Chapter 2

Choosing Targets for Assessment and Intervention

Improving Important Student Outcomes

Renee O. Hawkins
David W. Barnett
Julie Q. Morrison
Shobana Musti-Rao

This chapter was written to help guide professionals through key decision points in identifying problems that should be targeted for intervention and in determining how to measure the targets. Each of these decision points has a direct impact on student interventions and outcomes and is guided by available research and data-based problem solving.

Target variable selection refers to problem-solving teams identifying targets for intervention and ways to measure those targets, whether the problem occurs at an individual student, class, or schoolwide level. *Target variables are derived from constructs of educational risk and yield specific measures or observations to identify students in need of intervention support and to track intervention outcomes. Both target variables and measures need to be carefully selected by teams because the measures are used to construct the database for monitoring and evaluating intervention programs.* Thus target variables and measures set the course of action by teams and act as the "heart monitor" for educational services, allowing for timely modifications as needed by showing initial risk for academic or behavioral difficulties, as well as ongoing intervention effectiveness.

In most cases, target *variable* selection is used instead of the more traditional target "behavior" because data-based problem solving is increasingly linked to school, classroom, or setting characteristics that may be outcomes of problem solving and schoolwide programmatic changes. These intervention setting characteristics may be progress-monitored when students are referred or screened for concerns about their behavior or academic performance in school. For example, students may be referred for academic failure problems, but schools may need to measure not only student progress but also the amount and quality of instruction provided to students, as instruction may need to be changed and monitored. In keeping with traditional discussions, specific instructional procedures are the *intervention* for a targeted student, and the student's performance is the *behavior.* However, the need for ongoing selection, monitoring, and modification of instruction programs at school, class, and individual levels blurs the traditional distinctions in what is typically targeted for change (i.e., student behavior or instruction). Measurement focusing on targeted variables includes *behavior in environment* and functions of behavior, and in educational programming key features of in-

structional environments may be significant targets for change.

The importance of decisions for students and stakeholders related to target variable selection may be quite high, and teams will wonder about the adequacy of different variables or alternative methods of measuring variables. Technical adequacy (e.g., reliability and validity, sampling) of target variable measurement is discussed as a way to increase team confidence that sound decisions have been made.

A related task for problem-solving teams is identifying *students* requiring intervention based on specific risk indicators. Relying on target variable data related to school success, teams use data to decide not only which students are in need of support but also how many students and at which levels to intervene in order to effectively measure and interpret outcomes (i.e., school, class, group, and individual levels). *Risk indicators* are factors or measures that suggest the likelihood of students' school success or failure. Teams make efforts to ensure the accurate identification of students in need of intervention to prevent students from falling further behind and to improve the use of school resources through effective programming, as, for example, raising school or class performance if many students are at risk.

First, an overview of target selection basics and guidelines for choosing target variables is provided. Second, methods of selecting students for interventions are described, as the methods and outcomes affect which variables are targeted and the appropriate level of intervention and monitoring. Third, advances in target variable selection are discussed in the context of response to intervention (RTI).

Target Variable Selection

Generally, problem-solving teams start with broad targets for change and use a "funneling" process (Hawkins, 1986) to narrow the focus in selecting target variables. The measurement of target variables is direct, contextualized by settings, and functional: What is happening in a specific situation that is concerning? and What can and should be changed? Changes in socially significant performance are what matters. Data regarding environmental and instructional variables, as well as technical adequacy (e.g., reliability and validity evidence), may be needed to defend specific team decisions concerning variables targeted for change. Students and situations can be highly challenging, and target variable selection is approached step by step, repeating problem-solving steps as needed.

Overview: The Basics of Target Variable Selection through Problem Solving

Target variables are selected based on the use of intervention research and behavioral problem solving. School psychologists trained in consultation team with teachers and, as appropriate, specialized professionals, students, and parents to resolve problem situations. Problem solving may be used repeatedly to help meet long-term objectives, as for a student with comprehensive socially and educationally related disabilities (Kratochwill & Bergin, 1990).

In classic discussions, a student's unique characteristics and situations are guiding factors in target variable selection. Kanfer (1985) wrote: "Each client presents the clinician anew with the fundamental task of deciding on a focus for the most effective intervention" (p. 7). Target variable selection steps generally include problem identification and analysis and continue with plan development, plan implementation, and evaluation. There is a creative process in which alternatives are considered and a rigorous progression links all steps with the best available empirical evidence.

In the problem identification step, teams clarify the problem behavior and desired alternative. In problem analysis, decisions also may be made to monitor environmental variables related to problem situations as significant factors that contribute to differences between observed and desired behavior and performance become understood. In applied behavior analysis (ABA), from which fundamental intervention methods are derived, an emphasis is on the use of high-quality data for decision making and the evaluation of interventions using single-case methods (baseline followed by intervention; see Daly, Barnett, Kupzyk, Hofstadter, & Barkley, Chapter 29, this volume). In addition, functional relationships are considered when making

predictions about behavior. Through functional behavioral assessment (FBA; see Jones & Wickstrom, Chapter 12, this volume) and analysis methods, functional hypotheses are generated to understand relationships between target behaviors and environmental variables. A *functional hypothesis* is a proposed explanation as to the reason that problem behaviors occur and persist, such as gaining attention or escaping difficult tasks (expanded later in the chapter). ABA also emphasizes achieving *social validity*, which means in part that persons in close contact with students have a voice in considering the goals, methods, and outcomes of prevention and intervention programs (Wolf, 1978).

The Scope of Target Variables Has Appropriate Focus

Following careful selection of significant variables, teams must make measurement decisions. *Comprehensiveness* as used here refers to the many possible and intervention-relevant considerations related to academic and social performance in schooling (e.g., medical problems; home setting events that interfere with sleep, nutrition, homework completion; generalization of social or academic skills; etc.; Gresham, 2007). With the idea of *level of analysis*, picture using "zoom in" or "zoom out" when examining situations or behaviors. *Splitting* or *lumping* occurs as complex skills are used as variables targeted for change and as teams focus on specific hypothesized variables of importance in problem solving. For example, social competence is made up of many social skills (e.g., social problem solving, eye contact) that must be contextually and developmentally appropriate for intervention plans. Reading can be broken down into requisite skills such as vocabulary, phonemic awareness, and so on, based on functional and empirical hypotheses of what is needed. Teams monitor progress at the construct level (e.g., "reading and social behavior are improving") by using measures validly related to the improvements and the intervention methods (e.g., greater reading fluency through more practice time, fewer arguments with peers based on applying problem-solving skills).

Some concerns may require measuring a *constellation of behaviors* (Kazdin, 1985). For example, student anxiety or depression may have various degrees and expressions of overt behavior and covert events that may be exacerbated by incidents in school or home.

Stimulus and response patterns or *covariations* (Kazdin, 1985) may need to be measured. For example, to measure student compliance as a targeted variable, teams may need to look at the following: clarity of classroom rules or expectations; a student's fluency with behavioral expectations; various qualities of a teacher's request, such as whether it is said nicely but firmly, with eye contact, and in proximity to the student, whether it can be done without supports, and whether wait time is appropriate; student's behavior or compliance with requests; peer norms for compliance (Bell & Barnett, 1999); teacher's reinforcement of compliance; and sustained compliance. Decisions about what to target and selection of interventions to improve compliance are linked to what the data say about a student's compliance in context.

