# CHAPTER 1

# Introduction to Design and Analysis for Educational Intervention



A recent documentary television program showed two men trying to attract sharks, from the safety of a shark cage and a boat, using different kinds of sounds. They tried several sounds for short periods of time, each with no luck. The disappointing experiment was about to end when all of a sudden a large great white shark appeared. The two would-be scientists concluded that the sound being projected at that time was the exact auditory stimulus that would attract sharks. Many people would probably, and correctly, interpret this "experimentation" as completely lacking validity, yet this approach to empirically answering questions happens in K–12 schools every day. Sometimes this unscientific decision-making model is acceptable, but often examining data collected in a manner from which causality cannot be inferred can have disastrous results.

While it would be ideal if general education instruction and curricula would meet the needs of all children, this is all too often not the case. A significant percentage of students have some academic or social behavior difficulties at school due to reasons that range from various environmental issues to student disabilities. Regardless of the cause, it is the school's obligation to attempt to intervene in order to solve the problem when students do not progress at the same rate as their peers or are behaving in an inappropriate manner, even for students with disabilities. Thus, we engage in individual and small-group interventions to try to solve academic and social behavior problems for personal, professional, and ultimately, legal reasons.

## WHAT IS INTERVENTION?

Before we move on, it is critical that we define the model of "intervention" that we use throughout this book. Without doing so, we risk readers applying their own version of educational intervention as they read and reflect. An intervention is "a planned modification of

the environment made for the purpose of altering behavior in a prespecified way" (Tilly & Flugum, 1995). An intervention can be at the whole-school (e.g., curriculum) or individual level.

This book clearly adheres to the empirical foundation of interventions with the effectiveness being best documented by observing the outcome data. In our opinion, outcome data gathered in a scientific manner drive the intervention process, which suggests that educators are not only effective in selecting appropriate and scientifically based interventions but also in collecting academic and social behavior outcome data.

There are a number of excellent resources for intervention selection (e.g., Burns, Riley-Tillman, & Rathvon, 2017; Shapiro, 2011; *ebi.missouri.edu*; *www.interventioncentral.com*), but the data analytic aspect of an effective intervention has consistently been the focus of intervention resources. This creates a critical problem because although selecting research-based interventions increases the likelihood of a positive effect, it does not ensure it. Further, while collecting defensible outcome data for interventions allow educational professionals to judge the effectiveness of the intervention, those data must be collected in a systematic manner (experimental) and analyzed appropriately. In other words, without effective intervention methods and analysis, the final step of problem solving cannot be taken. This book attempts to fill that critical void by focusing on the design and analysis of data to evaluate the effectiveness of interventions. Below we discuss systematic school-based design and analysis to evaluate intervention effects.



## SCHOOL-BASED INTERVENTION AND ANALYSIS: MULTI-TIERED SYSTEMS OF SUPPORT

Multi-tiered systems of support (MTSS) is a problem-solving model that uses student outcomes to provide the most effective instruction and intervention to each student through the efficient allocation of educational resources. Readers are likely familiar with MTSS and other terms related to tiered systems of intervention, such as response to intervention (RTI) and positive behavioral interventions and supports (PBIS). Both RTI and PBIS are types of an MTSS, with RTI focusing on academic problems and PBIS on behavioral difficulties. Both RTI and PBIS use a three-tiered model (Tier 1: universal; Tier 2: targeted; and Tier 3: intensive), increasing levels of intervention, and provision of related educational services to each student based on need. Through this process, special education becomes the funding source for children who need the most significant levels of support (Burns & Gibbons, 2012) through an RTI framework, but resource-allocation decisions are also made when starting a Tier 2 intervention or when intensifying an intervention so that it is a Tier 3 rather than a Tier 2 intervention.

### *MTSS and Intervention*

Educators in settings that implement an MTSS model clearly place a significant value on intervention and data-based decision making. As such, it is critical to consider the varying nature of interventions in education. When one considers intervention from a whole-school perspective, the importance of intervention methodology is clearly not consistent through-

out each level of service delivery. Intervention begins with curriculum, and then becomes more and more individualized as it proves to be ineffective. At each stage, the amount of resources allocated to the intervention increases and the implications of intervention failure rise. Figure 1.1 presents the three-tiered model of MTSS with intervention, assessment, and services typically associated with each level (Burns, Jimerson, VanDerHeyden, & Deno, 2016).

