can be used to create defensible data sets for eligibility decisions by careful examination of intervention sequence, response to intervention, intensity, and outcome—key themes underlying resource allocations (Fuchs & Fuchs, 1998; Gresham, 1991). Teams will have to decide where to begin the assessment and intervention process, but the design logic can start with school- and class-wide interventions (Horner & Sugai, 2000; O’Shaughnessy et al., 2003; Tilly et al., 1999) and move to more intensified interventions as needed, suggested by IDEA, the No Child Left Behind Act, and the President’s Commission on Excellence in Special Education (OSERS, 2002).

## RTI Data Sets for Special Services Decisions

### *The Construct of Intervention Intensity*

The nomological net of the construct of intervention intensity is a broad one (Gresham, 1989, 1991; LeLaurin & Wolery,

1992; Noell & Gresham, 1993; Sechrest, West, Phillips, Redner, & Yeaton, 1979; Yeaton & Sechrest, 1981), and variables used to measure intensity for any particular case should be determined idiographically (Lentz et al., 1996). Related terms appear in the literature, and like other educational and psychological concepts, construct-method links will require further development and scrutiny (i.e., selecting and measuring intensity variables). *Intervention strength* is probabilistically defined by the likelihood that an intervention will change a problem situation (Gresham, 1991; Yeaton & Sechrest, 1981). *Resistance to intervention* (also referred to as *response strength*) is defined as “the lack of change in target behaviors as a function of intervention” (Gresham, 1991, p. 25); that is, a child does not respond positively to intervention plans. A lack of change requires greater intervention strength derived from re-planning targeted variables, interventions, and appropriate support for children and teachers (Lentz et al., 1996).

### *Selecting and Measuring Intensity Variables in Context*

Two classes of variables must be measured as part of an assessment that provides information about a child’s response to interventions. First, there must be socially valid child outcome variables that can be measured repeatedly across time. Second, the variables selected must allow for quantification of the intensity of any intervention. Intervention intensity is related to, but different from, the typical measurement of treatment integrity.

**Child Outcome Variables.** Direct assessment methods, such as curriculum-based measurement (Fuchs & Fuchs, 1986; Shinn & Bamonto, 1998) for academic problems and classroom observation (Shapiro & Kratochwill, 2000), are widely used to measure important target behaviors during instructional and other interventions and have often been proposed as basic methods for examining readiness for general education. Other, indirect measures, such as child- and setting-specific behavioral ratings, have been used as repeated measures when direct measures are impractical (e.g., Barnett et al., 2000). Ideally, school-based teams would plan interventions for teaching new skills or reducing disruptive behaviors based on results from functional and direct assessment methods. Practical questions for assessment of response to intervention include whether acquisition of target skills (academic or social) occurs more rapidly under one set of intervention conditions or another, or whether behavioral fluency (e.g., oral-reading fluency) increases differentially. Behavioral reduction questions (e.g., the elimination of dangerous or disruptive behaviors) can be assessed directly, but they should also be recast as questions of acquisition of alternative or replacement behaviors during evaluation. From these measures and a review of the literature, the following classes of suitable child responses for this process have been suggested (Barnett, Bell, et al., 1999; Bell & Barnett, 1999): active student engagement

(or disruptions), rate of skill acquisition or trials to a set performance criterion, and behavioral fluency (academic or behavioral). As part of examining a child's response to interventions, one or more of these classes would be selected to measure child response to intervention across time.

**Intensity Defined by Intervention Implementation Qualities.** Logistical characteristics of interventions that change as interventions change, and that represent hierarchies of discrepancies from typical classroom routines, intervention difficulty, necessary intervention resources, or other indices of intensity, have been proposed (Barnett, Bell, et al., 1999; Bell & Barnett, 1999; Noell & Gresham, 1993). Table 1 provides a conceptual framework for these variables. For any student, intensity would relate to changes (e.g., from less to more resources or time) for a relevant set of these variables, with no assumption that all need to be directly measured.

Although the details of the contextual measurement of intensity are dependent on many idiosyncratic variables, the most basic requirements are that

1. a task analysis of the intervention plan is conducted,
2. the events and behaviors that comprise the intervention are defined,
3. appropriate logistical indicators of intensity are selected and a plan to measure or estimate them is developed, and
4. the scheduling and extent of the actual episodes involving participation of the child and change agents are planned and checked to estimate intervention intensity (e.g., Ehrhardt, Barnett, Lentz, Stollar, & Reifin, 1996; Gresham, 1989; LeLaurin & Wolery, 1992).

In regards to intensity, conclusions will be made about the degree to which interventions differ from typical routines in terms of resources, time, involvement of professionals other than the classroom teacher, and other factors.

There are at least two intervention scheduling tactics that may need to be described and measured to estimate the construct of intensity:

1. Intervention teams might need to specify and then verify the intervention schedule by occasions per day and their length, including recording the times per day that an intervention embedded in classroom routines was carried out; indicating the duration in minutes of an intervention, such as working in small, remedial groups or initiating teacher-directed repeated readings; documenting the duration in minutes that a teacher monitored a child; and noting that a supportive intervention was in place, with key steps followed (e.g., an activity schedule or peer buddy system).

2. If the intervention was contingency-driven, it could be reported as a percentage of occasions that the intervention was used as planned (e.g., if the child's behavior met the definition of appropriate behavior, positive teacher attention was given). Although estimates of variables will be less than perfect, the basic units of analysis are (a) the clear specification of the plan in operational terms, (b) the scheduling of the frequency and duration of the intervention plan or contingencies for plan use, (c) the measurement of intervention contacts between change agent and child or environmental changes, and (d) the assessment of the infrastructure of an intervention (e.g., planning, consultation, professional involvement).

In summary, many variables may be used to measure intervention intensity. We narrow the variables considerably by emphasizing the scheduling of the intervention and contacts with change agents as the primary factors. However, we also leave the question open because teams may need to add other key variables to provide estimates of intensity. The defining characteristics of such variables are that they are measurable and signify intervention effort.

### *Hierarchies of Intensity*

An *intervention hierarchy* describes a series of interventions or components that are unified by response class (e.g., low rates of academic responding, disruptive social behaviors) and ordered in a planned sequence to resolve a problem situation. The interventions in a hierarchy may differ due to children's responses to progressive intervention steps (e.g., Harding, Wacker, Cooper, Millard, & Jensen-Kovalan, 1994; Haring & Eaton, 1978).

Practical considerations, as well as ecological and behavioral principles, guide the hierarchical ordering of treatments. For example, Harding et al. (1994) attempted to identify the easiest and least intrusive intervention procedure that led to the largest incremental improvement in appropriate behavior so that they could teach parents how to do the interventions at home. Daly and colleagues have been evaluating hierarchical instructional trials (e.g., Daly, Lentz, & Boyer, 1996; Daly, Martens, Hamler, Dool & Eckert, 1999). Within this research, an example of an increasing-intensity design involves sequencing interventions for oral-reading fluency problems based in part on how much adult supervision would be necessary to carry out the intervention. Procedures requiring less supervision or fewer modifications to the curriculum are tried early in the sequence.

