
Data-Based Decision-Making

Stanley L. Deno

The use of data to make decisions is both central and basic in the response to intervention (RTI) approach. While data are central to effective RTI, the procedures for basing RTI decisions on data are complicated and varied. The purpose of this chapter is to consider core ideas of data-based decision-making in RTI, to provide a perspective on issues related to data-based decision-making, and to recommend procedures for maximizing success in data-based decision-making.

Why Data-Based Decisions? A Personal Story

Early in my tenure as a faculty member at the University of Minnesota, I was fortunate to receive funding to develop a collaborative training program for special education teachers in a nearby Minneapolis elementary school (Seward Elementary). A primary feature of that project involved spending my days in the school to help create a noncategorical special education resource program by disestablishing two segregated special classes. At the time, the setting was referred to as a university field station, and my role included working there each day with six to eight university students. The goal for the university students was to learn how to function in,

what was then, a new special education program model in which students with high-incidence disabilities would spend most of their days in general education classrooms rather than special classes and receive supplementary instruction from resource teachers.

The role and procedures for functioning as special education resource teachers in such a setting were, as yet, undeveloped. And so the author's job was to develop both program and the role requirements for the resource teachers in training. The primary goal of our work there was to create a supportive academic program for all special education students that would enable them to acquire basic academic skills in reading, writing, and arithmetic. Since the students, now integrated, presented significant management challenges for their teacher, it was also necessary to develop the capacity of the teachers to address the social behavior of the special education students. In the approach to teacher education, each university student assumed responsibility for the programs of several of the Seward special education students who were now integrated into general education classrooms. My role was to provide direction and support for the university students as they assumed responsibility for designing the special education students' programs and that, typically, included supplementary tutoring.

Moving ahead in this project, the decision that was most challenging for me was the type of instructional program that the university students would use in attempting to increase the basic skills

S. L. Deno (✉)
University of Minnesota, Minneapolis,
MN 55455-0364, USA
e-mail: denox001@umn.edu

of their special education students. I soon discovered that, while I had my ideas about what they should do, the students were taking classes from other university professors who were recommending approaches to teaching that were quite different from what I was recommending and different from one another. The challenge that I faced was how to resolve this dilemma—teachers in training who were getting competing perspectives on how best to teach struggling students.

Most of what is written in the remainder of this chapter emerged from that early experience in trying to provide professional training when competing ideas and uncertainty exists regarding how best to solve the problems confronted by struggling students. *The decision then, as now, was to be sure everyone became committed to the problems to be solved rather than the solutions they were going to use.* In effect, I wanted the goals the teachers in training would be using to be predetermined and specific and for those goals to drive their work. I wanted the means for attaining those goals to be free to vary so that the teacher trainees were motivated to integrate and use what they considered the best approach they could create from what they had learned from their various professors and from their experience. I also wanted them to be free to modify the programs that they created when those programs seemed not to be effective.

To create the context for this approach to intervening with struggling students, it was clear that two critical needs existed. First, to have a way to establish the determinacy of the goals; and, second, to have procedures for evaluating whether or not instructional programs were leading to the attainment of those goals. It was assumed that with these two components in place, to unleash the creative energy and motivation of the teachers in training without placing their special education charges in academic jeopardy would be possible.

Data-Based Program Modification

The approach that emerged from the early work at Seward Elementary School was called “data-based program modification” (DBPM) and was

detailed in a book of the same name (Deno and Mirkin 1977). DBPM was conceived then, as now, as a sequence of decisions that are made when modifying a student’s program and a parallel set of evaluation activities that are intended to increase the likelihood that program modification would be successful. The evaluation procedures were designed to inform the decisions to be made, and the central feature of the evaluation procedures used for each decision was data collection. In the elaborated form of DBPM, each decision made includes specific measurement, evaluation, communication/collaboration, and consultation activities intended to improve the outcome of the decision.

DBPM Assumptions and Current Considerations

Presented below are the five assumptions in the DBPM manual that provide the basis for the approach. They are as important now as they were then, but the original wording has been modified to clarify their relevance to RTI. Each of the assumptions is discussed as it relates to data-based decision-making in RTI.

Assumption #1 The effects of program changes/interventions for individual students are typically unknown and should be treated as hypotheses.