In Tier 1, the standard intervention practices, or curriculum, are utilized in a somewhat uniform manner across all children in the school. While schools tend to spend a good deal of time and energy on the selection of curriculum, this selection process is not tailored to one individual student—rather, it is designed to encompass the general education population. The assessment of the effectiveness of this curriculum is also not individualized for each student—rather, standard grading methods and periodic standardized assessment procedures are utilized to judge how individual students are progressing. At this level, essentially all we know is that either a student is progressing acceptably, or is not. We can never be sure whether the observed progress or lack thereof is directly related to the general education program. Realistically, a student could be doing quite well due to other supports (e.g., tutoring, home support, superior abilities) that are unknown to the school. Further, a student could be failing due to reasons that are not directly related to the primary intervention.

In the case of a failing student, he/she is moved to Tier 2 so that schools can accomplish two goals. The first goal is obviously to attempt more rigorous interventions to help the child in relation to his/her academic or social behavior difficulties. The second goal is to administer the intervention in a more systematic manner in terms of selection, assessment, integrity, and intervention methodology so that we can better understand a child's academic or social behavior problem more thoroughly. In Tier 2, the first goal takes precedence. At



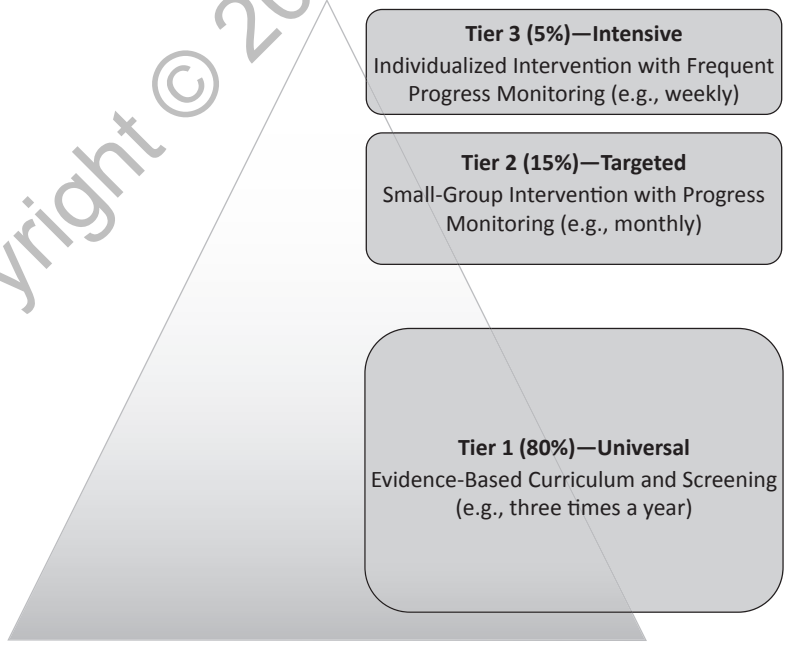


**FIGURE 1.1.** A continuum of assessment and intervention activities.

this stage, daily small-group interventions with weekly or biweekly progress monitoring are often the norm (Burns, Hall-Lande, Lyman, Rogers, & Tan, 2006). Tier 2 should service approximately 15% of the total school population.

Students who do not respond adequately to intervention in the first two tiers are moved into Tier 3. In Tier 3, thoroughly understanding a child's academic or social behavior problem becomes increasingly more important as assessment and intervention efforts attempt to identify the causal variables or functional relationship between the instructional setting and student learning. At this stage, a fully problem-solving approach is essential for the development of intervention. Children for whom interventions within Tier 3 are not effective, or for whom the successful intervention is so intense that it exceeds the resources associated with general education, are considered for special education eligibility in order to provide a free and appropriate public education (FAPE).

## *Special Education as a Foundation for MTSS*

Because the goal of MTSS is to identify an intervention that allows a child to be successful, special education is primarily seen as the funding avenue for children whose needs are so intense that they exceed the resources typically allocated to general education. Thus, the primary jurisdiction over MTSS functionally falls within general education, but the origins lie in special education.

The definition of special education has never changed—it is and always will be specially designed instruction, at no cost to the parents, to meet the unique needs of a child with a disability. However, there have been two difficulties in practice: how children with disabilities are identified, and how well the instruction is specially designed. We discuss both below.