Scheduling Targeted Variable Measurements for Progress Monitoring

After variables and measures are selected, teams decide *when* and *how much* to measure and *at what point to analyze the data*. Targeted variables are measured to establish a baseline of current performance and to closely monitor intervention effects or "what is happening" so that timely changes can be made to plans as needed. In practice, *decision rules* are set with team members about what constitutes adequate plan implementation, how long to try the plan, and measurable goals or criteria for performance. A decision rule is used to link data to instructional decision making through carefully planned instructional trials to see whether changes in methods or content are needed (e.g., Fuchs & Fuchs, 1986). Thus a decision rule is an agreement or plan to carefully try an intervention for a set time (or number of trials, etc.) to see what changes in plans may be needed based on the data. Decision rules can improve decision making by providing timely feedback to teams on "what works." New decision rules are reset after each point-of-intervention evaluation.

Schedules for collecting data on targeted variables may vary widely. The schedule for monitoring should be based on specific research with the intervention and target

variable measures and on the realities of situations. Also, teams should evaluate the amount of risk associated with the ongoing occurrence of the problem behavior for the targeted student, as well as for others in the environment. For guidelines, high-risk behaviors may be monitored every day to once per week or every 2 weeks for academic performance to allow for measurable growth to triannually for academic screening programs. As examples, it may be acceptable to monitor writing fluency weekly, whereas highly disruptive behavior or physical aggression toward peers may be monitored daily to quickly identify an effective intervention plan and ensure the safety of all students. Schedules are modified as needed based on what the data indicate (e.g., changes in level or trend). To help with the scheduling challenges, different data sources are used, and, as situations improve, follow-up measures become less frequent. For example, for challenging behaviors, a teacher daily report is used, along with periodic direct observation by a consultant (the data sources should not be combined but should show separate results). Schedules and organization of data collection also are linked to single-case designs (see Daly et al., Chapter 29, this volume).

Guidelines for Selecting Target Variables

Guidelines help teams with sound decisions regarding target variable selection. Target variables should be linked to direct measures of the problem that are reliable, sensitive enough to measure change resulting from the intervention, and related to valid outcomes (Macmann et al., 1996). Table 2.1 summarizes practical guidelines for selecting targeted variables and measures building on classic discussions (e.g., Hawkins, 1986; Kratochwill, 1985). Basic reliability and validity information and other measurement qualities are ways to help with the choices in target variable measure selection.

Target Variables Are Clearly Defined

Target variables are defined in observable and measurable terms and in ways that all members of the problem-solving team can understand. Operational definitions clearly and objectively describe the observable features of the behavior. They include examples and nonexamples of behavior and provide a complete picture of what the target behavior looks like (Hawkins & Dobes, 1975).

Target Variables Can Be Significantly Changed

Teams select target variables that can be meaningfully changed in that they are influenced by the environment. Target variable measures should be sensitive enough to reflect changes in behavior resulting from prevention and intervention programming. Many examples show why the idea of changeability is important. Personalities, temperament, intelligence, and self-concept are mentioned frequently as concerns or explanations in consultations with parents or teachers. However, these attributes are not easily modifiable as targets of interventions, and typically measures of these constructs are not useful for progress monitoring. Through effective problem solving, variables can be selected that satisfy concerns but that also yield measures that are practical and valuable for progress monitoring. Examples include targeting and improving academic and classroom functioning and social competence skills, as well as supports for teachers and students. Taking broadly stated concerns of parents and teachers, finding sound ways to select target variables related to valid concerns, and progress-monitoring interventions are basic functions of problem solving.

Target Variables Can Be Directly Measured

Target variables are directly linked to the problem situation by carefully selected measurement methods. Intervention research is used to help achieve confidence in measurement and intervention plans by using high-quality data to evaluate instructional and behavioral outcomes. For academic concerns, curriculum-based measurement (CBM) is a well-researched, reliable, and direct method for measuring student performance in core academic areas including reading, math, written expression, and spelling (Deno, Marston, & Tindal, 1985–1986; see also Marcotte & Hintze, Chapter 5; Burns & Klingbeil, Chapter 6; and Gansle & Noell, Chapter 7, this volume). Table 2.2 describes common CBM variables. CBM allows the

TABLE 2.1. Practice Guidelines for Target Variable Selection

Professional standard	What to look for
Target variable measures meet scientific and professional standards of "high-quality data"	• Validity: teams use variables linked to specific prevention and intervention research or establish the functional validity of the variables (i.e., demonstrate its validity for an individual; demonstrate causal relationship)
Problem solving is used to form empirically valid plans	• Targeted variables may include environmental, instructional, and student measures of change as needed. • Validity and level of inference: Outcomes are directly observable and meaningful. • Validity and sensitivity: Teams use measures that can track changes in behavior or performance in increments that are useful for ongoing and timely decisions. • Social validity: Consumers of services (i.e., teachers, parents, and students) also evaluate intervention goals, methods, and outcomes. • Reliability: Teams use measures with known and acceptable reliability or ensure reliability through ongoing checks (i.e., agreement checks between observers).
Cost–benefit and sustainability are considered in making selections	• Costs are estimated by also considering potential outcomes. High-quality data may be needed to obtain high-quality results that can produce ultimate "savings" for students and schools. Intervention failure is costly.
Decisions are monitored carefully	• Decision rules are used whereby teams set goals and try out interventions for an agreed-on number of sessions based on research with the intervention. • Graphs are used to show the ongoing decision process, including baseline (if possible) and results of each condition.
Does the intervention work?	• Interventions are examined through an internally valid research design. Alternatively, schools can use an "accountability design" by looking at changes in performance or behavior as measured by carefully selected target variables with the intervention in place (see Daly et al., Chapter 29, this volume).
How well does it work?	• Questions that can be addressed by teams include the size and significance of effects, as compared with benchmarks, peer norms, and judgments by consumers. These data lead to the next steps by teams. • Broader consequences are considered, including planned as well as unplanned outcomes that may be positive or negative or may occur over longer time periods.

frequent collection of data to evaluate interventions. For example, students selected for small-group math instruction based on low performance on math CBM continue to be monitored weekly using math CBM.