### *Instructional and Intervention Trials*

The idea of instructional and intervention trials includes two tactics: the logic of adding or subtracting intervention compo-

TABLE 1. Logical Characteristics of Interventions Related to Intensity

Category	Potential intensity variables
1. Intervention management and planning (measured by time teacher spent outside of typical routines)	Adults' monitoring of activities Teacher prompting Communication with stakeholders (e.g., parents) Progress-monitoring activities (e.g., assessment, graphing) Consultation and meetings between professionals
2. Activities embedded in typical classroom routines (measured by time spent beyond typical routines)	Modification of typical routines Modification of tasks or assessments Increased levels of assistance to students during class work Increased one-to-one interaction (e.g., additional practice within activities, different feedback system) Provision of contingencies (social or otherwise) for expected behaviors
3. Intervention episodes (outside of typical routines, measured by time/beyond typical routines)	Tutoring Social skill groups Counseling Additional remedial instruction (group or individual) Completely new instructional formats Provision of contingencies related to these efforts
4. Materials and other tangible resources (measured by cost or time to develop)	Additional practice materials Published remedial or new curricular packages
5. Change agents (nonprofessional vs. professional qualifications)	Peers Adult volunteers Paraprofessionals Certificated educators

nents to find the least intrusive intervention needed to meet the child's needs, based on direct measures of children's responses to interventions. First, intervention trials can be ordered by increasing intensity to help determine the most likely intervention alternatives for both social interventions and academic instructional decisions. Second, because challenging behaviors may demand immediate and comprehensive programming (e.g., for safety), a goal of decreasing intensity designs would be to rapidly reduce the intensity of a comprehensive intervention plan, consistent with normalization principles.

### *Response-to-Intervention Concepts Working Together*

To illustrate this point, a response-to-intervention "eligibility" data set would be based on both (a) a discrepancy between the educational performance in significant curriculum content (or other educational or social outcomes) of the referred child and that of typical children and (b) a desired change in performance that is not responsive to well planned and carried out interventions that are naturally sustainable within the educational service unit. We add to this the notion that hierarchically arranged or sequenced interventions may be used to

derive refined intervention plans to aid eligibility decisions. Within this conceptualization, at least three patterns or combinations of patterns derived from direct measurement of "intensity of the need for special support" are used for making decisions (Hardman et al., 1997):

1. An intervention may be successful but require extraordinary effort, time, or resources to be sustained.
2. An intervention may need to be in place extensively or throughout the school day for a child's appropriate inclusion.
3. Interventions that can be carried out by a teacher may be unsuccessful, requiring re-planning and special resources to be added to a situation. (p. 64)

The performance discrepancy and actual intervention data can be used to clarify the interventions and the special education services or supports needed to meet the needs of the child in the present environment, or to establish needed environmental modifications to do so. The team decision for special services eligibility occurs at this point—based not on a pattern of test scores but on a data set that shows a pattern of empir-

ically derived service-delivery needs based on response to intervention. Within a general approach of instructional and intervention trials sequenced by intensity are different single-case design strategies that form the bases of our discussion.

## Single-Case Designs Applied to Service-Delivery Questions

Single-case designs evolved because of the need to understand patterns of individual behavior in response to independent variables and, more practically, to examine intervention effectiveness. Design use can be flexible, described as a process of response-guided experimentation (Edgington, 1983), providing a mechanism for documenting attempts to live up to legal mandates for students who are not responding to routine instructional methods. The basic methods are

1. selecting socially important variables as dependent measures or target behaviors,
2. taking repeated measures until stable patterns emerge so that participants may serve as their own controls (i.e., baseline),
3. implementing a well-described intervention or discrete intervention trials,
4. continuing measurement of both the dependent and independent variables within an acceptable pattern of intervention application and/or withdrawal to detect changes in behavior and make efficacy attributions,
5. graphically analyzing the results to enable ongoing comparisons of the student's performance under baseline and intervention conditions, and
6. replicating the results to reach the ultimate goal of the dissemination of effective practices (Barlow, Hayes, & Nelson, 1984).

Single-case designs are a valid methodology for establishing empirical interventions (Stoiber & Kratochwill, 2000). Beyond research, single-case designs have been used for determining individually appropriate interventions (Wacker, Steege, & Berg, 1988) and organizing data from consultations that include accountability, service delivery, and professional preparation goals (Barnett, Air, et al., 1999; Barnett, Daly, et al., 1999; Kratochwill, Elliott, & Busse, 1995). However, a limitation with traditional design use in the context of educational programming is that the question of which intervention is the most effective and least intrusive may not be addressed. Interventions are based on a careful analysis of environment and response function and, pertinent to our discussion, often need to be "fine-tuned" (McComas & Mace, 2000). Many studies involve comparisons of intervention conditions, but these comparisons do not necessarily suggest that intervention plans should be refined to meet optimal levels of intensity.

## Organizing Designs by Increasing and Decreasing Intensity

Combinations and sequences of interventions are common in school practice (Ervin et al., 2001), and many designs can be used to compare intervention conditions by intensity. Parametric designs provide a basic way to establish behavioral response to intensity variables (ordering different values of the independent variable; Sidman, 1960). Multi-element (or alternating treatments) designs are widely used to rapidly try out various interventions that could be organized by different intensities and may be combined with other designs to examine intensity (e.g., a parametric design; Northup & Gully, 2001). Van Houten and Hall (2001) suggested an intensified treatment design, which looks like an A-B-B' design (or parametric design), with B' being a modification of B that calls for, for example, increasing the rate of teacher praise. Sainato, Strain, and Lyon (1987) used a changing-criterion design applied to rates of teacher instructional requests as a measure of intensity. Generally, *component analysis* describes studying the separate and combined effects of environmental variables that may be isolated, to reach the goal of identifying the minimal "package" or component necessary to obtain desired intervention effects (Vollmer & Van Camp, 1998).

Traditional notations may be used for describing designs of increasing and decreasing intensity. The baseline is Phase A; different intervention phases are labeled B, C, and so on. Combined treatments within an intervention phase are noted by discrete components (e.g., BC may stand for reinforcement plus feedback). The use of a prime (B'-B'') indicates that slight changes in intervention variables were made. Thus, an increasing-intensity design could be depicted as A-B-C, A-B-B'-B'', or A-B-BC, showing the progression of intervention changes by making modifications (B') or by adding components (C) consistent with an underlying hierarchy or continuum of intensity. For decreasing-intensity designs, a condition of relatively high intensity (BC) could be followed by a less intense condition (B); this design can also be demonstrated by a sequence such as B'' followed by B'. Similar notations can be used for accountability designs (e.g., A-B-C design) which lack control conditions (i.e., withdrawal and replication) but may still be used to evaluate plans in line with our review (Barlow et al., 1984). Given the many ways that intervention data may be organized, we emphasize two basic single-case design tactics that can help create data sets for special service-delivery decisions.