With the widespread emphasis on selecting evidence-based interventions in RTI, it is sometimes easy to forget that the evidence of effectiveness rests on outcomes for groups rather than individuals. Beyond that, in virtually all experimental comparisons, there is overlap between the distributions for treatment and comparison groups. In a practical sense, this means that some students receiving the more effective, evidence-based intervention actually had the same, or poorer, outcomes as students in the less effective intervention. What this means for implementing RTI interventions is that, depending on the effectiveness of the existing program, one can have confidence that the evidence-based intervention is likely to boost the progress, in general, for a group of students, but it cannot be said that will

be the case for every student in the group. One needs to be mindful, then, that as an evidence-based intervention is implemented that one cannot predict for whom the intervention will work, and for whom it will not.

Assuming that the purpose of RTI is to improve the likely outcomes for all students, the fact that one cannot predict with certainty the students who will benefit from an intervention means that, no matter how strong the evidence for an intervention, careful attention must be paid to the effects for individual students. This is sometimes difficult to accept when one is working hard on assuring the fidelity of an intervention, but it is a reality that must be addressed. It also forces on us the prospect of having to create and manage a more differentiated program at each tier of our RTI model in an effort to be responsive to the progress of all students. Complete individualization is not logistically feasible, of course, but placing students in programs for most of the school year without being responsive to individual intervention effects would be unacceptable.

Assumption #2 Single-case research designs with repeated measurement data are well suited for testing intervention hypotheses.

Assuring that one is using data to monitor intervention effects for individual students requires practical program evaluation procedures that provide the best possible evidence for determining whether a program is meeting its goals for individual students. At this point in the development of RTI models, as has been the case in the past, the best available evaluation procedures are those that involve repeated measurement of individual student performance. Repeated measurement of performance across time produces a record (a time series) depicting the extent to which a student is progressing satisfactorily toward performance goals (Skiba and Deno 1984). These data can be configured and displayed graphically to provide an easily interpretable picture of a student's past, present, and likely future performance (see Fig. 1, for example). Even though repeated measurement of performance produces a data record that enables evaluation of intervention effects, it is important to recognize that many questions exist regarding the best approach to implementing a repeated measurement system. One seemingly simple question is how frequently student performance must be measured to provide a sufficiently reliable and valid database for

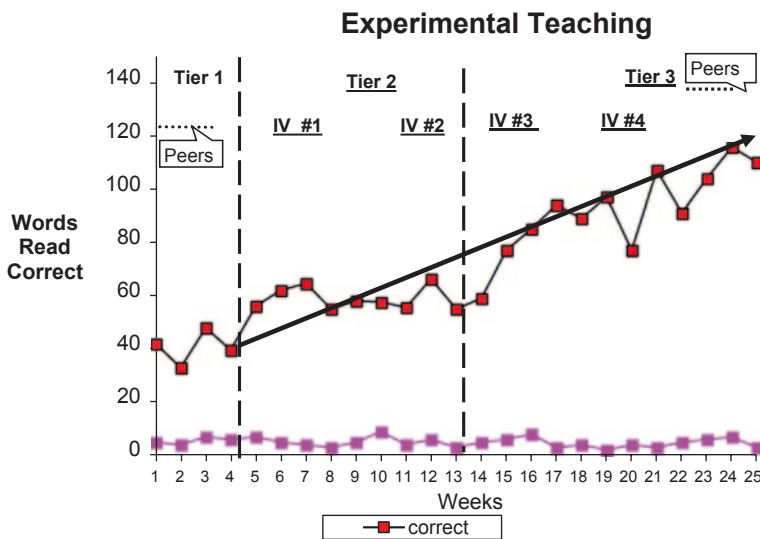


Fig. 1 Experimental teaching

testing individual intervention effects (Jenkins et al. 2005; Jenkins et al. 2009). This includes both practical and technical considerations. Since the system is to be logistically feasible, keeping the number of measurements to the minimum necessary to provide reliable and valid data is practically important. At the same time, making a judgment about whether a student is making sufficient progress as a result of an intervention requires a conclusion based on a student's rate of growth before and after the intervention has been implemented. Unfortunately, research on estimating growth from repeated measurement data has not sufficiently answered the question of how many measurements are required to reliably estimate growth (Ardoin and Christ 2009).

Assumption #3 Modifications in the general education program for individual students are hypotheses whose effects should be empirically tested.