For behavioral concerns, a high-quality data source for evaluating interventions is direct observation. There are several likely methods of collecting observational data, all based on selecting significant and consistent settings, times, conditions, or activities for observations. First, *time-sampling* procedures often are used to improve the technical adequacy of observational data. Observation sessions are divided into intervals (e.g., 10–30 seconds), and the variable of interest is recorded by set procedures (Cooper, Heron, & Heward, 2007). Observers record whether or not the behavior occurred continuously during the interval (e.g., 10 seconds) for *whole-interval recording*, at any point during the interval for *partial-interval recording*, or at the end of the interval for *momentary time sampling*. Second, in *event-recording* procedures, observers record features of behavior such as frequency. For both time sampling and event recording, the session is summarized using the data collected (e.g., student was engaged as a percent of intervals during a 20-minute session; the

TABLE 2.2. Academic Target Variables and Curriculum-Based Measurement

Variable	CBM
Reading fluency	Words correct per minute: Number of words correctly read aloud during 1-minute timed-reading probe
Math fluency	Digits correct per minute: Number of correct digits on timed (2–5 minutes) computation probe
Writing fluency	Total words written: Number of words written following a story starter during timed (3–5 minutes) probe
Spelling fluency	Correct letter sequences: Number of correct letter sequences during timed, dictated spelling probe

student talked out five times during a 20-minute lesson) over baseline and intervention sessions. Figure 2.1 shows an example of an observation system for engagement and how data would be graphed. As discussed in the following subsection, the graph also shows the results of a reliability check by a second observer, as noted by squares representing additional data points in the figure. Multiple variables can be measured simultaneously with more complex codes. For example, positive engagement is illustrated in Figure 2.1, a replacement behavior for inattentive, disruptive, or other concerning behavior that also may be measured in a code, along with instructional variables or a teacher's effective use of positive managerial practices. The variable of engagement also may be refined by measuring qualities of practice activities (e.g., Daly, Martens, Barnett, Witt, & Olson, 2007).

Target Variables Are Reliably Measured

At one level, teams agree on variables targeted for change and how to measure them and examine and resolve differences; this is the reliability or *consistency* of targeted variable *selection* across team members (Macmann et al., 1996). Evidence suggests that agreement on what to target may be a critical step, as team members may have different beliefs about causes of behavior and therefore about what to measure (e.g., Wilson & Evans, 1983). As selected, and throughout the problem-solving process, teams check the reliability of target variable measures. Reliability is estimated for some targeted variable measures (i.e., CBM) based on prior research. Ongoing reliability checks, also known as *agreement* checks, allow problem-solving teams increased confidence in measuring targeted variables (e.g., agreement on performance, frequency, duration, discrepancy from typical peer performance) and intervention effects. Reliability checks involve comparing the results of two observers independently coding or scoring the same sample of behavior for consistency. For example, two professionals may co-observe a classroom and compare data at the end of the observation session. Permanent products, such as a completed math CBM probe, may be independently scored and compared. Additional ongoing samples of CBM or observations can improve the reliability of individual decisions.

Technical adequacy checks for educational programming—as when targeted variables are curriculum, instructional skills, and behavioral management—are known variously as intervention adherence, fidelity of implementation, or intervention integrity. These checks are typically based on agreement indices showing the consistency of steps as carried out compared with implementation plans, scripts, and schedules (Barnett et al., 2007).

The operational definition of the target variable and the assessment system selected can significantly affect reliability. The reliability of data on the target variable and the validity of decisions made based on those data are improved when a precise behavioral definition is established. If the definition of the behavior is unclear, data are more likely to be unreliable, and teams will not be able to interpret the effects of interventions with confidence.

The method selected to assess the target variable also affects the reliability of data (Cooper et al., 2007). For example, for behaviors without a discrete beginning and end, such as student engagement, a time-sampling approach (as discussed earlier) would be most appropriate. Using a frequency count for such behaviors would likely re-

Code	Behavior	Definition	Recording Method
✓	Engagement	Student is attending to assigned task/activity by writing, reading, raising hand, asking or answering questions, talking to peers on topic, listening to the teacher or peers, looking at academic material	10-second, whole-interval recording

Minute	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

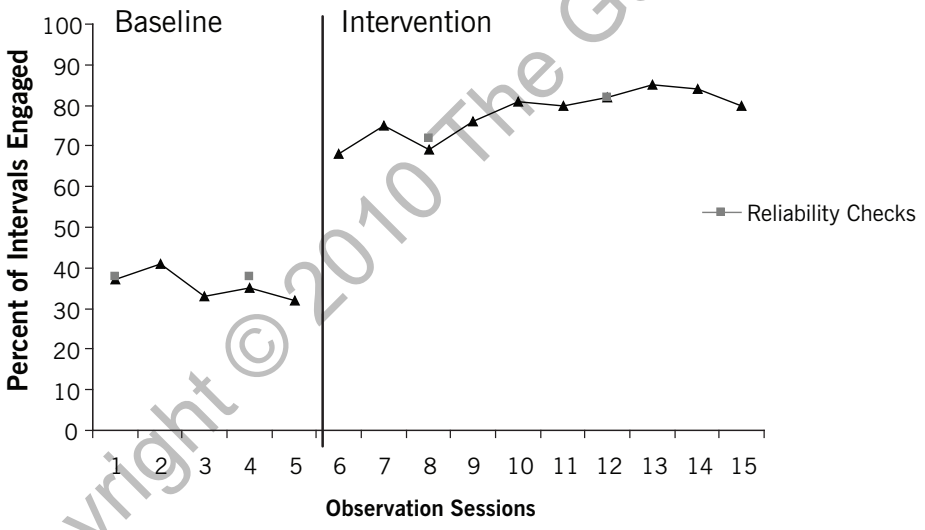

FIGURE 2.1. Example of observation code and graphed data.

sult in low reliability, as the observer would have difficulty determining when one occurrence of engagement ended and another began. Self-reports of teachers' adherence to intervention plans may not be equivalent to observational data by a consultant. Adequate training of those responsible for data collection and reliability checks can help improve reliability. Figure 2.1 provides an example of an operational definition for engagement and shows how reliability data can be coplotted to evaluate consistency of decisions that would be made by primary and other observers. Intervention adherence data also may be coplotted or summarized, along with student outcome data (Barnett et al., 2007).

Guidelines from research (e.g., Kennedy, 2005) suggest that reliability checks should occur for approximately 20–33% of the sample of observations across baseline and intervention, often using a criterion of at least 80% agreement. However, early in problem solving, more frequent reliability checks may

help teams evaluate the adequacy of operational definitions of targeted variables and of the data collection system and provide feedback to observers. The frequency of reliability checks also may depend on the amount of risk or severity of the problem behavior. For problem behaviors that require intense interventions, reliability checks can increase confidence in decisions concerning the use of resources to produce positive student outcomes.

Target Variables Are Linked to Meaningful Present and Future Outcomes

Direct and reliable measurements are necessary, but validity evidence related to positive outcomes for students adds even stronger criteria to target variable selection. Intervention research guides teams toward target variables that are linked to both short- and long-term positive outcomes (Kazdin, 1985). To accomplish this task, the selection of target variables and measures should be associated with evidence-based intervention methods that lead to meaningful change (Barnett et al., 2007). Teams weigh evidence by being up to date on specific intervention research to accurately judge current risk and make predictions about future consequences and to help select measurement methods. For example, numerous studies link specific intervention procedures with improved performance on curriculum-based assessment and measurement variables (Shapiro, 2004; Shinn, Walker, & Stoner, 2002). Thus, although increasing "engagement" is frequently selected as a target variable, the quality of practice opportunities afforded by increased engagement is the likely active ingredient in intervention success and can be progress-monitored (Daly et al., 2007). Teams may select opportunities to respond to academic stimuli (e.g., Greenwood, 1991) as a target variable leading to interventions that focus on providing students additional guided, independent, and generalized practice of skills, progress-monitored on progressively more natural and difficult material, all of which are linked to positive outcomes (e.g., Daly et al., 2007). As examples, selecting oral reading fluency as a target variable would lead to empirically supported interventions, such as repeated readings, peer tutoring, previewing strategies, taped-word procedure, and so on, and, through more practice on familiar and nonfamiliar material, to improved chances of school success (see Linan-Thompson & Vaughn, Chapter 16, this volume). Furthermore, although referral concerns may be extremely specific (e.g., hitting), targeted variables may be broader to ensure more significant positive behavior change (e.g., problem solving for increased social competence). Other criteria include possible benefits not only to the student but also to others in the environment, such as teachers and peers. The social validity of target variables is established when team members and data sources agree that targets, methods, and goals for change are acceptable (Wolf, 1978).