### *Identifying a Child with a Disability*

We have long understood that children who are referred for special education eligibility assessments are likely to be placed into special education, often due to variables other than educational need or consistency with eligibility criteria (Maki, Burns, & Sullivan, 2017; Potter, Ysseldyke, Regan, & Algozzine, 1983; Ysseldyke & Thurlow, 1984). As a result, special education policy recommendations have focused on using MTSS to enhance eligibility decisions (Donovan & Cross, 2001; Gresham, 2002b), which culminated with the 2004 provision in federal special education law that a local education agency "may use a process that determines if the child responds to scientific, research-based intervention as a part of the evaluation procedures" (Public Law 108–446 § 614 [b][6][A]; § 614 [b][2 & 3]). Assessing the extent to which a student responds to research-based interventions is both the hallmark of MTSS and of experimental decision making. The effect on identifying disabilities is still a matter of debate and future research.

### *Specially Designed Instruction*

Special education has a history of questionable practices in developing specially designed instruction. Previously used intervention heuristics emphasized intuitive appeal over

empirical data. Interventions for children with disabilities in the 1960s, 1970s, and early 1980s were frequently developed by matching instructional modality with the student's preferred learning style (i.e., auditory, kinesthetic, or auditory learner) or by identifying areas of psycholinguistic difficulty (Kavale, 2001). However, meta-analytic research found such small effects for these interventions that relying on them was presented as an example of philosophical and clinical beliefs overshadowing research data (Kavale & Forness, 2000).

Over the past 25 years, research has consistently supported that interventions have a greater likelihood for success if based on formative data, implemented with integrity, and evaluated with empirical data (Codding, Hilt-Panahon, Panahon, & Benson, 2009; Daly, Martens, Dool, & Hintze, 1998; Martens & Gertz, 2009; Szadokierski, Burns, & McComas, 2017). This movement toward empiricism in intervention represents a more scientific approach to special education that relies on experimental rather than correlational data (Reschly & Ysseldyke, 2002). The emphasis on student outcomes is also apparent in efforts to implement and study MTSS. Instructional decisions based on student learning data inexorably lead to techniques with a strong research base and away from popular but less effective approaches, such as learning styles.

## *Problem Analysis as a Foundation for MTSS*

Deno and Mirkin (1977) proposed that specially designed instruction can happen only through data-based program modification, or using outcome data to evaluate program effects and make appropriate modifications as needed. Thus, effective special education programming is based on analysis of outcome data, which is also the basis of MTSS.

In an MTSS, as student needs become more intense, measurement becomes more precise and frequent, and problem analysis becomes more in-depth. Problem analysis involves identifying the important elements of a problem by examining possible causes, seeking additional information, and framing possible solutions (Thomson, 1993). In Tier 1, problem analysis simply involves identifying whether a problem exists. Often the data gathered in Tier 1 can be similar to the shark seekers discussed earlier. If the sound attracts the shark (i.e., students are learning), then there is no need to consider the situation any further because it does not matter *why* students are being successful at this point. Thus, the measurement can be simplistic and the assumption of causality can be minimal.

Students for whom the general curriculum and instruction are not effective require a slightly more in-depth analysis. Remembering that Tier 2 essentially involves small-group interventions, the primary analysis task is to identify a category of deficit and then monitor the effectiveness of the interventions to address that deficit. Monitoring the progress requires sound data collection techniques, but a lack of sufficient progress heightens the need for psychometric adequacy and experimental rigor. As students' needs become more intense, the need for experimental control and assumptions of causality become more prominent. Students for whom the Tier 2 intervention was not successful are then provided with individualized intervention, which requires that the primary problem-analysis focus in Tier 3 is to identify the causal variables of the deficit based on idiosyncratic behavior and idiographic principles.

## THE IMPLICATIONS OF AN MTSS MODEL ON INTERVENTION METHODOLOGY

The role of academic and social behavior intervention, which has always been significant, is magnified in an MTSS environment. In this setting, the success or failure of an intervention (i.e., the student's response to an evidence-based intervention [EBI]) is a critical diagnostic tool as long as the intervention is (1) functionally relevant, (2) evidence based, and (3) implemented with integrity (Riley-Tillman, Burns, & Gibbons, 2013). These three criteria highlight the many critical issues educational professionals must deal with when implementing an MTSS model. First, the intervention must be appropriate for the problem. If an intervention is not functionally relevant, then a lack of response is predictable rather than a signal of a more serious problem. Once a function has been proposed, it is critical that the intervention is considered "evidence based," or supported by research as an effective remedy to a particular referral concern. As with the use of functionally irrelevant interventions, using interventions that have no empirical support is not considered evidence of a larger problem but rather preliminary evidence that the intervention is not effective. Finally, the intervention must be implemented with integrity. If the intervention is altered or degraded, then that is the most likely reason for the lack of success.