## Increasing-Intensity Designs

Increasing-intensity designs evaluate the least amount of intervention necessary to accomplish objectives, by making step-by-step decisions as interventions are attempted, with more intensive interventions used as necessary. The logic of increasing-intensity designs was illustrated by Sheridan, Kratochwill, and

Elliott (1990). During an initial treatment phase, parents and/or teachers assisted four socially withdrawn children in developing, practicing, and recording a specific goal for initiating a peer contact. During the second treatment phase, these procedures were extended to all opportunities for peer initiations throughout the day, which were monitored by the child in a daily journal. Results indicated that the frequency of social initiations approached or exceeded that of matched peers only after the second phase was introduced.

A component analysis that exemplifies an increasing-intensity design was provided by Rhymer, Dittmer, Skinner, and Jackson (2000). The math performance of four students showed only slight increases after implementation of a brief, fluency-building intervention. Substantial improvements over baseline conditions, however, were observed for three students when performance feedback was added to the fluency-building intervention.

McConnell et al. (2002) provided an example of a hierarchical increasing-intensity design. During an initial treatment, materials and structured activities produced moderate increases in the expressive language of four children with language delays. For two children, substantial increases were observed when structured activities were replaced by parent coaching that incorporated strategies for facilitating expressive language. All three of these studies demonstrate how the intensity of treatments may be increased by extending, adding, or altering procedures until intervention goals are met.

### *Decreasing-Intensity Designs*

Concerns can be high- or multi-risk (National Information Center for Children and Youth with Disabilities, 1999), extreme in the challenges that children present and the range of teachers' capacities to help with those challenges. Children may face harsh consequences of school failure or highly disruptive social behaviors, or a classroom may have more than one child with unusually challenging behaviors. Children receiving special services, often representing a multifaceted intervention, need to be reintegrated into more typical environments. Thus, decreasing-intensity designs begin with a multicomponent intervention or comprehensive package that has been shown to be effective and is predicted to meet children's immediate objectives for change. Experimental control is ascertained by systematically withdrawing intervention facets, ideally until the intervention is natural and ecologically sustainable (e.g., Anita & Kreimeyer, 1988; Dooley, Wilczenski, & Torem, 2001; Rusch, Connis, & Sower, 1979; Rusch & Kazdin, 1981; Sainato, Strain, Lefebvre, & Rapp, 1990). A family of designs have been described in this way (Kazdin, 1982). A sequential withdrawal design involves "gradually withdrawing different components of a treatment package to see if behavior is maintained" (Kazdin, 1982, p. 213). A partial withdrawal design to study response maintenance is linked to a multiple-baseline design (across behaviors, persons, or situations): "The intervention is first withdrawn from only one of the behaviors (or

baselines) . . . . If withdrawing the intervention does not lead to a loss of the behavior, then the intervention can be withdrawn from other behaviors (or baselines) as well" (Kazdin, 1982, p. 215).

Systematic, or partial, withdrawal allows for the opportunity to examine the maintenance of intervention effects. If an element of a treatment package is withdrawn and treatment effects are maintained, this component may have affected treatment initially but is no longer necessary. If treatment effects are not maintained, this element remains necessary for maintaining behavioral effects. The withdrawal of treatment elements in a stepwise manner is used to identify those elements that are no longer necessary.

As an example, Sainato et al. (1990) developed a treatment package to help integrate preschoolers with disabilities into kindergartens by increasing their independence. Beginning with three combined components (i.e., reinforcement, matching teacher observations with children's self-observations, and children's self-assessments), a sequential withdrawal design was used to successfully maintain appropriate behaviors, ending with the self-assessment strategy alone. Dooley et al. (2001) presented an example of a decreasing-intensity accountability design that mirrors practices commonly used with children displaying high-risk concerns. Following a baseline for disruptive behaviors and compliance for a young child diagnosed with pervasive developmental disorder, a two-component intervention package was implemented. After 5 days, the more intrusive reinforcement component was withdrawn and compliance levels were maintained only by an activity schedule, at levels similar to those of the combined package. This type of design (i.e., sequential withdrawal) allows decisions about necessary services to be empirically validated.

### *Innovations in Brief Empirical Problem Analysis*

Recently, suggestions have been made regarding empirical methods that may help teams more rapidly validate intervention hypotheses or suggest validity evidence for interventions. Brief functional analysis is based on brief exposures of manipulated conditions, with replications of results conducted to confirm hypotheses (Cooper et al., 1992). This model has been used extensively to test the functions of problem behavior (e.g., access to peer attention). Harding et al. (1994) employed brief experimental analysis to identify potential treatment components for challenging behaviors. Similarly, Daly and others (e.g., Daly & Martens, 1994; Daly et al., 1998; Daly et al., 1999; Daly, Murdoch, Lillenstein, Webber, & Lentz, 2002; Daly, Witt, Martens, & Dool, 1997; Eckert, Ardoin, Daly, & Martens, 2002; Martens, Eckert, Bradley, & Ardoin, 1999; VanAuken, Chafouleas, Bradley, & Martens, 2002) have used brief experimental analysis to identify effective reading strategies. In brief experimental analysis, a series of independent hypothesis-derived empirical treatments or combinations are implemented as needed in ascending order of some relevant

dimension, such as intrusiveness, ease, or difficulty. Analyses are based on rapid, single exposures of interventions for only a few sessions (i.e., less than 3 data points), and brief withdrawals and replications are used to strengthen inferences.

### Designs Applied to Special Education Eligibility Decisions

Figure 1 displays a hypothetical increasing-intensity design for Abby, a second-grade student. After parent consultation, review of statewide proficiency and other test scores and full academic records (including prior interventions), and a school-wide screening of important early literacy skills, all representing multifaceted information, the multidisciplinary team narrowed the primary area of concern to basic reading skills and proceeded with intervention trials to clarify the level of supports necessary to sustain adequate progress.

In an initial assessment, Abby read 25 correct words per minute (CWPM), which was approximately 53% below the average rate of five randomly selected second-grade peers. After a stable baseline (A), the first intervention phase (B) was introduced, linking incentives to language workbook com-

pletion and performance on daily running records administered by her teacher. The second intervention phase (C) involved modeling of oral reading and increased practice through peer tutoring. Four days of week, Abby and a peer took turns reading aloud to one another for 25 minutes. At the end of this phase, a second peer administration revealed that Abby's slope of improvement (1 word per week) was below that of her peers (2 words per week). The third intervention phase (D) increased successful engagement aimed at acquisition of new words and drill and practice for fluency. Four days per week, a reading specialist provided 25 minutes of phrase and word drills to Abby in a small-group context. First, miscues from initial readings were repeatedly presented in context. Second, isolated words that featured unmastered letter patterns generated by the classroom teacher were repeatedly presented. In response to phrase and word drills, Abby's slope of improvement (2.67 words per week) exceeded that of her peers at this point (2.33 words per week). When the previous intervention was reinstated, Abby's performance stabilized and that of her peers continued to increase.