Functional Hypotheses Are Used to Critically Examine Possible Targeted Variables

Through FBA, information from a variety of methods, including interviews, questionnaires, rating scales, and direct observation, is used to hypothesize functional relationships between problem behaviors and features of the environment. Behavior or performance can "look" the same on the surface but can occur for very different reasons. FBA methods are a means of identifying these reasons to create effective interventions. Functional information about variables is used to develop intervention plans to decrease problem behavior and increase appropriate behavioral alternatives (see Jones & Wickstrom, Chapter 12, this volume; Gresham, Watson, & Skinner, 2001; Watson & Steege, 2003). The intensity of the FBA varies depending on the severity of the problem behavior. From relying solely on interviews to conducting multiple direct observations, teams can tailor the FBA process to meet students' needs. Furthermore, functional hypotheses can be directly evaluated to more clearly establish function. To establish *function* means that specific reasons for challenging behavior are tested. In a functional analysis, antecedent and consequence variables are experimentally manipulated to verify the function of behavior (Gresham et al., 2001; Horner, 1994). In some cases, teams also may conduct a brief experimental analysis in which different intervention conditions are presented and the effects compared to increase the validity of intervention selection

decisions (see Daly, Hofstadter, Martinez, & Andersen, Chapter 8, this volume). In this way functionally significant target variables linked to interventions can be clarified.

Prioritizing and Combining Target Variables

Teams consider research indicating which target variables are associated with positive outcomes and linked to specific interventions. The idea of *keystone* variables prioritizes those having relatively *narrow targets for change with the possibility of widespread benefits* to clients (e.g., Barnett, Bauer, Ehrhardt, Lentz, & Stollar, 1997). Common examples include teachers' effective instruction and managerial skills and students' reading fluency, engagement with practice opportunities, social problem solving, compliance with adult requests, and independence with classroom routines through self-management. Selecting a keystone variable as the initial target may result in positive accompanying effects that reduce the need for additional interventions.

In many cases, students exhibit more than one problem behavior, presenting more than one possible target variable. Team members can prioritize targets based on a number of considerations or include more than one target variable. First, teams may consider the severity of problem behaviors. Dangerous and high-risk behavior would be targeted immediately. Behaviors most significantly discrepant from those of peers may be targeted early, providing more time for intervention efforts to have effects.

Sometimes teams may elect to target more than one variable right from the start. For example, a student may demonstrate academic skills deficits in math and reading. Both academic areas are keys to school success and may warrant immediate intervention. In such cases, teams must be careful to ensure that they have the resources necessary to target both variables meaningfully or develop plans in a sequence based on relative risk (e.g., reading, then math). As another example of possible multiple target variables, the relationship between poor academic performance and increased rates of problem social behavior has been well documented (Sugai, Horner, & Gresham, 2002). When a student is referred for academic and behavior problems, options for teams include deciding to intervene with academics and seeing whether social behavior changes without direct intervention, or vice versa, before implementing two distinct intervention plans. To help with this decision, teams would carry out a functional assessment to plan target variables based on hypotheses, confirmed with data, about whether or not a student has the needed skills to perform academic tasks or whether student performance variables need to be targeted (e.g., planning reinforcement).

Selecting Students for Intervention

Should schools select students based on concerning behaviors or performance, and then figure out target variables, measures, and interventions? Or should schools first select key variables and measures related to behavior and performance and educational risk, then screen all students and select students for interventions based on results? Both strategies have merit, and recent developments in screening and decision making now make both within reach. This section describes methods of student selection for intervention services, applying the foundation already discussed in target variable selection.

Schools often select students for intervention based on a concerning behavior or performance as typically determined by teachers or parents or by a student's self-referral. There are advantages to receiving referrals directly from those having the most knowledge about a situation and applying problem-solving steps to identify significant variables and to achieve needed outcomes. At the same time, the process of individual referral has led to great variability in who is selected to receive intervention services and what happens to them. The unfortunate tradition has involved waiting for students to fall behind peers or to fail and then applying cultural, local, or personal ideas about failure and what to do about it, including what to target for change. It is very common in schools for students referred for academic or behavior problems to be tested, classified, and placed in special programs. Inconsistent guidelines about selection, idiosyncratic and indefensible measurement decisions, weak systems-level interventions such as grade re-

tention, group and individual interventions uneven in quality and of often-unknown effectiveness, and the lament "he or she just fell through the cracks" have been commonplace. In the end, the system of individual referrals, diagnostic testing for educational problems, and resulting classification and placement has been widely criticized with respect to systematic and effective special services to students (e.g., Heller, Holtzman, & Messick, 1982). Additionally, many argue that this flawed process has led to the overrepresentation of some minority groups in special education and that strengthening prevention, educational, and behavioral interventions without unnecessary and potentially stigmatizing labels is highly promising (e.g., Hosp & Reschly, 2004; Newell & Kratochwill, 2007; Skiba et al., 2008).

This section includes a discussion of various approaches to selecting students for intervention services, including strengths and weaknesses of teacher nominations, use of curriculum-based norms, and indicators of risk. Decision rules also are needed in cases in which intervention assistance is needed not for an individual student but for the class or even the school, and these decisions are informed by estimates of prevalence or base rates of the targeted variable.

Methods of Selection

Identifying Students in Need of Intervention Using Teacher Nominations

Teachers are significant participants in problem solving, and their observations about student performance are vital to the process because of their frequent and unique contacts with students under natural classroom demands. Teachers generally show a moderate to high level of accuracy in reporting student academics and behavior (e.g., Feinberg & Shapiro, 2003; Gresham, Reschly, & Carey, 1987). However, variations among teachers' goals, expectations, and tolerances for student behavior and academic performance can lead to different reasons for referral across teachers and referral rates. Factors such as the performance or behavior of peers in a class can affect how a teacher perceives an individual student and the likelihood that the student will be referred or not (e.g., VanDerHeyden & Witt, 2005). Also, teachers unknowingly may be interacting with students in ways that exacerbate problem behavior or low performance. Regarding intervention decisions, VanDerHeyden, Witt, and Naquin (2003) showed that teachers' predictions of who will and will not have an adequate response to intervention are not very accurate, but many teachers also may have limited specific intervention experience. Nonetheless, when used in conjunction with direct measures of student performance (e.g., academic performance data, direct observation behavior data), information obtained from teacher observations can help effectively identify students in need of intervention support. To achieve the quality of data needed for accurate student selection, teacher information is supported with data on student performance relative to peers (locally and nationally), such as CBM and independent observations.