Fortunately, each of the aforementioned criteria (functionally related: Beavers, Iwata, & Lerman, 2013; Szadokierski et al., 2017; evidence based: Burns & Ysseldyke, 2009; Stormont, Reinke, & Herman, 2011; integrity: Fryling, Wallace, & Yassine, 2012; Sanetti, Gritter, & Dobey, 2011) has received a good deal of attention. Regrettably, these elements are necessary but not sufficient for an MTSS process to work effectively. Even if a functionally relevant EBI is conducted with integrity, we still may not simply conclude that the child was responding (or not responding) to that particular intervention as opposed to other extraneous factors. This statement requires the use of fully experimental methods.

Burns and Gibbons (2012) suggest that MTSS is a series of problem-solving decisions in which measurement becomes more frequent and precise, and problem analysis becomes more exact as the student's needs become more severe. Thus, as the decisions being made with the intervention outcome data become more serious, it is essential to start considering the issue of causality. Specifically, educational professionals must document both whether academic or social behavior changes have occurred, and whether the intervention in question was responsible for those substantiated changes. The issue of causality becomes most relevant when academic or social behavior problems are dramatic or when the results of the intervention will be used for special education eligibility decision making. At this step it is critical that one can attest that it was the intervention that resulted in a change or that the intervention truly did not show an effect.

Some have suggested that a documented "lack" of an intervention effect triggers special education services in an MTSS model (Fuchs, Mock, Morgan, & Young, 2003; Speece, Case, & Molloy, 2003). While this is logical from an educational perspective, it is rather unique experimentally (as researchers tend to look for an effect rather than a lack of an effect). To be able to take this next step, we need to employ a defensible intervention methodology within the framework of a single-case research design.

## SINGLE-CASE DESIGN AND ANALYSIS

Single-case designs (SCDs) are a class of experimental methodology that has been utilized for decades in a number of disciplines, including psychology and education (Kazdin, 1982). While known by a number of different names, such as *single subject*, *intrasubject*, and *N = 1*, we use the term *single case* throughout this book. Before we move on, it is important to note that a number of the names used for this class of experimental design imply defining features that are not always accurate. For example, the terms *single subject* and *N = 1*, which suggest that this class of methodology is to be used with only one subject, are incorrect, as SCDs are often used with small groups. Another term used for this class of design, *intersubject replication*, can also be a bit misleading for two reasons. First, as Kazdin noted in his classic SCD book, some designs actually depend on comparing intervention effects across subjects. Second, while replication is a key figure in the classic research application of these designs, in an applied educational setting we do not always have a replication phase. It is our belief that the most common feature of SCD is the very philosophy on which it is based—baseline logic—which we further explore in Chapter 2.

Perhaps the best manner of introducing the concept of SCD is to consider what this class of design was developed to accomplish (Riley-Tillman & Walcott, 2007). SCD is a collection of experimental methods that are developed to essentially document three things:

1. If there is an observable and important change in some dependent variable.
2. If the observed change in the outcome data postapplication of the independent variable is a result of the application of the independent variable.
3. If this change is something that is generalizable across time, setting, and target.

When applied to an educational environment in general and educational intervention specifically, the three general purposes of SCD read as follows:

1. Did the outcome variable (e.g., percent of time on task) change when the intervention was implemented?
2. Was the observed change due to the implementation of the intervention and *only* the implementation of the intervention?
3. Can the information learned from this educational intervention be generalized to other similar educational problems and settings?

Although each of these three purposes are discussed further in Chapter 2, at this point it is important to know that SCDs allow for educational professionals to defensibly document the effect of interventions, assess the role of the intervention in the observed change, and decide whether that information has some general programming utility. Obviously, documenting intervention effectiveness is of critical importance in today's educational environment. Gone are the days where it was acceptable for an educational professional to simply state that some intervention was or was not effective. In the accountability era, such statements, particularly if they result in educational planning, must be backed up with data.