Taking into consideration these results as well as prior evidence of the chronicity and severity of her reading problems, the team decided that Abby met criteria for special ser-

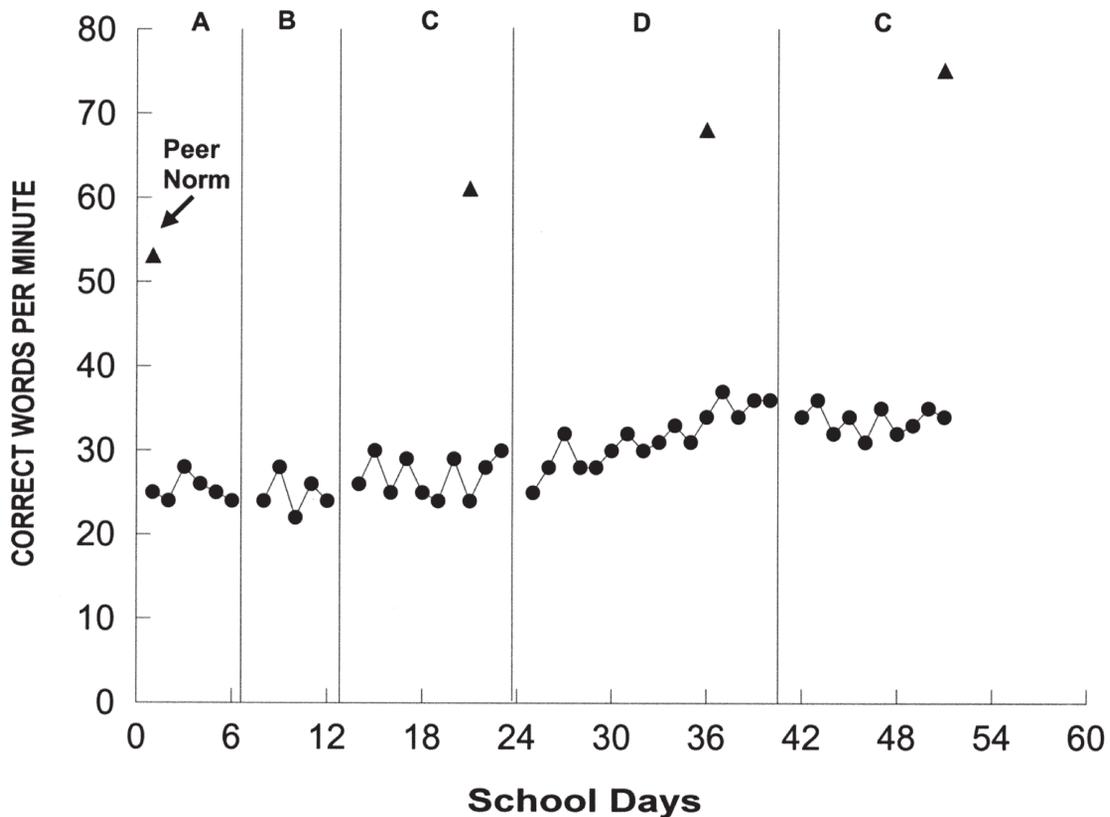


FIGURE 1. An increasing-intensity design: the number of correctly read words per minute across baseline (A) and Phases B (class-wide incentives), C (modeling and repeated practice), and D (word and phrase drills) for Abby.

vices (SLD) because grade-level learning rates were approximated only in response to a uniquely designed and specialized intervention. Changes in intensity are reflected by increases in intervention management requirements (teacher's record-keeping, team meetings, oral-reading probes, contingency management), embedded activities (peer tutoring that also required developing materials), and added expertise of change agents (from a peer model to a reading specialist, remedial materials).

Figure 2 shows a hypothetical decreasing-intensity accountability design and part of the data for an intervention-based evaluation for Tim, a first-grade student referred for high rates of dangerously disruptive behavior (i.e., climbing and jumping off furniture), aggressive behavior, and other developmental concerns. The problem-solving team (including parents) decided that an appropriate assessment for disability evaluation and intervention planning could be based on his extensive medical and educational histories; curriculum-based assessments; observations by teachers (general and special education) and the school psychologist; review of prior, less formal attempts by teachers to improve behaviors; and recent brief intervention trials. Measures shown included child disruptive behaviors and teacher contacts (e.g., necessary monitoring and proximity to child, directions or prompts).

Following baseline observations (A), including a descriptive functional analysis and the hypotheses that disruptions were related to lack of behavioral repertoires for the classroom routines (skill hypothesis) and that teacher attention was maintaining inappropriate behavior, a multicomponent intervention (BCD) was implemented to help with classroom behaviors. The interventions used were an activity schedule (B), classroom tokens and reinforcement (C), and daily home-school notes showing performance in each activity along with at-home positive practice of selected school routines (D). The activity schedule involved clarifying activities and management routines for all of the children. After disruptive behavior reached low levels, Phase D was eliminated. The behavior was maintained at near-goal levels, and Phase C was withdrawn. After an increase in disruptive behaviors, Phase C was reintroduced, resulting in the BC phase and a reduction of disruptive behaviors to low levels. Professionals and parents considered all existing information available, as well as the observations, interventions, and outcomes, to determine whether the interventions were of sufficient intensity and quality (supported with technical adequacy information) for the child to be eligible for special services as a "child with a disability."

