2016

Data-Based Decision Making in School Counseling: Utilizing Multiple Single-Case Indicators to Evaluate Interventions

Ryan J. McGill
College of William and Mary

Kelly S. Kennedy
Chapman University, kkennedy@chapman.edu

Randy T. Busse
Chapman University, busse@chapman.edu

Recommended Citation

McGill, R.J., Kennedy, K.S., & Busse, R.T. (2016). Data-based decision making in school counseling: Utilizing multiple single-case indicators to evaluate interventions. *The Practicioner Scholar: Journal of Counseling and Professional Psychology 5*(1): 104-121. Retrieved from http://www.thepractitionerscholar.com/article/view/15972/10380

Data-Based Decision Making in School Counseling:
Utilizing Multiple Single-Case Indicators to Evaluate Interventions

Ryan J. McGill, The College of William and Mary
Kelly S. Kennedy, Chapman University
R. T. Busse, Chapman University

Abstract

As the field of professional school counseling continues to move toward a data-based decision making model of service delivery, there is a need for dissemination of best practice methods for evaluating whether school-based counseling interventions are effective. In that vein, the purpose of this article is to review several methods of data-based decision making within a single-case outcome evaluation model, as well as their potential applications for school counseling interventions. To aid practitioners, the potential use of these methods is demonstrated in a case example and accompanying graphic displays.
Keywords: school counseling, single-case design, intervention outcomes, effect sizes, evaluation

Data-Based Decision Making in School Counseling: Utilizing Multiple Single-Case Indicators to Evaluate Interventions

Since *A Nation at Risk* (National Commission on Excellence in Education, 1983) was published over 30 years ago, the landscape of public education has been dramatically influenced by the standards-based reform movement. Such calls for greater accountability have influenced all aspects of the educational system, including professional school counseling (Dollarhide & Lemberger, 2006; Sink, 2009). In response, school counselors have been encouraged to utilize data-based decision making models to demonstrate the effectiveness of their service delivery and individualized interventions (Hatch, 2013, Issacs, 2003; Whiston & Sexton, 1998).

Data-Based Decision Making in School Counseling

Data-based decision making serves as a foundation of the most recent iteration of the American School Counselor Association national model for school counseling programs (ASCA, 2012). Specific standards addressed under the proviso of accountability include the ability to provide stakeholders with *quantitative* [emphasis added] evidence of program success, the ability to disaggregate important information from available data sources, and the ability to evaluate the effects of individual and small group interventions on the social and academic performance of students. ASCA (2012) defined the use of data as an "accountable method to align the school counseling program with the school's academic mission" (p. l6). In addition to promoting the collection, analysis, and dissemination of data by all school counselors, such activities are essential requirements for school counseling and guidance programs seeking to be identified as Recognized ASCA Model Programs (RAMP).

Although data-based decision making has been embraced by accrediting agencies and professional school counseling associations (e.g., ASCA, CACREP), data analysis strategies often employed by many school counselors fail to meet established national standards (Dollarhide & Saginak, 2012). As an example, in a survey of contemporary school counseling practice, Dimmitt and colleagues (Dimmitt, Carey & Hatch, 2007) found that the most prevalent form of data collection and analysis amongst practitioners was documenting the number of student contacts by individual counselors. However, student contact data do not provide practitioners with information relevant for evaluating the effectiveness of interventions, because such data are not tied to actual intervention outcomes (e.g., attendance, referral rates, and academic performance). In a more recent survey of 114 school counselors working in programs that had earned RAMP designation, Young and Kaffenberger (2011) found that almost 20% of respondents indicated that they rarely utilized formal assessment to evaluate the effectiveness of interventions. To address these limitations, school counselors have been encouraged to utilize more robust methods for assessing intervention outcomes such as the use of single-case design methods (American School Counselor Association, 2012; Foster, Watson, Meeks, & Young, 2002; Lundervold & Belwood, 2000) as well as recognized statistical techniques for evaluating change within the human services, such as effect size estimators (Brady, Busse, & Lopez, 2014; Perdrix et al., 2012).