The second and third goals of SCD are a bit more advanced, but have become critical in modern education. As noted above, when the success or failure of an intervention is used as evidence for special education eligibility for a struggling student, clearly we should be sure that it was indeed the intervention that created the effect rather than some random external factor. To be sure of the relationship between the intervention and the outcome data, experimental control is critical. Experimental control is what SCD is built to accomplish, as discussed in Chapter 2. While this idea of documenting causality may seem novel, to avoid this issue with high-stakes cases is hardly defensible.

The final goal of SCD is to collect evidence regarding the generalization of some intervention finding. In other words, can this same intervention be used in the future with the same child, or can this intervention be used with another child who has a similar issue? As shown in Chapter 2, the only way to gather evidence to support that a current intervention effect will work in the future or with another child is systematic replication. SCD gives educational professionals a path to conduct systematic replication, and thus a path to defensible claims as to the generalizability of intervention results.

## WHY IS SCD IDEAL FOR EDUCATIONAL PRACTICE?

Perhaps the most attractive element of SCD is that it is a family of experimental methods that are designed for use with one student (or a small group of students). If you pick up a journal that publishes educational research, you will undoubtedly find that most studies compare groups of subjects in some manner. This type of experimental design is focused on collecting data from a large number of subjects that are considered to be a good representation of some larger population. This form of experimentation is effective for the development of generalizable knowledge. Research using one child is typically not as suited for this purpose. For example, it would be risky to say that outcome data on one child with some specific reading problem would be predictive for other children with the same reading problem (Kazdin 1982; Kratchowill & Williams, 1988; Walcott & Riley-Tillman, 2007). The one child being observed could be unique in a number of ways that might make him/her an atypical representative for the larger population of children with that specific reading problem. On the other hand, if you were to collect data from say, 150 children with that specific reading problem, their outcome data would start to describe the larger population (all children with that specific reading problem), but might miss important idiosyncrasies for the individual student. Moreover, while a group orientation is logical if you are attempting to produce knowledge about a population, most educational professionals work at the level of one child or a small group of children.

Through systematic replication of observed intervention effects (across setting, situations, and interventionists) we gain confidence that the intervention will be effective in future applications if used with targets similar to those represented in the large-group efficacy studies. While this is clearly valuable information, it does not mean that we can confidently assume that an EBI will always have a positive effect on an individual child who is a member of that group. To illustrate why this is the case, consider an intervention that is consistently effective (e.g., showed a positive intervention effect with 90% of children stud-

ied). Even with such a consistently effective intervention, it was ineffective in 10% of the cases. A practitioner can never truly know before implementation if he/she is working with a child who falls in that 10% and thus the intervention would likely prove ineffective. In the end, the label of EBI does not imply that an intervention will work with a similar child, but rather that it is probable that it will be effective (assuming it was selected appropriately). The implications of this reality are critical. Even a well-selected EBI that is delivered with perfect treatment integrity may prove ineffective for an individual child. As a result, in practice it is critical that educational professionals determine whether an EBI proves effective for each child after implementation. This obligation is even more critical when a child's response to an EBI will be a part of special education eligibility determination. Indeed, understating what an EBI is and what the label of EBI really means, it is clearly not defensible to simply select, implement, and assume that the child will respond. When the level of interest is the specific child, then SCD is the ideal choice of research methodology.

The use of SCD allows for educational professionals to "know what they know, and know what they don't know" in terms of RTI. By adhering to an appropriate SCD, educational professionals can make defensible judgments about the impact of interventions (independent variables) on outcome data (dependent variables) with only one student. As discussed above, in an environment where the stakes of a successful intervention are dramatically higher (particularly, schools using an MTSS model of service delivery), being able to use a defensible methodology to determine whether an intervention was effective is critical.

There are other advantages to SCD beyond clearly demonstrating the effectiveness (or lack thereof) of an intervention. First, it is a highly feasible method of conducting applied research as it does not require the use of a control group or randomization of subjects. Obviously, in schools it is not typically appropriate to have a "control group" of students who are intentionally denied an intervention. One can only imagine a parent's response when he/she finds out that his/her struggling child is receiving a placebo intervention. While an SCD requires baseline data to be collected (and thus, some delay in implementing an intervention), there is no need for a group of children in need to be denied the appropriate services. Ethically, this is a significant advantage of single-case methodology. In addition, SCD is highly flexible. As discussed in Chapter 2, single-case logic (or baseline logic) can be used to build a number of different designs. This range of options makes it much more likely that there will be an appropriate design for each situation. In essence, SCD and analysis is the ideal technology for educational professionals to make more defensible educational decisions in an MTSS model.