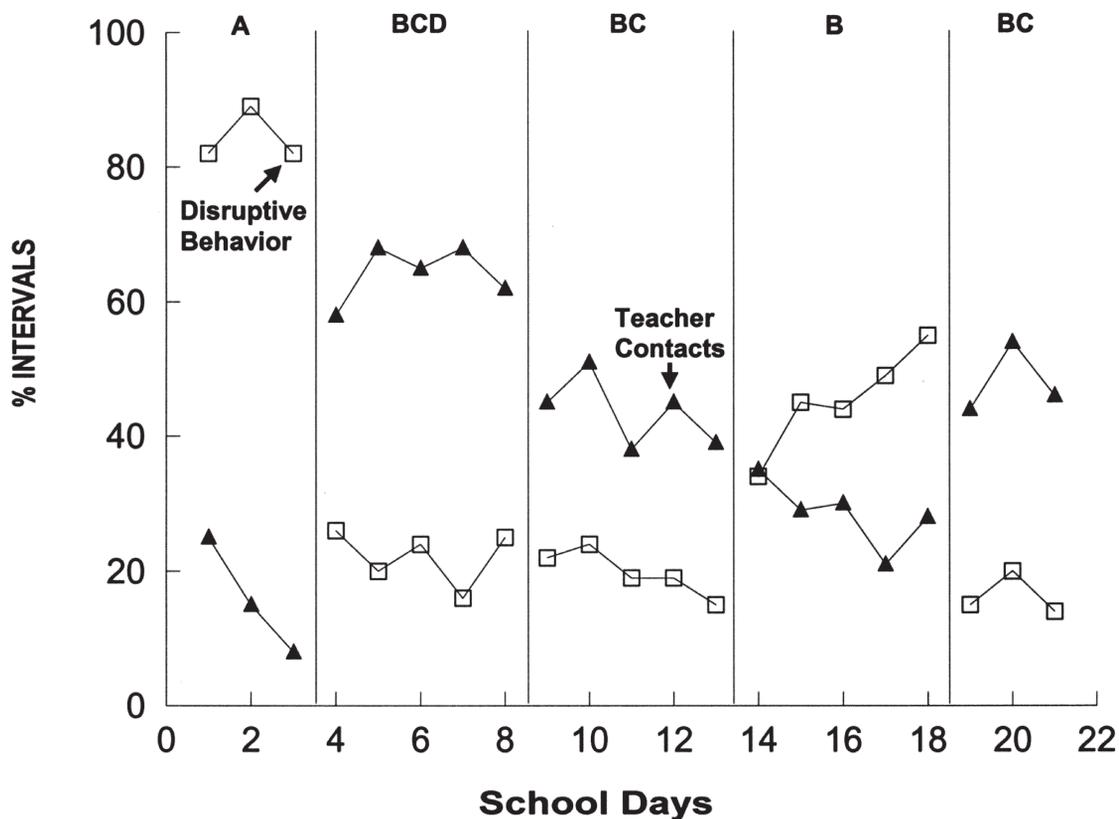


FIGURE 2. A decreasing-intensity design: the percentage of observed target responses across baseline (A) and Phases B (activity schedule), C, (classroom tokens and reinforcement), and D (daily home-school notes showing performance in each activity and at-home positive practice of selected school routines) for Tim.

Table 2 summarizes ways that graphic analysis can be used as an index of intervention intensity for the above cases. First, intervention effectiveness may be evaluated in terms of ongoing student performance in relation to intervention components. Second, intervention intensity may be evaluated in terms of the quality and quantity of intervention components, measured directly (see Figure 2) or indirectly through treatment checklists. Third, intervention teams may use the “intercept” of effectiveness and intensity for further planning or to determine the need for special services. Technical checks (e.g., intervention adherence and observer agreement on key measures) may be co-plotted on figures to aid team members’ decisions (Ehrhardt et al., 1996).

## Challenges and Questions

There are challenges to using intervention data sets for teams considering eligibility decisions. Beyond those discussed earlier, significant variables may be difficult to identify or change because different school contexts and levels of skill and tolerance displayed by change agents may influence plans (Gerber, 1988). Another question is how much an intervention-based model “costs” (i.e., how much professional time it takes to implement). However, several studies show that cost may be highly variable and related to team fluency in problem solving and not necessarily child characteristics, yet is still similar overall in time to “testing” children (e.g., Barnett, Air, et al., 1999). Other questions, addressed next, may relate to design requirements, risks in behavioral programming, and criteria for a technically adequate data set for special services eligibility decisions.

## Design Control and Experimental Validity Questions

In single-case research, the researcher determines the design and makes changes based on ongoing data analysis. Compliance with the intent of IDEA and the consultation foundation results in different questions of design control. Team decision-making and negotiation determines what is measured, what is changed, and what design components are used to address questions about special services.

In practice, designs may be used for decision purposes that weaken internal validity arguments. For example, when two or more interventions are introduced, the intervention sequences, combinations, or scheduling (i.e., time between interventions) may make it difficult to determine which effects are attributable to specific interventions (termed *multiple treatment interference*). Level and trend changes may be immediate and significant following the introduction of the intervention, contributing to internal validity evidence. Also, several tactics may be used by teams to strengthen internal validity arguments. These include using interventions with established empirical evidence and applying control conditions (e.g., baseline in untreated conditions, brief withdrawals followed by reintroduction of intervention). Plan evaluation, to some degree, can take precedence over internal validity for specific components. Evidence may be marshaled for the effectiveness of a plan, and this is an adequate criterion for many educational decisions.

## Potentials and Risks

First, design efficacy depends on reasonable selections of (a) a response class and sensitive measures related to school suc-

TABLE 2. Hypothetical Examples: The Graphic Analysis of Intervention Intensity

Analyses	Figure 1: Increasing design	Figure 2: Decreasing design
Intervention effectiveness	The data display oral-reading fluency rates consistent with grade-level expectations only in the context of Phase D.	The data show that classroom behavior improved during Phase BCD, maintained by Phase BC.
Intervention intensity	Phase D is characterized by 100 min of treatment per week, delivered outside of routine instruction by a trained adult specialist (modifications of classroom activities).	Class-wide activity schedules were incorporated into teacher routines. Classroom tokens and reinforcement delivered required brief teacher contacts with student during 48% of intervals for academic activities (child monitoring).
Functional discrepancy analyses and team judgment	Figure shows there was significant performance discrepancy between target child and peer norms when Phase D was not in place. Intervention intensity is consistent with uniquely designed and specialized instruction.	Graph shows significant discrepancy in the two measured variables. Other interventions need to be planned (e.g., academics, self-regulation) to further reduce teacher monitoring. Data showed that a significant amount of teacher time was spent implementing interventions.

cess and (b) an initial intervention that has empirical support for effectiveness, with careful analysis conducted of presenting problems—poor selection of significant variables for change, weak initial interventions, or low intervention adherence can lead to false conclusions about necessary intervention intensity. A positive response is a characteristic of an intervention context, and intervention strength always relates to idiosyncratic variables. For example, a reading intervention (e.g., repeated reading) may cause significant gains in reading fluency with one student and little improvement with a second student, even when problem analyses are similar. Second, unnecessarily intensive multicomponent strategies can be costly and restrictive for the student, and relatively simple interventions (e.g., contingency management) may be successful even when behavior or academic performance is highly problematic. Third, the use of functional analysis and graphic analysis of data holds potential for interpretive errors (e.g., Gresham et al., 2001).

Measurement targets for many intervention decisions include reference to analysis of academic environments (Lentz & Shapiro, 1986; Martens & Kelly, 1993; Shapiro, 1996). Depending on the presenting problems, the first intervention phase (commonly referred to as first phase, tier, or level; e.g., Fuchs & Fuchs, 1998; O'Shaughnessy et al., 2003) in a design of increasing or decreasing intensity may be conducted class-wide to ensure that the class is overall effectively instructed and managed, with the next phase focused on residual concerns not affected by the class-wide intervention. Class-wide interventions pertain to interventions that target instruction, activities, routines, or times of the day rather than the behavior of an individual child (e.g., Paine, Radicchi, Rosellini, Deutchman, & Darch, 1983). The idea is that improving classrooms caused learning to increase and overall disruptive behaviors to decrease. Examples include designing an effective physical layout for the classroom, improving instructional and managerial methods and routines, clarifying and teaching guides or rules and expectations, and limiting ineffective requests or prompts (e.g., Fowler, 1986; Fuchs & Fuchs, 1998; Paine et al., 1983; VanDerHeyden, Witt, & Gatti, 2001). If class-wide interventions are successful and sustainable, there is no need for more structured individual special services, and the amount of intervention effort calculated for an individual child would be zero—with a potential of high payoff for many.