Curriculum-Based Approaches to Selection

Introduced earlier, CBM is commonly used to select students for academic intervention programs and to monitor student progress during intervention (Deno et al., 1985–1986). CBM is now used widely for academic screening (e.g., Ardoin et al., 2004; Glover & Albers, 2007). Advantages of CBM for screening include brevity (i.e., 1–5 minutes), repeatability, and sensitivity to student progress. For example, in CBM reading, students read aloud a grade-level passage for 1 minute as the administrator records the words read correctly and incorrectly. CBM is interpreted by using various norms and performance criteria from research, discussed next.

Identifying Students Using National Norms

Historically, comparing student performances with national norms from published norm-referenced tests has guided decisions about student need for intervention and/or special education. National norms provide information about the relative performance of students compared with same-age and same-grade peers. However, national norm groups do not necessarily reflect the educational and social environment of a particular school, classroom, and/or student, and they do not directly indicate the degree of possible risk for academic failure. Furthermore,

the use of national norms may present problems with respect to interpreting the performance of some culturally and linguistically diverse students. National norms must enable meaningful comparisons with school and student demographics and must be useful in setting goals and evaluating progress. National norms are used with other norms, such as school, grade, classroom, or peer norms, and with valid criteria for identifying students at risk, depending on the prevention and intervention purpose.

Today, large-scale norms are available for most CBM measures (available from DIBELS [*dibels.uoregon.edu*] and AIMSweb [*aimsweb.com*]) based on data from schools subscribing to the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) and AIMSweb data systems. Although these norms can provide a broader point of comparison for student performance, schools must take into account differences in student populations and resources, which contribute to significantly different performance. Schools included in the DIBELS and AIMSweb databases may not be representative of national student performance, even though they are geographically diverse. Schools subscribing to these systems are more likely to emphasize reading achievement, including adopting a research-based curriculum and using screening and progress monitoring (Good, Wallin, Simmons, Kame'enui, & Kaminski, 2002).

Identifying Students Using Local Norms

A local norm is a description of a school population's performance on a set of tasks developed to represent students from that particular school or school system (Habedank, 1995). The rationale for developing local norms is that behavior and academic performance are products of the ongoing interactions between students and their specific and unique environments. Local norms can be used to evaluate the performance of schools and classrooms over time when compared with national norms and risk indicators and to establish appropriate short-term goals for low-performing schools, classrooms, or students.

For example, schools may use CBM screening data to determine the average oral reading fluency of students at each grade, or schools may use archival data to determine the average number of office referrals per student (i.e., per month, semester, year). Local norms allow the comparison of an individual student's performance with the performances of peers within the same instructional context. As such, local norms provide a more direct and appropriate point of comparison than national norms for many intervention decisions, including student selection. However, local norms should be interpreted along with valid risk indicators that can reliably estimate the likelihood that a student will be successful or require intervention (e.g., Kame'enui, Good, & Harn, 2005) or that a class or group, and not necessarily an individual student, would be the focus of intervention efforts.

In summary, when selecting students for intervention assistance and when setting achievement goals for schools and individual students, it is important to consider national and local norms linked to valid indicators of educational risk. Local norms can be used to accurately identify struggling students within the context of the specific school setting. In addition, once students are selected for intervention, local norms can set initial performance goals that are attainable, and goals can be gradually increased to reduce risk based on national norms and empirically derived performance criteria associated with school success. Behavioral target variables, measurement methods, and goals likewise are set within a local context (e.g., numbers of students with disruptive behaviors in a classroom or other school context; Bell & Barnett, 1999).

Identifying Students Using Valid Indicators of Educational Risk

Researchers have identified numerous indicators of educational risk that may contribute to a student's school performance. Students also move in and out of risk situations. Thus schools cannot possibly assess all potential indicators of risk. However, by carefully selecting risk indicators with strong empirical support, problem-solving teams can increase the chances that they are correctly identifying many students who will need intervention assistance to achieve school success. Academic failure is preventable to a degree by early screening, with accurate risk appraisal

and effective programming. Although local and national norms can provide valuable information about student performance relative to peers, the relative performance of the student may be less significant than risk estimates.

When selecting students for intervention, data should allow an empirical prediction or likelihood either that the student will be successful with additional supports or that the supports are not needed at that point. The selection of students should be both norm and criterion-referenced, taking into account comparisons between target student and peer performance, as well as comparisons with specific performance levels that are predictive of need for intervention or continued success. A *benchmark* is an empirical method of indicating that a student is on track if the current level of instruction is continued; similarly, levels of risk can be indicated for specific performances on measures (Kame'enui et al., 2005). Risk can also be estimated from repeating CBM measures and determining whether at-risk students are catching up to peers and grade-level benchmarks by noting changes in level and trend (or slope of progress) of performance.

As examples, DIBELS benchmarks are based on research correlating performance on various early-reading measures with later literacy outcomes. The benchmarks provide a criterion from which to evaluate student performance. Unlike screening based on comparisons only with national or local norms, criterion-based screening provides problem-solving teams with empirical estimates of risk levels that can be used for school planning (Kame'enui et al., 2005). Also, the AIMSweb system can help problem-solving teams conduct criterion-based screening by reporting percentile ranks for performance levels on various CBM measures across grade levels. Knowing that there is empirical evidence that performance on the CBM measures is linked to short- and long-term academic outcomes, teams can select students for interventions, set goals, and monitor progress using these data.

Why Base Rates Are Important

Base-rate estimates can help make the most of screening programs by appropriately focusing instruction or intervention efforts, including what variables to target, as well as methods of screening, selection, and program design (Macmann & Barnett, 1999; VanDerHeyden & Witt, 2005). Base rates are estimates of the prevalence of an objectively defined characteristic, such as risk for reading failure; social risk, such as dropping out of school; or a diagnostic category, such as learning disabilities. These specific base rates are estimated for a population or setting, such as a school (Meehl & Rosen, 1955). When deciding which students need intervention services, schools should consider base-rate estimates of the proportion of students expected to demonstrate academic or behavioral difficulties of interest. If base rates are very high or low, screening itself and program decisions need to be altered. For example, based on past graduation rates, two schools estimate the base-rate occurrence of dropping out of high school. School 1 has a base rate of 10%, whereas School 2 has a base rate of 60%. For School 1, with a relatively low base rate for dropout, intervention would focus on individual and small groups of students who are at risk for dropping out. In contrast, based on the high base-rate estimate for dropping out at School 2, planning would emphasize schoolwide prevention programming. In such a case, the focus is not only on individual students but also on the school as a system and on what can be done to effectively screen and better support the student population to increase graduation rates. By considering base rates, teams can evaluate an early screening process to ensure that students who need services are not being overlooked and that students who do not need intervention are not unnecessarily receiving additional support (Glover & Albers, 2007). When classrooms have high or low rates of academic problems, considering base rate helps ensure that appropriate screening methods are used and that interventions and support programs are targeting school needs effectively by addressing target variables and interventions, as appropriate, at class, group, or individual levels (e.g., Newell & Kratochwill, 2007; Skiba et al., 2008; VanDerHeyden & Witt, 2005).