It goes without saying that the best possible RTI model rests, first, on a core educational program for all students (tier 1) that is likely to be maximally effective. It should be the goal of every RTI model to minimize the number of students who will be identified as needing a tier 2 intervention. Having said that, even the most effective tier 1 programs will contain students whose needs are not being sufficiently addressed in the core program and who are at risk for academic difficulty. In such cases, students will be placed in an alternative or supplementary program (tier 2) focused on reducing their risk of failure. As this is done, it is important to be mindful that any program differing from the general education program rests, in fact, on the hypothesis that the alternative program will benefit students placed in that program. As stated earlier, no program is going to benefit all of the students all of the time. Thus, placement in an alternative program requires careful monitoring of program effects.

It is important to be reminded of this need to monitor the effectiveness of tier 2 programs because progress monitoring requires a commitment of time and resources that is tempting to avoid. In some cases, it might seem that a stu-

dent should experience the alternative program for half or, perhaps, an entire school year to determine its effects. In such a case, it might seem reasonable to wait until the end of the program to evaluate its effects. Waiting for such a long duration to test the intervention effects, however, risks the very real possibility that some students are not benefitting from the intervention and that a large segment of the school year might pass before changes are made to those students' programs. A responsive RTI model includes procedures for frequently evaluating intervention effects throughout the school year rather than waiting for most of a school year to evaluate intervention effects.

Assumption #4 To use single-case designs for testing intervention hypotheses requires specification of the "vital signs" of educational development that can be routinely obtained in and out of school.

Forty years ago, simple and direct measures of student performance that could be routinely obtained to monitor the academic health of young students did not exist. Student performance was informally appraised by teachers through daily work with their students and occasional formal assessment was accomplished through school-wide standardized achievement testing. Accountability requirements were virtually nonexistent and no requirement existed to assure that students given remedial assistance actually benefited from that assistance.

In the intervening years, substantial amounts of effort and resources have been devoted to developing technically adequate and logistically feasible measures of student academic growth. Now, a menu of options exists from which educators can select progress-monitoring measures that enable "vital sign" monitoring and can serve as the basic data for making RTI programs responsive to student performance (National Center on Response to Intervention 2012). A review of the National Center on Response to Intervention "tools chart" reveals that different approaches have been developed to monitor student progress. The center's evaluation of those measures is intended to establish that all included measures

can, potentially, be used within an RTI model. As is also clear from the chart, however, most of the progress-monitoring measures are based on the general outcome measurement (GOM) approach (Fuchs and Deno 1991) on which the curriculum-based measurement (CBM) approach is based (Deno 1985).

Research on this approach spanning more than 30 years has supported the technical adequacy and the utility of the GOM/CBM approach for generating the student performance data necessary for making a wide range of programming decisions (Espin et al. 2012). A very large number of juried empirical and nonempirical articles on GOM/CBM have been published in the professional literature of school psychology and special education (Speece 2012), and many states include those measures in their statewide testing. Nevertheless, several important issues on the use of these procedures require consideration.

The first issue to consider is whether the technical adequacy of GOM/CBM data is sufficient for making the high-stakes decisions that are required in RTI. CBM was developed as a tool for teachers to use in formatively evaluating their instruction with students who were academically disabled. The original research hypothesis to be tested using CBM data was that teachers could increase student achievement by collecting CBM data and using it to make decisions about whether to continue or change their instruction. Initial research supported this hypothesis (Fuchs et al. 1984). Since that early work, GOM/CBM has been increasingly used to make a wide range of decisions, including screening for alternative programs, writing annual individualized education program (IEP) goals, eligibility for special education services, evaluating program effects, and as evidence in legal proceedings (Espin et al. 2012).

As has been pointed out by others (Davis et al. 2007), placement of students at different levels in an RTI model is a high-stakes decision as it is practiced in most schools. CBM/GOM procedures are often used in the screening of students to identify those at risk for failure. While GOM/CBM procedures have been, and can be, adapted to a variety of uses, the extension to making

decisions where the stakes are higher than that for which the measures were initially developed must be done carefully to assure that the data are valid for the decisions being made. Evidence exists that when using only GOM/CBM data to screen grade 1 students to identify those who will be placed in tier 2, a significant number of errors are made (Jenkins et al. 2007). A better approach is to use GOM/CBM as one measure in a battery of assessments (Compton et al. 2006; Jenkins and O'Conner 2002). Doing so minimizes classification errors to a degree that both assures that students who require intervention will receive it and that resources will not be spent on students who will succeed without intervention.