## HOW READY ARE EDUCATIONAL PROFESSIONALS?

It is clear that intervention practices in schools are changing and there are technologies such as EBIs, problem-solving models, direct academic/social behavior assessment strategies, and single-case methodology and analysis that exist for educational professionals to utilize in this transition. Unfortunately, it is somewhat questionable how prepared teachers and other educational professionals are to use these technologies. Preservice teachers

report receiving very little instruction in behavior practices or direct academic assessment (Begeny & Martens, 2006) and practicing special education teachers frequently report using instructional practices with little research support (Burns & Ysseldyke, 2009). Moreover, graphing procedures remain remarkably underutilized in elementary and secondary classrooms, and are rarely taught to preservice teachers (Espin, Wayman, Deno, McMaster, & Rooij, 2017). These studies only confirm the widely held belief that national implementation of an MTSS model will be a difficult task (VanDerHeyden et al., 2016). Clearly one of the major challenges in transition to an MTSS model is a significant amount of training for teachers and other educational professionals in some of the most basic skills associated with MTSS implementation.

While these issues are of concern, there are clearly individuals in the schools who should be the conduits for distributing information and training educational professionals to adopt MTSS. Given that group-based data management is a critical component of MTSS (Kovaleski & Pedersen, 2008), one or two individuals with expertise in SCD and analysis should be able to use this technology to assist in individual cases, as well as model it for other members of the team. While this task can seem daunting, in our experiences, such demonstrations are effective training opportunities that can help team members develop the skill set to effectively launch an MTSS model.

## CONCLUDING COMMENTS

The purpose of this chapter was to consider the role of SCD in a modern model of special education service delivery. Clearly, as the stakes of using academic and social behavior intervention increases, the integrity of such intervention must also increase. As such, it is critical that educational professionals steep themselves in the literature focusing on intervention selection and the measurement of intervention effects. This book is not a comprehensive guide of intervention selection or the measurement of the effects of an intervention—therefore, professionals who desire development in reference to those two topics should consult texts on intervention selection (e.g., Burns, Riley-Tillman, et al., 2017), academic assessment (e.g., Hosp, Hosp, & Howell, 2014), and social behavior assessment (e.g., Chafouleas, Riley-Tillman, & Sugai, 2007).

While the selection, implementation, and monitoring of educational interventions are necessary elements of a defensible and effective MTSS model, they are not sufficient. MTSS can best be conceptualized as the systematic use of assessment data to most efficiently allocate resources in order to improve learning for all students (Burns & VanDerHeyden, 2006). Moreover, in order to inform a valid MTSS framework, data need to model academic growth, distinguish between ineffective instruction and unacceptable individual learning, inform instructional decisions, and be sensitive enough to detect treatment effects (Gresham, 2002a). Thus, the terms *MTSS* and *data* are essentially synonymous, but it is not the collection of data that makes an effective system—rather, it is the systematic use of these data. In our opinion, MTSS cannot happen without defensible data collection *and* valid analytical techniques for which to consider data from individuals and groups of students.

This book is not intended to be a complete overview of the philosophy and basis of SCD nor is it intended as a guide for traditional researchers. There are a number of excellent

texts that already accomplish those goals (e.g., Cooper, Heron, & Heward, 2007; Kazdin, 1982; Kennedy, 2005; Tawney & Gast, 1984). Rather, this book is intended to be a guide for educational professionals who desire and are now more often required to make educational intervention an applied research endeavor. As such, the remainder of this book provides a detailed description of SCD, the analysis of intervention data, and guidance as to how to use this technology in schools. In the next chapter, we consider baseline logic in order to have a full understanding of SCD. Chapters 3 and 4 build on the foundation of baseline logic to present a number of SCDs. Chapters 5 and 6 outline the numerous strategies that educational professionals can use to effectively analyze intervention outcome data. Chapter 7 presents a unique application of SCD for assessment of interventions called *brief experimental analysis*. Chapter 8 discusses the use of SCD in educational research. Finally, in Chapter 9, additional guidelines are discussed to enhance the use of SCD and analysis in MTSS environments.