Following class-wide intervention, small-group or embedded interventions (a second phase, tier, or level) can be used to address children's additional needs (Daugherty, Grisham-Brown, & Hemmeter, 2001; Wolery, 1994). Embedded discrete trial interventions provide supports and adaptations for small groups or individual children integrated into classroom routines and activities and may include increasing opportunities to practice academic, language, or social skills; modification of entry and transition routines; or peer tutoring. Finally, as a third phase or tier, based on insufficient response to class-wide or small-group/embedded interventions, interventions become more specialized and individualized. As learning or

performance objectives are met, design intensities are decreased through the service-delivery phases or tiers.

We acknowledge that our suggestions depend on good problem-solving and accurate data-based hypotheses guiding the selection of initial interventions or effective broadband interventions. The successful application of increasing- or decreasing-intensity designs depends on the existence of many preconditions and steps outlined in the federal laws and initiatives reviewed earlier and technical adequacy concepts applied to problem-solving.

## Technical Adequacy

The crux of many discussions of reform has been to shift the focus of disability analysis in special education evaluation from psychometric criteria to behaviors in natural settings as a basis for intervention planning and service-delivery allocation. Different technical adequacy concepts apply. One requirement is the ongoing measurement of a problem or instructional situation to distinguish between those that respond positively to intervention efforts and those that cannot be logistically supported without special services. An intervention-based data set for eligibility for special services would include patterns of discrepancies over time, demonstrating necessary interventions (including supports and services) required to address the discrepancies and thus aid the judgment of teams. The contextual measurements and designs simultaneously demonstrate performance discrepancies, children's response to interventions, need for more or less intense services, and qualities of needed services. The following methods of demonstrating technical adequacy have been reviewed by Macmann et al. (1996; decision reliability and validity): (a) demonstrating reasonableness of target variable selection and measurement tactics leading to the accurate and reliable description of a functional discrepancy, (b) showing defensible intervention selection (i.e., empirically demonstrated in a sound sequence), (c) documenting intervention adherence, and (d) measuring intervention outcome. Technical adequacy in response to intervention designs has evolved to include (e) applying decision rules for within and between tier or level intervention changes (i.e., applying level and trend analyses) and (f) judging intervention intensity.

The legal requirement is met by using data organized by a design to help establish present levels of performance, proactive or prevention conditions suggested by IDEA (e.g., Drasgow & Yell, 2001; Fuchs & Fuchs, 1998; Tilly et al., 1999), and the educational needs (required components of IEPs) to support service delivery in the least restrictive environment, to enable children to meaningfully participate in typical classroom activities. In fact, the idea of increasing-intensity assessment defines and clarifies the least restrictive environment prior to the need for classification. Given adequate technical adequacy, increasing- and decreasing-intensity designs can be used as decision aids, to determine the least intrusive and most efficient and effective strategies.

Teams may need to consider (a) local, class, or micro-norms used for goal-setting and progress-monitoring for typical children (Bell & Barnett, 1999; Fuchs & Fuchs, 1998) and (b) a data path for children referred because of problematic behavior or performance. The amount of intervention and performance or behavioral gain (level and trend) in comparison to typical peers may form part of the basis for the analysis of intensity. Although the idea of local norm use may raise some questions, there are empirical, practical, and legal foundations for using local norms to address resource questions. First, there is evidence that students who are identified as having learning disabilities are the lowest achieving students, according to local norms (Shinn, Good, & Parker, 1999). Second, repeated assessment of local norms provides estimates of expected growth rates or trends. Thus, decisions regarding a referred child's progress can be based on converging or diverging trends, rather than absolute discrepancies at one point in time (Tilly et al., 1999). Third, the U.S. Supreme Court has used local community standards when defining the least restrictive environment clause of IDEA (*Board of Education v. Rowley*, 1982). Finally, the use of local norms does not rule out comparisons with norms from other data sources, which may be important.

Questions also may be raised about multifaceted evaluation requirements in light of typically narrower targets associated with single-case designs. A "single target behavior" is not necessarily problematic in our view because a multifaceted evaluation, as defined by IDEA, includes all prior information (e.g., any standardized test scores, disciplinary referrals, developmental and educational histories) and clarifies needed information. Thus, IDEA requirements pertain to participation in general education and the idea of incremental decision-making. Although children's concerns may be multifaceted, teams often target one or a few response classes directly related to general educational programming in order to clarify the intervention targets and services that a team deems necessary for the child to participate in general education.

## Conclusions

One of the most compelling current ideas in special education is that providing accurate empirical appraisals of children's responses to significant intervention qualities within typical settings will help address resource questions directly related to disability evaluation. We argue that through an empirical, step-by-step process, the least restrictive environment can be judged in a valid manner. Using instructional and intervention trials as data organizers within increasing- and decreasing-intensity designs may allow for the analysis of appropriate intervention intensity and help map out the actual decisions made by school-based intervention teams.

There are no formulas that are both simple and technically defensible for identifying educational disabilities, especially when service-delivery questions are raised, and there is

nothing easy about the process we suggest. Decisions need to be made in the context and setting of a child's school by persons who are knowledgeable about children, resources, and issues of how to analyze the amount of effort and intensity required to accelerate the child's academic performance or sustain appropriate behavior. It is the responsibility of a team of professionals and parents to justify the decision based on an acceptable evaluation plan and supporting data. Single-case designs and direct assessment data can supply the basic database for these decisions, with the foundation being positive change in response to the least and most natural intervention. Each successive intervention plan is hierarchically linked to resources and intervention elaboration *or* reduction, withdrawal, or fading, through the process of problem-solving, collecting data, measuring children's responses to interventions, and graphing results. The intervention-based data set guides decisions about special services. Our goal has been to outline how scientifically based methods and evaluation designs could be integrated in response to intervention models, to stimulate and guide practice and future research on the technical qualities of different operational definitions of response to intervention.

## AUTHORS' NOTE

We would like to thank Anna Adams and Seena Skelton for their helpful feedback.