In summary, student selection is based on improving *accuracy* and *usefulness* of targeted variables, measurement methods, and decisions about who needs help and what is helpful to students. In some cases, *schools* or

classrooms may be selected for intervention if performance of many students is alarming, such as high rates of school failure or discipline referrals.

The RTI Context

RTI (response to intervention) changes the landscape of target variable selection due to its purposes and methods. At present, RTI is an option identified in Federal law (Individuals with Disabilities Education Improvement Act, 2004) for local educational agencies to help identify students with specific learning disabilities (SLD), but RTI's possible impact is much broader (e.g., Batsche et al., 2005). In contrast to starting with a student referral and figuring out target variables and next steps, the defining quality of RTI is an approach to decision making using universal early screening and outcomes of empirically defensible prevention programs and sequences of interventions as the database for service delivery determination. Selecting target variables and students for intervention is based on objective criteria derived from research. Concepts and measures of risk (e.g., poor reading fluency, challenging behaviors) are supported with data indicating that targets can be influenced by environmental changes *and* have evidence of positive outcomes for children. Research-based prevention programs and interventions are used to judge needed program qualities in schools. Thus target variable selection, as well as the selection of students for intervention services, starts with the premise of effective schools, research-based constructs of risk, and research on what works.

First, schools using RTI screen all students and offer appropriate services without delays. That is, schools assess the performance of all students on systemwide, high-priority target variables and assign students identified as "at risk" to valid instructional programs or interventions. This is in contrast to the approach taken by many schools in the past, in which target variables were identified idiosyncratically by the person making the referral and, more specifically, after a student had been referred for assistance. Second, a student's RTI intervention progress through established and research-based tiers of services may be used as evaluation data for more specialized service decisions or for decisions to fade intervention assistance when no longer necessary. These intervention data would be used instead of diagnostic test results collected at one point in time and questionably related to interventions. The dataset is different and would include detailed information on the research-based interventions implemented, reliable and valid data on the student's response to interventions, and evidence that interventions were carried out carefully. The result of a tiered intervention progression is a valid data-based description of targeted variables and needed interventions based on prior outcomes that can be used for planning next steps, as necessary, at all levels (school, class, group, and individual).

RTI is evolving, but generally, the first tier of RTI models is intended to be universal, school- and classwide, influencing the greatest number of children through prevention, sound curriculum, and evidence-based instructional and classroom managerial practices. Guidelines suggest that effective schoolwide supports should meet the needs of 80–90% of students in a given student population, with 10–20% of students requiring additional support (e.g., Kame'enui et al., 2005; Sugai et al., 2002). Students requiring additional support are served in a second tier consisting of short-term empirically based *selected* or *targeted* interventions (e.g., Batsche et al., 2005). More common examples of Tier 2 programs are based on standard protocols for valid instructional interventions that increase practice opportunities in small groups (i.e., reading skills) based on curriculum, data, and decision rules from Tier 1 (e.g., Vaughn, Wanzek, Woodruff, & Linan-Thompson, 2007; see also Linan-Thompson & Vaughn, Chapter 16, this volume). For the approximately 1–5% of students who are not sufficiently helped by the first two tiers, Tier 3 includes more intensive individualized services, or services delivered to smaller groups of students, and a focus on increased practice of specific skills related to the Tier 1 curriculum. Figure 2.2 shows a typical tiered model.

Data and team decisions would demonstrate need for intervention changes that increase or decrease in intervention intensity (e.g., time, specialized resources). Student performance ideally would be tracked

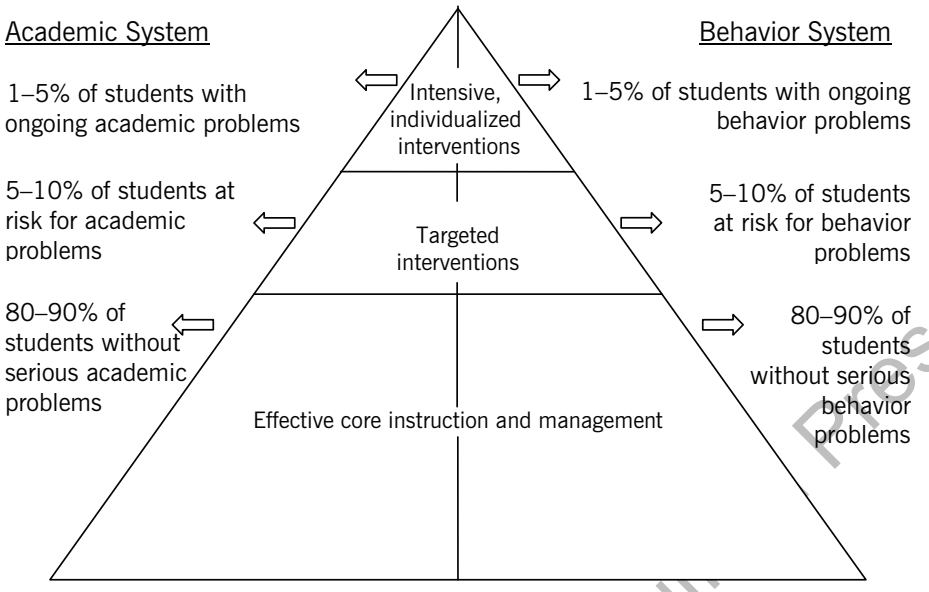

FIGURE 2.2. Typical tiered model.

or monitored in ways that are as close as schools can come to high-quality intervention research *within each tier* in order to expect results similar to the research from which it was derived. Multiple goals may be set that include immediate as well as long-term objectives, such as outcomes for success in typical environments and, ultimately, successfully maintained performance of targeted variables and generalization of responses to new situations. Teaming and problem solving are used to increase the chances of success for RTI by addressing planning and the logistics of intervention implementation (e.g., how often, where, who will implement, schedule for reliability and intervention adherence checks), as well as outcomes at the levels of both school and individual student. Problem solving also is used to help with needed instructional variations and unique student situations.

Similarly, the framework of positive behavior support (PBS) conceptualizes prevention and intervention efforts for social behaviors, calling for high-quality programming built on empirically validated interventions and tiered services (Sugai et al., 2002). First, an effective schoolwide system is developed and implemented. In Tier 1, classrooms are well designed and managed, behavioral expectations are directly taught, and all students receive competent instruction on key social skills. In Tier 2, more practice is given based on a valid curriculum or intervention program in a group or embedded format, as are interventions with specific, troublesome classroom activities such as transitions. Tier 3 is based on intensified and individualized plans related to social behaviors. RTI and PBS are integrated in some RTI models and have many commonalities (Batsche et al., 2005).

RTI and Target Variable Selection

Universal Screening

The success of RTI is dependent on the early and accurate identification of students at risk (Compton, Fuchs, Fuchs, & Bryant, 2006). Universal screening defines the initial variables that will be used for intervention decisions in RTI. Variables may include instruction in and mastering of key early literacy skills, rate and level of improvement in skills, and variables related to instructing and supporting social competence in schools.