To address the aforementioned gaps in data-based decision making within the field of school counseling, the purpose of this article is to present six single-case outcome methods for evaluating individual intervention data that may be useful in the practice of professional school counseling. Although there are a variety of methods for assessing the effects of applied interventions, we have chosen to focus particularly on single-case techniques because of their practicality, flexibility, and use in pre-referral intervention systems such as response to intervention (RTI).

Single-Case Design and Treatment Effectiveness

Single-case design (SCD) is a versatile methodology that has been utilized to evaluate clinical and practice-based outcomes in several fields of applied psychology. In SCD, the effects of a treatment or intervention on an individual are assessed across time using an outcome measure (e.g., rating scale data, observational data), wherein an individual serves as their own control to compare performance across baseline/pre-test and intervention phases (Barlow, Nock, & Herson, 2008). After the implementation of a well-designed intervention, SCD can be a valuable analytical tool for evaluating various treatment outcomes that does not require overly technical sophistication on the part of the user (Riley-Tillman & Burns, 2009). As a result, SCD potentially provides school counselors with a useful framework for making defensible judgements about the efficacy of applied interventions and is considered by some (e.g., Brown, Steege, & Bickford, 2014; Riley-Tillman & Walcott, 2007) to be the best method for providing these data in school-based settings. It should be noted that although SCD can be utilized to assess group-level effects, doing so violates the fundamental logic of the method as it is most versatile when used to appraise individual-level effects (Riley-Tillman & Walcott, 2007).

Despite the widespread use of SCD within the school-based intervention literature, there is a lack of consensus as to how best to evaluate such data (McGill, 2016; Parker, Vannest, &

Davis, 2011; Vannest, Davis, & Parker, 2013). Within the past decade, several strategies have been proposed to evaluate single-case outcomes. We begin with an overview of six selected types of data collection available to counselors that may be most useful, approachable, and applicable in school settings, including: Goal Attainment Scaling, Visual Analysis, Visual Analysis Rating, Trend Analysis, Percentage of Non-overlapping Data, and Reliable Change Index. Each of these assessment methods are described, along with the strengths and limitations of each approach and their utility for assessing specific school counseling outcomes. We conclude the article by demonstrating the practical application of a suggested framework for evaluating single-case outcomes within the context of a school counseling case study that employs multiple single-case outcome indicators.

Before we begin, it is important to provide some theoretical context. In practice and research, evaluators have traditionally assessed treatment outcomes using a dichotomous yes/no framework to determine whether an intervention was effective (Busse, McGill, & Kennedy, 2015). As per Jacobson and Truax (1991), we believe that school counselors (and others in the field of education, e.g., school psychologists, special educators) would be better served utilizing a framework in which the utility of an intervention is based upon an examination of the magnitude of its effect on social or academic behavior. Briefly, an effect size is a quantification of the magnitude of difference between pre-intervention (baseline) data and post-intervention data that allows practitioners to evaluate progress toward specified treatment goals, and to evaluate the magnitude of intervention outcomes (Swaminathan & Rogers, 2007). School counselors can then utilize this information to determine whether an intervention should be continued, discontinued, or modified in some way to continue progress toward intervention goals for a student.

Selection of Evaluation Tools

Before presenting the various ways in which a school counselor can evaluate their data, it is important to discuss the importance of selecting the potentially most practical methods or tools to represent change or growth for a counseling intervention. Even if a counselor selects more sophisticated methods for analysis, the method(s) will not support the efficacy of an intervention if quality, relevant data were not collected. In all cases, efforts should be made to utilize measures or indicators that provide reliable (i.e., consistent and with minimal error), and valid (i.e., measuring the intended constructs) representations of the targeted behaviors that school counselors are aiming to change via their interventions (Schmidt & Hunter, 1996).

As counselors make decisions regarding how to evaluate their interventions, they must of course keep in mind the unique needs of every student. Considerations such as the referral party (e.g., parent, math teacher, self-referral) and exact area of intervention focus (e.g., study skills, anger management skills) should help in the measurement selection process. In accordance with the ASCA model (2012), school counselors should be cognizant of collecting relevant outcome (e.g., attendance, behavior, achievement) and perception (e.g., attitude, beliefs, reflection of competencies gained) data.