## REFERENCES

- Anita, S., & Kreimeyer, K. (1988, December). Maintenance of positive peer interaction in preschool hearing-impaired children. *The Volta Review*, 325-337.
- Barlow, D. H., Hayes, S. C., & Nelson, R. O. (1984). *The scientist-practitioner: Research and accountability in clinical and educational settings*. New York: Pergamon.
- Barnett, D. W., Air, A. E., Bell, S. H., Gilkey, C. M., Smith, J. J., Stone, C. M., et al. (1999). Evaluating early intervention: Accountability methods for service delivery innovations. *The Journal of Special Education*, 33, 177-188.
- Barnett, D. W., Bell, S. H., Gilkey, C. M., Lentz, F. E., Jr., Graden, J. L., Stone, C. M., & Smith, J. J. (1999). The promise of meaningful eligibility determination: Functional intervention-based multifaceted preschool evaluation. *The Journal of Special Education*, 33, 112-124.
- Barnett, D. W., Daly, E. J. III, Hampshire, E. M., Hines, N. R., Maples, K. A., Ostrom, J. K., & Van Buren, A. (1999). Meeting performance-based training demands: Accountability in an intervention-based practicum. *School Psychology Quarterly*, 14, 357-379.
- Barnett, D. W., Lentz, F. E., Jr., & Macmann, G. M. (2000). Child assessment: Psychometric qualities of professional practice. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment* (2nd ed., pp. 355-386). New York: Guilford Press.
- Bell, S. H., & Barnett, D. W. (1999). Peer micronorms in the assessment of young children: Methodological review and examples. *Topics in Early Childhood Special Education*, 19, 112-122.
- Board of Education of Hendrick Hudson Central School District v. Rowley. 176, 102 S. Ct. 3034, 73, (1982).
- Cooper, L. J., Wacker, D. P., Thursby, D., Plagmann, L. A., Harding, J., Millard, T., & Derby, M. (1992). Analysis of the effects of task preference, task demands, and adult attention on child behavior in outpatient and classroom settings. *Journal of Applied Behavior Analysis*, 25, 823-840.
- Daly, E. J. III, Lentz, F. E., Jr., & Boyer, J. (1996). The instructional hierarchy: A conceptual model for understanding the effective components of reading interventions. *School Psychology Quarterly*, 11, 369-386.

- 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. 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, Murdoch, A., Lillenstein, L., Webber, L., & Lentz, F. E., Jr. (2002). An examination of methods for testing treatments: Conducting brief experimental analyses of the effects of instructional components on oral reading fluency. *Education and Treatment of Children*, 25, 288-316.
- 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.
- Daugherty, S., Grisham-Brown, J., & Hemmeter, M. L. (2001). The effects of embedded skill instruction on the acquisition of target and nontarget skills in preschoolers with developmental delays. *Topics in Early Childhood Special Education*, 21, 213-221.
- Deno, S., & Gross, J. (1973). The Seward-University Project: A cooperative effort to improve school services and university training. In E. N. Deno (Ed.), *Instructional alternatives for exceptional children* (pp. 104-121). Arlington, VA: Council for Exceptional Children.
- Dooley, P., Wilczenski, F. L., & Torem, C. (2001). Using an activity schedule to smooth school transitions. *Journal of Positive Behavioral Interventions*, 3, 57-61.
- Drasgow, E., & Yell, M. L. (2001). Functional behavioral assessments: Legal requirements and challenges. *School Psychology Review*, 30, 239-251.
- Eckert, T. L., Ardoin, S. P., Daly, E. J. III, & Martens, B. K. (2002). Improving oral reading fluency: An examination of the efficacy of combining skill-based and performance-based interventions. *Journal of Applied Behavior Analysis*, 35, 271-281.
- Edgington, E. S. (1983). Response-guided experimentation. *Contemporary Psychology*, 28, 64-65.
- Ehrhardt, K. E., Barnett, D. W., Lentz, F. E., Jr., Stollar, S. A., & Reifin, L. H. (1996). Innovative methodology in ecological consultation: Use of scripts to promote treatment acceptability and integrity. *School Psychology Quarterly*, 11, 149-168.
- Ervin, R. A., Radford, P. M., Bertsch, K., Piper, A. L., Ehrhardt, K. E., & Poling, A. (2001). A descriptive analysis and critique of the empirical literature on school-based functional assessment. *School Psychology Review*, 30, 193-210.
- Fowler, S. (1986). Peer-monitoring and self-monitoring: Alternatives to traditional teacher management. *Exceptional Children*, 52, 573-581.
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53, 199-208.
- Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice*, 13, 204-219.
- Gerber, M. M. (1988). Tolerance and technology of instruction: Implications for special education reform. *Exceptional Children*, 54, 309-314.
- Gresham, F. M. (1989). Assessment of treatment integrity in school consultation/prereferral intervention. *School Psychology Review*, 18, 37-50.
- Gresham, F. M. (1991). Conceptualizing behavior disorders in terms of resistance to intervention. *School Psychology Review*, 20, 20-36.
- Gresham, F. M. (2001). *Responsiveness to intervention: An alternative approach to the identification of learning disabilities*. Paper prepared for the OSEP Learning Disabilities Initiative, Washington, DC: U.S. Department of Education, Office of Special Education Programs.
- Gresham, F. M., Watson, T. S., & Skinner, C. H. (2001). Functional behavioral assessment: Principles, procedures, and future directions. *School Psychology Review*, 30, 156-172.
- Gresham, F. M., & Witt, J. C. (1997). Utility of intelligence tests for treatment planning, classification, and placement decisions: Recent empirical findings and future directions. *School Psychology Quarterly*, 12, 249-267.
- Harding, J., Wacker, D. P., Cooper, L. J., Millard, T., & Jensen-Kovalan, P. (1994). Brief hierarchical assessment of potential treatment components with children in an outpatient clinic. *Journal of Applied Behavior Analysis*, 27, 291-300.
- Hardman, M. L., McDonnell, J., & Welch, M. (1997). Perspectives on the future of IDEA. *Journal of the Association for Persons with Severe Handicaps*, 22, 61-77.
- Haring, N. G., & Eaton, M. D. (1978). Systematic instructional procedures: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton, & C. L. Hansen (Eds.), *The fourth R: Research in the classroom* (pp. 23-40). Columbus, OH: Merrill.
- Horner, R. H., & Sugai, G. (2000). School-wide positive behavioral support: An emerging initiative. *Journal of Positive Behavioral Interventions*, 2, 231-232.
- Individuals with Disabilities Education Act Amendments of 1997, 20 U.S.C. § 140 (26).
- Kazdin, A. E. (1982). *Single case research design*. New York: Oxford University Press.
- Kratochwill, T. R., Elliott, S. N., & Busse, R. T. (1995). Behavioral consultation: A five-year evaluation of consultant and client outcomes. *School Psychology Quarterly*, 10, 87-117.
- LeLaurin, K., & Wolery, M. (1992). Research standards in early intervention: Defining, describing, and measuring the independent variable. *Journal of Early Intervention*, 16, 275-287.
- Lentz, F. E., Jr., Allen, S. J., & Ehrhardt, K. E. (1996). The conceptual elements of strong interventions in school settings. *School Psychology Quarterly*, 11, 118-136.
- Lentz, F. E., Jr., & Shapiro, E. S. (1986). Functional assessment of the academic environment. *School Psychology Review*, 15, 346-357.
- Macmann, G. M., Barnett, D. W., Allen, S. J., Bramlett, R. K., Hall, J. D., & Ehrhardt, K. E. (1996). Problem solving and intervention design: Guidelines for the evaluation of technical adequacy. *School Psychology Quarterly*, 11, 137-148.
- 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.
- Martens, B. K., & Kelly, S. Q. (1993). A behavior analysis of effective teaching. *School Psychology Quarterly*, 8, 10-26.
- McComas, J. J., & Mace, F. C. (2000). Theory and practice in conducting functional analysis. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment* (2nd ed., pp. 78-103). New York: Guilford Press.
- McConnell, S. R., Rush, K. L., McEvoy, M. A., Carta, J., Atwater, J., & Williams, R. (2002). Descriptive and experimental analysis of child-caregiver interactions that promote development of young children exposed prenatally to drugs and alcohol. *Journal of Behavioral Education*, 11, 131-161.
- National Information Center for Children and Youth with Disabilities. (1999). *Interventions for chronic behavior problems*. Washington, DC: Author.
- Noell, G. H., & Gresham, F. M. (1993). Functional outcome analysis: Do the benefits of consultation and prereferral intervention justify the cost? *School Psychology Quarterly*, 8, 200-227.
- Northup, J., & Gully, V. (2001). Some contributions of functional analysis to the assessments of behaviors associated with attention deficit hyperactivity disorder and the effects of stimulant medication. *School Psychology Review*, 30, 227-328.
- O'Neill, R. E., Horner, R. H., Albin, R. W., Sprague, J. R., Storey, K., & Newton, J. S. (1997). *Functional assessment and program development for problem behavior: A practical handbook* (2nd ed.). Pacific Grove, CA: Brookes/Cole.
- O'Shaughnessy, T. E., Lane, K. L., Gresham, F. M., & Beebe-Frankenberger, M. E. (2003). Children placed at risk for learning and behavior difficulties. *Remedial and Special Education*, 24, 27-35.