As an example, teams conduct universal screening for academic performance three times a year using norm- and/or criterion-referenced (derived locally and/or nationally) cut points to identify students who

need intervention (Ardoin & Christ, 2008; Good, Gruba, & Kaminski, 2002). Typically, a CBM probe or a median score from three CBM probes administered at one time in key areas (e.g., reading and math) is used, and students scoring below the cut point are considered for Tier 2 intervention (e.g., Ardoin et al., 2004). Other RTI models include universal screening only at the start of the school year to identify a group of students who show potential at-risk indicators. Students in this group are then closely monitored to determine need for preventive intervention (Fuchs & Fuchs, 2007). To determine response to Tier 1 instruction and need for Tier 2 intervention, Fuchs and Fuchs (2007) recommend using a "dual discrepancy" criterion based on both (1) student growth defined as differences in performance over time that show up as a *slope* on a graph (i.e., words read per minute plotted every week showing changes in reading fluency over time); and (2) the level of performance (e.g., mean level of a target student compared with peers or other norms). In this model, a student is selected for Tier 2 intervention if, after at least 5 weeks of progress monitoring in the general curriculum, his or her slope of improvement and final level of performance are both at least 1 standard deviation below those of peers. As a different example, VanDerHeyden, Witt, and Gilbertson (2007) use class CBM data to decide various next steps that may include a brief (10-minute) classwide academic intervention carried out for 10 days to help with screening decisions if the class performance is low. As the preceding examples suggest, there is not one RTI model at present, but there is a strong consensus on early universal screening for key instructional, curricular, and social variables and use of the measures of these variables for monitoring progress.

Tiered Variables

Although there are variations in RTI models, target variables and students are selected and analyzed by teams using specific procedures described by the RTI model for each of the aforementioned tiers. As introduced, RTI and PBS are characterized by structural components, or tiers, that organize school, classwide, group, and individual target variables and sequential decision points that are analyzed by looking at student outcomes (Gresham, 2007; Sugai et al., 2002). Thus school teams analyze school, class, group, and individual contexts based on research on improving reading and social behaviors and reducing challenging behaviors. Table 2.3 shows RTI tiers and common target variables that may be used by schools or teams at each tier. Within RTI, in addition to targeted variables related to direct measures of students' academic skills and behavior, instructional and classroom variables (curriculum, adherence to the curriculum, qualities and prevalence of instruction, classroom management, discipline programs, etc.) may be targets of intervention. Tier 1 variables include those related to scientifically based instruction for academic skills and social behaviors (Kame'enui et al., 2005; Sugai et al., 2002; Vaughn et al., 2007). Similar to Tier 1, selection of Tier 2 target variables would yield measures related to the specific academic or social concern. Furthermore, accompanying the increasing intensity of interventions from Tier 1 to Tier 2 would be an increase in the intensity of progress monitoring (e.g., more frequent assessment of target variables, more refined measures of academic progress in early literacy skills or social behavior, more frequent reliability and intervention adherence checks). At Tier 3, the collection of data on target variables would intensify further as teams try to use resources efficiently while still promoting positive student outcomes. Other important Tier 2 and 3 variables are included in plans for generalization and maintenance of skills and performance of skills in typical educational settings (Tier 1). For example, when implementing a reading intervention, teams may assess for generalization by monitoring reading fluency on both practiced and unpracticed reading passages and may plan to improve generalization in other instructional contexts, such as math word problems. Social behavior targets taught in Tier 2 groups would be progress-monitored in classroom and other school settings. Across all tiers, data on the degree to which RTI procedures are implemented as intended and that show the quality of outcomes are required; these are characteristics of services not typically applied to traditional referral decisions.

Problem solving related to functional and testable hypotheses about student academic

TABLE 2.3. Examples of Target Variables and Measures across Tiers

	Academic		Behavior	
Tier	Variables	Measures	Target variables	Measures
1	• Reading • Math • Content-area achievement • Instruction	• DIBELS triannual benchmark data • Math and reading triannual CBM benchmark data • Achievement test scores • Opportunities to respond to academic stimuli • Adherence to the curriculum	• Disruptive behavior • Instruction	• Office discipline referrals • Teacher referrals • Opportunities to practice appropriate social behavior • Adherence to classroom management procedures
2	• Reading • Math • Content-area achievement • Instruction	• Weekly DIBELS, math and reading CBM progress-monitoring data • Homework and classwork completion & accuracy • Test scores • Opportunities to respond to academic stimuli	• Disruptive behavior • Engagement • Compliance • Peer interactions • Instruction	• Weekly direct observations of engagement, compliance, peer interactions • Weekly teacher report of behavior • Opportunities to practice appropriate social behavior
3	• Reading • Math • Content-area achievement • Instruction	• Twice weekly DIBELS, reading CBM progress-monitoring data • Homework and classwork completion and accuracy • Test scores • Opportunities to respond to academic stimuli	• Disruptive behavior • Engagement • Compliance • Peer interactions • Instruction	• Twice weekly direct observations of engagement, compliance, peer interactions • Daily teacher report of behavior • Opportunities to practice appropriate social behavior

and social learning and performance, which may be useful at all RTI tiers, are critical at Tier 3 (individualized and intensive), with more challenging and complex academic and social problem behavior. In other words, beyond increasing specific practice, when previous intervention attempts have failed and the environmental variables contributing to the problem behavior are unclear, problem solving and FBA methods described in the chapter may provide information that can lead to the efficient identification of effective interventions.

Conclusions: Achieving Confidence in Decisions

Target "behavior" selection is in keeping with traditional discussions of problem solving for students with concerning behavior. However, with the number of challenges faced by schools and in line with academic and social interventions that are based on systematic changes in instruction and environment, we use the broader term of *target variable selection*. Target variable selection, measurement, and schedules of measurement create the data for intervention decisions. Students may come to the attention of professionals because of concerning behaviors or performance, after which target variables are selected; and target variables may be selected in advance by schools, with measures then used for screening and decision making. Target variables organized by RTI enable progress monitoring at various levels to address specific questions: at the school (what's working, what isn't); classroom (more or less teacher support, quality of instruction); and for students (change intervention or tier, quality of needed interventions and supports). Technical adequacy was stressed as a way to improve the validity and reliability of the decision-making process. The basic reason for technical adequacy is to get a "handle" on the overall confidence that teams can have in a complex process of decision making.

References

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum-based measurement slope estimates using data from tri-annual universal screenings. *School Psychology Review*, *37*, 109–125.

Ardoin, S. P., Witt, J. C., Suldo, S. M., Connell, J. E., Koenig, J. L., Resetar, J. L., et al. (2004). Examining the incremental benefits of administering a maze and three versus one curriculum-based measurement reading probes when conducting universal screening. *School Psychology Review*, *33*, 218–233.

Barnett, D. W., Bauer, A., Ehrhardt, K., Lentz, E., & Stollar, S. (1997). Keystone targets for change: Planning for widespread positive consequences. *School Psychology Quarterly*, *11*, 95–117.

Barnett, D. W., Hawkins, R. O., Prasse, D., Macmann, G., Graden, J. L., Nantais, M., et al. (2007). Decision making validity in response to intervention. In S. Jimerson, M. Burns, & A. VanDerHeyden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 106–116). New York: Springer Science.

Batsche, G., Elliott, J., Graden, J. L., 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.

Bell, S. H., & Barnett, D. W. (1999). Peer micronorms in the assessment of young children: Methodological reviews and examples. *Topics in Early Childhood Special Education*, *19*, 112–122.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, *98*, 394–409.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson.