Other issues related to using GOM/CBM in an RTI framework exist, as well. Repeated measurement of individual student academic performance creates the opportunity to measure student growth and current level over the relatively short durations of an intervention. When teachers are using the measures informally to aid in their judgments about how students are progressing, the stakes are relatively low. However, when growth is being used to determine whether students are benefitting from their current tier placement or whether a move to a more intensive level is required, this is a high-stakes decision. As with using GOM/CBM as a measure for screening, using the repeated measurement data it produces to estimate growth is fraught with technical problems that have not yet been adequately solved. Growth trends estimated from GOM/CBM tend to be quite inconsistent and unreliable, with large confidence intervals that make data interpretation difficult (Ardoin and Christ 2009; Christ and Coolong-Chaffin 2007). Despite the problems associated with making high-stakes decisions using GOM/CBM growth data, the GOM/CBM approach is one of the only viable alternatives for examining individual student growth since other achievement measures tend to be insensitive to change over relatively short durations (Marston et al. 1986; Skiba et al. 1986). Continuing research and development designed to improve the technical characteristics of progress measures is underway and promises to address this important issue (Christ and Ardoin 2009).

Assumption #5 Testing intervention hypotheses requires well-trained professionals capable of drawing valid conclusions about program effects from the “vital signs” data.

The fifth assumption focuses on the heart of data-based decision-making: the decision-maker’s use of data to make intervention decisions. As was true in the mid-1970s, the key element in successful, data-based programs is the people responsible for organizing and implementing programs and using data to increase their effectiveness. In the early development of data-based programs using vital sign progress monitoring, the decision-makers were special education teachers using the data-based approach to intervene with individual special education students. As the data-based approach has become more widespread, the training and competence of an increased number of professionals has come into play. Now, classroom teachers, school psychologists, and program administrators, among others, are required to examine student performance data to make decisions ranging from individual student progress to program evaluation. Regrettably, collecting and interpreting student performance data has never been a major component of teacher education programs (Stiggins 1993). Almost certainly, the same might be said with regard to the training of school administrators. The one group of educational professionals who are highly trained in assessment is the school psychologist. Not surprisingly, school psychologists often become key personnel in data collection, management, and interpretation in RTI. Typically, however, their work is done at the program level, rather than the individual student level, simply because the ratio of school psychologist to students is limited. Consequently, teachers still remain at the front line in the interpretation of individual student data.

Since teachers bear the primary responsibility for monitoring and interpreting individual student progress data, it is important to consider how well they are able to do this task. Espin and colleagues recently examined the accuracy and quality of teachers’ interpretations of the typical progress-monitoring graphs used in RTI (Espin et al. 2012). The results of their research

provide both good and bad news. Teachers’ interpretations were compared to the interpretation of experts, and while some of the teacher’s interpretations were consistent with the experts, others differed widely. Espin and colleagues described the poorer interpretations as lacking coherence, specificity, and accuracy. Obviously, there is room for improvement in the interpretation of student progress graphs. Even as this point is made, however, one needs to remember that, at best, the inconsistency of most progress measures for modeling individual student growth means that even experts in interpreting such data are using a database that is insufficient for making high-stakes decisions.

Using Decision Rules to Make Intervention Decisions To take full advantage of the repeated measurement data derived from progress monitoring requires not only that the data be collected and summarized but also that those managing interventions be responsive to the data in an effort to build a more effective intervention. Unfortunately, even when progress-monitoring data are systematically collected throughout the course of the academic year, it is common to find that teachers might make only one or two deliberate changes to increase intervention effectiveness (Deno et al. 1988). The problem one confronts is how to assure that the data collected during intervention are used as much to evaluate and improve the intervention as they are to evaluate the student.