- Paine, S. C., Radicchi, J., Rosellini, L. C., Deutchman, L., & Darch, C. B. (1983). *Structuring your classroom for academic success*. Champaign, IL: Research Press.
- Pennypacker, H. S., & Hench, L. L. (1998). Making behavioral technology transferable. *The Behavior Analyst, 20*, 97–108.
- Reschly, D. J., & Ysseldyke, J. E. (2002). Paradigm shift: The past is not the future. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 3–20). Bethesda, MD: The National Association of School Psychologists.
- Rhymer, K. N., Dittmer, K. I., Skinner, C. H., & Jackson, B. (2000). Effectiveness of a multi-component treatment for improving mathematics fluency. *School Psychology Quarterly, 15*, 40–51.
- Rusch, F. R., Connis, R. T., & Sower, J. (1979). The modification and maintenance of time spent attending to task using social reinforcement, token reinforcement and response cost in an applied restaurant setting. *Journal of Special Education Technology, 2*, 18–26.
- Rusch, F. R., & Kazdin, A. E. (1981). Toward a methodology of withdrawal designs for the assessment of response maintenance. *Journal of Applied Behavior Analysis, 14*, 131–140.
- Sainato, D. M., Strain, P. S., Lefebvre, D., & Rapp, N. (1990). Effects of self-evaluation on the independent work skills of preschool children with disabilities. *Exceptional Children, 56*, 540–549.
- Sainato, D. M., Strain, P. S., & Lyon, S. R. (1987). Increasing academic responding of handicapped preschool children during group instruction. *Exceptional Children, 56*, 540–549.
- Sechrest, L., West, S. G., Phillips, M. A., Redner, R., & Yeaton, W. (1979). Some neglected problems in evaluation research: Strength and integrity of treatments. In S. Sechrest, S. G. West, M. A. Phillips, R. Redner, & W. Yeaton (Eds.), *Evaluation studies: Review annual* (Vol. 4, pp. 15–35). Beverly Hills, CA: Sage.
- Shapiro, E. S. (1996). *Academic skills problems: Direct assessment and intervention* (2nd ed.). New York: Guilford Press.
- Shapiro, E. S., & Kratochwill, T. R. (Eds.). (2000). *Behavioral assessment in schools: Theory, research, and clinical foundations* (2nd ed.). New York: Guilford Press.
- Sheridan, S. M., Kratochwill, T. R., & Elliott, S. N. (1990). Behavioral consultation with parents and teachers: Delivering treatment for socially withdrawn children at home and school. *School Psychology Review, 19*, 33–52.
- Shinn, M. R., & Bamonto, S. (1998). Advanced applications of curriculum-based measurement: “Big ideas” and avoiding confusion. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 1–31). New York: Guilford Press.
- Shinn, M. R., Good, R. H., & Parker, C. (1999). Noncategorical special education services with students with severe achievement deficits. In D. J. Reschly, W. D. Tilly, & J. P. Grimes (Eds.), *Special education in transition: Functional assessment and noncategorical programming* (pp. 81–105). Longmont, CO: Sopris West.
- Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. Boston: Authors Cooperative.
- Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93*, 735–749.
- Stoiber, K. C., & Kratochwill, T. R. (2000). Empirically supported interventions and school psychology: Rationale and methodological issues. Part I. *School Psychology Quarterly, 15*, 75–105.
- Tilly, D. W., Reschly, D. J., & Grimes, J. (1999). Disability determination in problem solving systems: Conceptual foundations and critical components. In D. J. Reschly, W. D. Tilly, & J. P. Grimes (Eds.), *Special education in transition: Functional assessment and noncategorical programming* (pp. 221–251). Longman, CO: Sopris West.
- U.S. Department of Education, Office of Special Education and Rehabilitative Services. (2002). *A new era: Revitalizing special education for children and their families*. Washington, DC: Author.
- VanAuken, T. L., 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*, 163–179.
- VanDerHeyden, A. M., Witt, J. C., & Gatti, S. (2001). Descriptive assessment method to reduce overall disruptive behavior in a preschool classroom. *School Psychology Review, 30*, 548–567.
- Van Houten, R., & Hall, R. V. (2001). *The measurement of behavior: Behavior modification*. Austin, TX: PRO-ED.
- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (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.
- Vollmer, T. R., & Van Camp, C. M. (1998). Experimental designs to evaluate antecedent control. In J. K. Luiselli & M. J. Cameron (Eds.), *Antecedent control: Innovative approaches to behavioral support* (pp. 87–111). Baltimore: Brookes.
- Wacker, D. P., Steege, M., & Berg, W. K. (1988). Use of single-case designs to evaluate manipulable influences on school performance. *School Psychology Review, 17*, 651–657.
- Wolery, M. (1994). Implementing instruction for young children with special needs in early childhood classrooms. In M. Wolery & J. S. Wilburs (Eds.), *Including children with special needs in early childhood programs* (pp. 151–166). Washington, DC: NAEYC.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214.
- Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology, 49*, 156–167.