Daly, E. J., Martens, B. K., Barnett, D. W., Witt, J. C., & Olson, S. C. (2007). Varying intervention delivery in response to intervention: Confronting and resolving challenges with measurement, instruction, and intensity. *School Psychology Review*, *36*, 562–581.

Deno, S. L., Marston, D., & Tindal, G. (1985–1986). Direct and frequent curriculum-based measurement: An alternative for educational decision making. *Special Services in the Schools*, *2*, 5–27.

Feinberg, A. B., & Shapiro, E. S. (2003). Accuracy of teacher judgments in predicting oral reading fluency. *School Psychology Quarterly*, *18*, 52–65.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*, 199–208.

Fuchs, L. S., & Fuchs, D. (2007). A model for implementing responsiveness to intervention. *Teaching Exceptional Children*, *39*, 14–20.

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology*, *45*, 117–135.

Good, R. H., Gruba, J., & Kaminski, R. A. (2002). Best practices in using dynamic indicators of basic early literacy skills (DIBELS) in an outcomes-driven model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (4th ed., pp. 699–720). Bethesda, MD: National Association of School Psychologists.

Good, R. H., Wallin, J. U., Simmons, D. C., Kame'enui, E. J., & Kaminski, R. A. (2002). *System-wide percentile ranks for DIBELS benchmark assessment* (Technical Report No. 9). Eugene, OR: University of Oregon.

Greenwood, C. R. (1991). A longitudinal analysis of time, engagement, and achievement in at-risk versus non-risk students. *Exceptional Children*, *57*, 521–535.

Gresham, F. M. (2007). Evolution of the response-to-intervention concept: Empirical foundations and recent developments. In S. Jimerson, M. Burns, & A. VanDerHeyden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 10–24). New York: Springer Science.

Gresham, F. M., Reschly, D. J., & Carey, M. (1987). Teachers as "tests": Classification accuracy and concurrent validation in the identification of learning disabled children. *School Psychology Review*, *16*, 543–563.

Gresham, F. M., Watson, T. S., & Skinner, C. H. (2001). Functional behavioral assessment: Principles, procedures, and future directions. *School Psychology Review*, *30*, 156–172.

Habedank, L. (1995). Best practices in developing local norms for problem solving in the schools. In A. Thomas & J. Grimes (Eds.), *Best practice in school psychology (3rd ed., pp. 701–715). Washington, DC: National Association of School Psychologists.*

Hawkins, R. P. (1986). Selection of target behaviors. In R. O. Nelson & S.C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 331–385). New York: Guilford Press.

Hawkins, R. P., & Dobes, R. W. (1975). Behavioral definitions in applied behavior analysis: Explicit or implicit. In B. C. Etzel, J. M. LeBlanc, & D. M. Baer (Eds.), *New developments*

in behavioral research: Theory, methods, and applications. In honor of Sidney W. Bijou (pp. 167–188). Hillsdale, NJ: Erlbaum.

Heller, K. A., Holtzman, W. H., & Messick, S. (Eds.). (1982). *Placing children in special education: A strategy for equity.* Washington, DC: National Academy Press.

Horner, R. (1994). Functional assessment contributions and future directions. *Journal of Applied Behavior Analysis*, *27*, 401–404.

Hosp, J. L., & Reschly, D. J. (2004). Disproportionate representation of minority students in special education: Academic, demographic, and economic predictors. *Exceptional Children*, *70*, 185–199.

Kame'enui, E. J., Good, R. G., III, & Harn, B. A. (2005). Beginning reading failure and the quantification of risk: Reading behavior as the supreme index. In W. L. Heward, T. E. Heron, N. E. Neef, S. M. Peterson, D. M. Sainato, G. Cartledge, et al. (Eds.), *Focus on behavioral analysis in education: Achievements, challenges, and opportunities* (pp. 68–88). Upper Saddle River, NJ: Pearson.

Kanfer, F. H. (1985). Target selection for clinical change programs. *Behavioral Assessment*, *7*, 7–20.

Kazdin, A. E. (1985). Selection of target behaviors: The relationship of treatment focus to clinical dysfunction. *Behavioral Assessment*, *7*, 33–47.

Kennedy, C. H. (2005). *Single-case designs for educational research.* Boston: Allyn & Bacon.

Kratochwill, T. R. (1985). Selection of target behaviors in behavioral consultation. *Behavioral Assessment*, *7*, 49–61.

Kratochwill, T. R., & Bergin, J. R. (1990). *Behavioral consultation in applied settings: An individual guide.* New York: Plenum.

Macmann, G. M., & Barnett, D. W. (1999). Diagnostic decision making in school psychology: Understanding and coping with uncertainty. In C. R. Reynolds & T. B. Gutkin (Eds.), *Handbook of school psychology* (3rd ed., pp. 519–548). New York: Wiley.

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.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, *52*, 194–215.

Newell, M., & Kratochwill, T. R. (2007). The integration of response to intervention and critical race theory-disability studies: A robust approach to reducing racial discrimination in evaluation decisions. In S. Jimerson, M. Burns, & A. VanDerHeyden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 65–79). New York: Springer Science.

Shapiro, E. S. (2004). *Academic skills problems: Direct assessment and intervention* (3rd ed.). New York: Guilford Press.

Shinn, M. R., Walker, H. M., & Stoner, G. (Eds.). (2002). *Interventions for academic and behavior problems: II. Preventive and remedial approaches.* Washington DC: National Association of School Psychologists.

Skiba, R. L., Simmons, A. B., Ritter, S., Gibb, A. C., Rausch, M. K., Cuadrado, J., et al. (2008). Achieving equity in special education: History, status, and current challenges. *Exceptional Children*, *74*, 264–288.

Sugai, G., Horner, R. H., & Gresham, F. M. (2002). Behaviorally effective school environments. In M. Shinn, H. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems: II. Preventive and remedial approaches* (pp. 315–350). Bethesda, MD: National Association of School Psychologists.

VanDerHeyden, A. M., & Witt, J. C. (2005). Quantifying context in assessment: Capturing the effect of base rates on teacher referral and a problem-solving model for identification. *School Psychology Review*, *34*(2), 161–183.

VanDerHeyden, A. M., Witt, J. C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a response to intervention (RTI) model on identification of children for special education. *Journal of School Psychology*, *45*, 225–256.

VanDerHeyden, A. M., Witt, J. C., & Naquin, G. (2003). Development and validation of a process for screening referrals to special education. *School Psychology Review*, *32*, 204–227.

Vaughn, S., Wanzek, J., Woodruff, A. L., & Linan-Thompson, S. (2007). Prevention and early identification of students with reading disabilities. In D. Haagar, J. Klingner, & S. Vaughn (Eds.), *Evidence-based reading practices for response to intervention* (pp. 11–27). Baltimore: Brookes.

Watson, T. S., & Steege, M. W. (2003) *Conducting school-based functional behavior assessments: A practitioner's guide.* New York: Guilford Press.

Wilson, F. E., & Evans, I. A. (1983). The reliability of target-behavior selection in behavioral assessment. *Behavioral Assessment*, *5*, 15–32.

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.