A Caveat: The Data Do Not Speak for Themselves An oft-heard fallacy is that “the data speak for themselves.” This, of course, is not true. The data are only numbers without meaning until one interprets them. In a short article on this topic, Behn (2009) makes the point by using the familiar example of a glass of water filled to its midlevel. The optimist describes the glass as half full, while the pessimist claims it is half empty. Both are using the same data point, but each interprets it differently. When we collect data on student performance, the situation is precisely the same. Student scores are meaningless until we apply our perspective to those scores. Even

as we summarize data using simple descriptive statistics like averages and percentages, we need to be aware that our history with those statistics tends to operate in such a way that we are immediately attributing meaning and significance to those numbers. When we screen students, the numbers that are generated become valuable to us when we apply a decision framework to those numbers for purposes of identifying those students who are progressing well and those for whom additional attention will be necessary. We need to be fully aware that the decision framework that is used rests on both assumptions and values regarding the importance of students performing at different levels. While it is useful to have available standards and cut scores to enable us to make decisions, we cannot be misled into thinking that the data speak for us. Carrying out system-level directives does not change the fact that we are accepting the values and the decision frameworks designed by others and, in so doing, we complicit in making the decisions.

Professional Expertise in Decision-Making As we consider the expertise required to make data-based decisions in RTI, a question to consider is, how to increase the likelihood that the decisions made improve the outcomes? A pervasive assumption is that the best possible approach to data interpretation and decision-making is one rooted in the collective judgment of experienced decision-makers. However, this assumption has been both questioned and contradicted in the clinical psychology literature. In his classic book, *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*, Paul Meehl examined the accuracy of clinicians in predicting individual outcomes. As is true for educators, most judgments about the clients of clinical psychologists involve predicting future outcomes from a variety of data sources (records, interviews, test scores, etc.). Once the data from these sources is quantified, two options exist. The first option is for professionals and others with a vested interest in the outcomes to examine the data and make a prediction regarding the best treatment based on their interpretation of the data (i.e., clinical prediction). The other approach that

Meehl examined was to use a formula created to make the prediction (i.e., a statistical prediction). In the latter case, once the formula has been created, that prediction decision is the product of the formula rather than the professionals interpreting the data. It is important to note that, even though a formula is being used, professional judgments can be included in the formula. However, in the statistical approach, the professional judgment does not determine the decision, the formula does. After reviewing the research where the two approaches were compared empirically, Meehl found that the evidence supported a statistical, rather than clinical approach, in making predictions. As might be expected, Meehl's book has become the center of controversy in the field of psychotherapy, but his conclusions remain supported after more than 50 years (Grove and Lloyd 2006).

Meehl's findings are important for developing data-based RTI programs that involve predicting student outcomes. The very nature of RTI involves a host of significant decisions, including identifying students who are at risk of continuing academic difficulty and who will benefit from intervention, the types of interventions that will best serve students, whether students are benefiting from intervention, whether students should be moved to a more intensified level of intervention, and so on. As one seeks to make data-based decisions, it would be well to consider whether there is an opportunity to use more statistical, rather than a clinical, approaches to making those decisions. Certainly, from the standpoint of accuracy, an argument can be made that we should try to do whatever we can to increase the accuracy of our predictions. From the standpoint of efficiency, using formula-based approaches could well reduce the amount of personnel time involved in making the many RTI decisions, potentially, increase responsiveness to student progress data. While using a formula-based approach might seem mechanistic and impersonal, as is discussed in the next section, such an approach is already being used by many practitioners using progress-monitoring data.

Decision Rules The use of formula-based approaches with progress-monitoring data is

described in the literature as “data decision rules” (Ardoin et al. 2013). Teachers have been trained in the use of data decision rules for more than 30 years (White and Haring 1980), and decision rules were part of the effective intervention reported previously (Fuchs et al. 1984). Those decision rules are generally based on data in the progress-monitoring graph where current student performance is plotted against the desired line of progress based on students’ year-end goals. The slope of that line of progress is established by drawing a line that connects students’ initial level of performance to the level established in the goal (see Fig. 1). Basically, two different types of decision rules based on that progress line have been used to prompt teachers to change programs or goals (Ardoin et al. 2013). The first simply specifies that the intervention should be changed if a prespecified number of consecutive data points fall below the line. The second requires that the slope of student scores be estimated and that slope then compared to the slope of the desired line of progress. If the slope is flatter than the slope of the progress line, then the rule directs a change in the intervention. If the slope is similar to the slope of the line of progress, then the rule is to continue the existing intervention. If the slope of student scores exceeds that of the desired progress line slope, then the rule directs either raising the goal (thereby increasing the slope of the progress line) or moving to a less intense intervention.