To allow for evaluation methods that measure magnitude of change over time, it is often helpful to collect multiple data points. The approaches described in this article require ongoing data collection, often referred to as progress monitoring. To enable progress monitoring, counselors should select relatively short, practical measures, which often rules out the use of commonly known school-based assessment measures (e.g., behavior rating scales such as the Behavior Assessment System for Children, 3rd Edition [BASC-2]; Reynolds & Kamphaus, 2015; state-wide achievement measures), as the use of these measures on a daily or weekly basis likely would be too cumbersome for most parents, students, and teachers to complete (and too expensive for most administrators to approve) and may be beyond the training level of many school counselors. It is helpful to remember that an emphasis on progress monitoring data does not preclude the use of more comprehensive pre and post-tests. Robust evaluation procedures include multiple types of data; examining behavior for intervention from the perspective of different informants (e.g., student, counselor, teacher, parent), and in different settings (e.g., home, school, with peers). Thus, pre and post-tests could be a useful addition to an evaluation that also includes progress monitoring measures.

Published and Pre-Existing Measurement Tools

A number of reliable and valid progress monitoring tools are available. Measures such as the Behavior and Emotional Screening System (BESS; Kamphaus & Reynolds, 2007) and Brief Progress Monitor (BPM; Achenbach, McConaughy, Ivanova, & Rescorla, 2011) are short, psychometrically sound, multi-rater measures of a variety of behavioral, social, academic, and emotional concerns. In addition to these types of measures, many counseling books and curricula include their own evaluation measures. For example, the Strong Kids series (e.g., Merrell, 2007) includes questionnaires that ask students about their knowledge of the Strong Kids material, and their use of these skills in situations outside of the counseling context. We acknowledge that the use of standardized brief rating systems may not be cost effective for practitioners and many practitioners may not have training in the use and interpretation of such methods of assessment. Although certainly more cost effective, many of the questionnaires and rating forms that accompany much of the published counseling curricula have yet to be validated. We caution practitioners to be attentive to such limitations when utilizing these measures to evaluate intervention effects. Depending on the particular measure that is administered, rating scales are valuable tools which can be employed to assess both “process” and “outcome” data, as defined by the ASCA model (2012).

School-Based Data

As an alternative or supplement to more formal measures, there are a variety of data available in school settings that can be utilized to evaluate counseling outcomes. For example, schools readily collect data regarding attendance, work completion, behavior referrals, and test or assignment scores. In selecting data, school counselors should be cognizant of including those measures that are sensitive to small increments of change over time. For example, if a counselor wishes to measure the impact of their eight-week study skills group on a student’s academic achievement, GPA may not be a useful source of data, whereas the percentage of assignments completed per class and weekly quiz scores for that student may be much more useful. Additionally, to be of use in evaluating outcomes, counselors should focus on data that are relatively frequent. Consider a student who is referred for individual counseling for anger management because she receives four office referrals a week for behavior concerns. Tracking the number of weekly office referrals she has may be an effective way to capture any progress she makes with an intervention. In contrast, some students may be referred for this same intervention who have only been sent to the office for one (albeit serious) incident. For these students, the number of weekly behavior referrals would not be useful data as there is insufficient data to inform adequate problem identification (e.g., Manassis, 2012).

Scaling Questions

An individualized source of data available to many school counselors are the responses to scaling questions often utilized in individual counseling interventions. Approaches such as Solution-Focused Brief Therapy (SFBT; Murphy, 2008) encourage counselors to have clients quantify (i.e., rate on a scale from 1 to 10) their perceptions regarding a variety of issues (e.g., sadness experienced, sense of competence) during individual counseling sessions. Data from scaling questions can be excellent indicators of student progress. Murphy (2008) provided an example in which SFBT strategies (e.g., scaling questions) were utilized to assess therapeutic rapport. Such uses are an example of employing scaling questions for the purposes of assessing perception and/or process data related to an intervention. For instance, a counselor could have a counselee provide scaled responses to statements such as “I controlled my anger this week when provoked,” “I advocated for myself when I needed help in class this week,” or “I used my coping strategies when I started to feel down about myself,” within every session to track the counselee’s perceptions of their progress regarding the skills on which they are working in counseling. With this backdrop, we turn now to present several complementary single-case outcome evaluation methods that may be useful in school counseling practice and research.

Goal Attainment Scaling

Goal Attainment Scaling (GAS; Kiresuk & Sherman, 1968) is a rating scale approach that easily can be applied to school counseling practice. GAS was originally developed as a method to evaluate the effectiveness of individualized mental health services in a clinical setting. The GAS methodology involves: (a) selecting a behavior, (b) defining as objectively as possible the behavior of interest, and (c) creating a scale to evaluate outcomes with positive (improvement) and negative (behavior worsened) values (Elliott & Busse, 2004). Individual treatment outcomes are then rated on levels of progress on a 5 point or 6 point scale (e.g., -2 to +2, with -2 indicating a problem is much worse, -1 indicating a problem is somewhat worse, 0 indicating no improvement from baseline, +1 indicating some improvement, and +2 indicating the intervention goal has been met). GAS ratings can be completed by school counselors, students, parents, aides and teachers. Multiple ratings also can be synthesized within the context of a team evaluation format.

The strengths of the GAS method are that it is time efficient, it can be utilized as a primary or supplementary measure to evaluate a wide variety of school counseling outcomes, and it can be utilized to assess generalization of individual intervention effects (i.e., a teacher can rate student progress in the classroom while a concurrent intervention is delivered by the school counselor in a different setting). The limitations of GAS are it relies on subjective ratings of goal attainment and some behaviors do not lend themselves to accurate ratings in the absence of direct assessment (e.g., school engagement, attitude toward classroom teacher; See Brady, Busse, & Lopez, 2014; Coffee & Ray-Subramanian, 2009; and Roach & Elliott, 2005 for useful examples and discussion of the GAS method.) The authors believe that the GAS method is particularly useful for assessing perception data, as defined within the current ASCA model.

School counselors frequently conduct informal check-ins or mini "counseling" sessions with individual students for a variety of referral problems. GAS can be utilized to provide evidence of the effectiveness or generalization of these sessions to other settings (e.g., the classroom). For example, if a counselor checks in with a student for 15 minutes a week because of issues related to homework completion, they could develop a GAS scale utilizing the rating system described above and administer it to the classroom teacher at the end of each week to determine whether the weekly check-ins were resulting in better performance in the classroom.

A particular strength of the GAS metric is that is forces counselors to think critically and realistically about the goals for an individual intervention during that intervention's early stages. In setting up the GAS metric for any student's goals, a school counselor needs to take into consideration factors such as the student's baseline level of performance, level of ability (if applicable), etc. to make determinations regarding what would constitute some improvement and what might constitute a goal being fully met. This level of subjectivity also is a limitation of GAS, in that one school counselor's metric and subsequent ratings may be somewhat different than another's for a similar intervention or case. Examples of GAS metrics for a student with academic study skills goals are provided in Table 1.

Visual Analysis

The most widely used method for evaluating single-case time-series intervention data is visual analysis (a.k.a., visual or 'ocular' inspection). Visual inspection requires the simultaneous evaluation of a single-case graph for trend, variability, mean shifts, and immediacy of intervention effects to make a subjective decision as to whether an intervention had the desired impact on performance, as evidenced by differences in the data plotted between the baseline and intervention phases. A typical SCD data display utilized for visual analysis is provided in Figure 1. In best case scenarios, visual inspection is a rather simple process because the graphed data can immediately convey to the evaluator that the intervention was effective (Parsonson & Baer, 1992). In our experience in real life counseling, these situations are relatively rare.

Visual analysis can be utilized to assess the effects of interventions on any outcome of interest for which data can be plotted graphically. For example, if a student is referred to a school counselor due to concerns related to class attendance, the counselor can compile weekly attendance data from school records to create a baseline and, once an intervention has been implemented (perhaps for a month), to gauge its effectiveness. If attendance improves then the intervention is determined to have been successful; if not, modifications to the intervention or an alternative intervention may be warranted. Determining how much improvement is needed to demonstrate the effects of an intervention is a subjective process that involves professional judgment. For example, using the attendance vignette above, for a student who is chronically absent, incremental attendance gains as low as 5% or 10% may be considered successful progress, whereas other referral problems (e.g., citations for fighting) may require near or almost complete elimination of the target behavior for an intervention to be deemed effective.