As has already been discussed, trend estimates from progress-monitoring data tend to be very unreliable. To overcome this difficulty, simple paper-and-pencil procedures have been developed that improve agreement in trend estimates based on the existing data (Skiba et al. 1989; White and Haring 1980). Improving agreement across those interpreting the progress-monitoring data is important, of course, but if the data themselves are unreliable, agreement in interpretation will not increase the reliability of the database for the decision to be made. Indeed, as Ardoin and colleagues concluded in their review, the present state of the empirical evidence regarding the effectiveness of formula-based decision rules is insufficient to recommend any particular decision

rule for guiding when to change or maintain an intervention (Ardoin et al. 2013).

As computers have become readily available to collect, store, and analyze student progress data, additional possibilities for using formula-based decision-making have become possible (Fuchs et al. 1989a). Now, instead of using paper-and-pencil procedures to estimate growth, computers can quickly provide consistent information about student performance. The fact that computer technology can be used to enhance teacher planning and improve student outcomes was established by Fuchs and colleagues more than 20 years ago (Fuchs et al. 1989a, b). Surprisingly, little has been done in the intervening years to build decision rules into computer-based systems. As a recent summary of the use of technology with GOM/CBM makes clear, web-based systems are now available to facilitate progress monitoring (Lembke et al. 2012). At the same time, the development and use of technology has focused more on data summaries and presentations and tend not to include decision rules designed to direct change in student programs. Perhaps this is a good thing, given the findings of Ardoin and colleagues that the evidence base for decision rules is limited. At the same time, the availability of computer technology should enable researchers and developers to design effective decision rules that can be incorporated into those now widely used web-based systems.

An alternative approach to using computer analyses of student performance to enhance intervention is illustrated in the work of Fuchs and colleagues (Fuchs et al. 2005). The focus of that work has been on adding diagnostic information to the progress monitoring so that teachers derive specific information about the profile of skills that students have mastered and those still needing attention. These data are then used to both group students with similar skill sets and to design interventions related to those skills. That group of researchers has been able to demonstrate that developing supplementary diagnostic data can indeed enhance instructional effectiveness. At the same time, collecting data on skills thought to contribute to developing generalized competence in reading, writing, and mathematics

can also result in a shift in attention from dependent variables to independent variables.

Distinguishing Between Dependent and Independent Variable Data

One of the most basic ideas taught to beginning researchers is the distinction between independent and dependent variables. Independent variables are those factors that the researcher manipulates directly in an effort to determine whether those manipulations produce a change in a measured outcome. In educational interventions, the independent variables are those changes (interventions) made in students' programs to produce higher levels of achievement. In contrast, dependent variables are not directly manipulated by the researcher. They are the outcomes, or the effects, that enable determination of the power of the independent variables. This research, of course, is an exploration to determine whether cause-and-effect relationships exist between the independent and dependent variables. In education, the primary dependent variables are the achievement scores collected annually. Most often these are state standards tests or standardized achievement tests. In RTI, the dependent variables need to be more routinely available than the annual state achievement or standards tests so that judgments can be made more frequently regarding the effectiveness of the interventions. That is one of the primary reasons why progress-monitoring data are collected and, in RTI, the dependent variables are typically progress measures.

It takes only a short step to see that in our work we are continually involved in manipulating variables in the environment in an effort to positively impact student outcomes. Thus, teaching, in general, can easily be viewed in terms of independent and dependent variables; and RTI, in particular, can be viewed as applied research in which combinations of independent variables are sought to produce desired changes in the dependent, progress-monitoring data.

Why is this distinction between independent and dependent variables important for us when considering data-based decision-making, and

how does it relate to our work? The answer is that high-quality RTI programs rest on collecting and using the best possible dependent variable data. All of the most important decisions made in implementing RTI are grounded in the assumption that the data used to make program change decisions are reliable, valid, and possess consequential validity (Messick 1989). Before we begin implementing our interventions, our independent variables, we must be sure that we have measures of student outcomes in place that provide us with the information necessary to evaluate intervention effects. Further, to the degree possible, we should strive to be sure that continuity exists in the dependent variable data used to make the decisions about the effectiveness of our programs. Consequently, to the extent possible, we should collect and use the same data to identify students who are academically at risk, to monitor their progress during intervention, to evaluate the effectiveness of the interventions that have been created, and to move students to a different level within the RTI framework. If students are identified as at risk by using a battery of measures, then it will be most useful if the ongoing dependent variable data are included in that battery and collected throughout intervention and program evaluation. Because all measures possess some degree of unreliability, using different measures for different decisions increases the likelihood that measurement error will be compounded and that inaccurate decisions will be made as a function of this compounding of error.

While dependent variable data can aid us in deciding whether a program is succeeding, it is not designed to provide information about potentially effective independent variables that might be included in an intervention. A simple analogy based on the vital signs is the thermometer. It is used regularly to measure body temperature, but it provides no information on what to do if the temperature is too high. Similarly, a bathroom scale can help us to monitor weight gain, but it does not provide direction regarding what to do should we seek to change our weight. When GOM/CBM is used as the primary dependent measure, like the bathroom scale, it does not direct us to techniques or interventions that might

alter the GOM/CBM scores. That information must come from other sources.

In an effort to provide decision-makers with data that can be used aid in designing interventions, some programs incorporate criterion-referenced tests that are used in initial identification of students at risk. For example, a criterion-referenced math test might be used to determine skills students have not yet mastered so that practice on the skills might be included in an intervention for those students. Another example is an early reading assessment that identifies decoding skills that students have, and have not, mastered. It is important to emphasize, however, that even though data have been collected revealing lack of skill mastery, the impact of including practice on those skills remains a hypothesis that must be tested by determining whether teaching those skills results in changes in the dependent measures. In other words, do the students' scores increase as they learn the skills?

When data are collected on skills thought to be necessary for a student to become generally competent in reading, writing, or arithmetic, it is relatively easy to become primarily focused on the skills that have been identified and to lose track of performance on the dependent measures. When that occurs, it is easy to mistakenly base evaluations of program success on the extent to which the identified skills have been mastered rather than whether real change has occurred in the generalized competence sought. This happens most often when students are assessed with a measure that includes a variety of subtests that have associated benchmark scores.

Suppose, for example, an initial reading assessment is conducted that includes measures of phonemic awareness, letter naming, onset phonemes, blending and nonword identification, and that each subtest has a "mastery" benchmark score. Given this set of initial reading skills, one might create an intervention targeting the skills in the assessment. And, as one implements the intervention, it might seem sensible to evaluate intervention effects by collecting data on whether students are mastering the various skills targeted in the assessment. The question that must be answered in this case is whether skill mastery

constitutes independent variable manipulation or whether they are the primarily dependent variables. The answer is that the data collected on mastery of those skills enable us to determine whether the instructional approach has been implemented, not whether the students have improved in their general reading competence.

To test the hypothesis that the students meeting benchmarks on specified pre-reading skills will become better readers requires not only that we measure their attainment of those skills but also that we use a measure that possesses validity as a general measure of reading. As the students achieve benchmark performance on the skills, we should be able to observe a corresponding increase in our dependent measure. Referring back to my personal experience in training preservice teachers, our first step in building a data-based program was to create the dependent variables, the measures of student performance that the trainees would use as a basis for testing the effects of their ideas about how best to teach their students. To apply research techniques in creating effective interventions, they needed to be free to use alternative independent variables when their initial efforts were not effective. Focusing on the dependent variable enabled them to think flexibly about alternatives that might effect change in the dependent variable, rather than being committed to particular intervention approaches.

It would not be surprising if many were to think that the reading skills identified in the above example would be the appropriate dependent measures in an early reading intervention. They might argue that these should be the outcomes toward which interventions are directed. To take that position, however, requires research that measures of these skills possess criterion validity as reading measures. A good example is the case that can be made for two other measures that might serve as the dependent measures of reading in first grade. Both isolated word recognition and reading aloud from text can be measured throughout the entire first-grade year and more closely represent actual reading. For that reason they, for example, would function well as dependent measures for evaluating intervention effects in beginning reading (Fuchs et al. 2004).

Handbook of Response to Intervention
The Science and Practice of Multi-Tiered Systems of
Support

Jimerson, S.R.; Burns, M.K.; VanDerHeyden, A.M. (Eds.)
2016, XXV, 740 p. 89 illus., 26 illus. in color., Hardcover
ISBN: 978-1-4899-7567-6