The strengths of the visual analysis method are that it allows practitioners to determine the effects of their interventions in an efficient manner and does not require overly technical sophistication to be effectively utilized, although training is required to differentiate aspects of treatment outcome and, as such, school counseling as a field perhaps should make training in visual analysis a priority. The limitations of visual analysis are that it may be insensitive to small changes that are observed in outcome data and that it is not robust in the face of significant variability or overlap in the data. For example, if baseline data are highly variable, it may be difficult for an evaluator to determine if an intervention was effective if only small improvements in the behavior were observed. It should also be noted that practitioners must consider the way in which data are graphically presented as the graphing method may convey an effect when in fact little or no change occurred across phases. Finally, as all elements of a visual array (e.g., trend, variability, etc.) have to be inspected simultaneously, there are no "golden rules" for what to do when one or more of these elements is viewed as discrepant from the others. For example, what is the practitioner to make of large positive effects in the presence of a baseline trend? Although guidelines for synthesizing visual analysis features in these circumstances are noticeably absent from the professional literature (Kratochwill, Levin, Horner, Swoboda, 2014), it has long been suggested (e.g., Baer, 1977; Fisher, Kelley, & Lomas, 2003) that the practical value of an intervention may be questionable if positive effects are not immediately apparent when initially inspecting graphs.

While several tutorials for constructing the graphs needed for visual analysis are available, (e.g., Dixon et al., 2009; Zaslofsky & Volpe. 2010), these reviews have been limited to illustrating basic functions in older versions of Microsoft Excel. This is problematic given the fact that the applications for visual and statistical analyses within the Excel platform are somewhat limited (Bulté & Onghena, 2013). In the time that has elapsed since the last of the aforementioned Excel tutorials was published, there have been tremendous advances in the development of software for data analysis. Most germane to the present discussion is the recent development of a series of software packages for the evaluation of single-case data in the R programming environment (R Development Core Team, 2016). One of these packages (SCVA) is particularly noteworthy as it was developed specifically to provide users with more robust tools for conducting visual analysis (see Bulté & Onghena, 2012; McGill, 2016 for demonstrations). Given the persistent concerns (e.g., Ninci et al., 2015) that have been raised about the consistency and quality of judgements resulting from visual inspection of data in applied practice, the development of more advanced software for the visual analysis of single-case data is a welcome addition to the practitioner's toolkit.

Visual Analysis Rating

Busse (2005) proposed a more structured method for visual analysis of data termed *Visual Analysis Rating* (VAR). VAR incorporates the logic of GAS in that practitioners utilize a rating scale format to visually evaluate graphed data. Raters utilize the same -2 to +2 GAS format with -2 indicating a strong negative effect, 0 indicating no change, and +2 indicating a strong positive intervention effect. In completing a VAR rating, practitioners must take into consideration whether the change between baseline and treatment is in the desired direction, generally immediate, discernible, maintained, and whether there is a relatively large and stable average level of change. The strengths of the VAR method are that it is simple to use, it provides a more systematic method of analysis than simple visual inspection, and it provides for rating progress toward goal attainment. The major limitation is that the method is in need of empirical validation and, as such, should be used with caution (See McGill and Busse, 2014 for a school-based application of VAR and other methods presented in this article toward evaluating a small group reading intervention).

Trend Analysis

The purpose of trend analysis is to examine the degree to which data are moving in a desired direction and toward a desired goal. The most commonly used method of trend analysis is ordinary least squares (OLS). OLS is a statistical regression method that creates a line of best fit that provides an estimate of direction (trend) and rate of improvement as indicated by the slope (ascent or descent) of the best fit line. OLS trend lines are easily created with basic graphic software and available online programs.

In trend analysis the regression line (a straight line of best fit for the data) is drawn for the baseline and intervention data, which allows for analysis of the slope of the line. When the slope of the trend line for the baseline phase is compared to the slope of the intervention data, it becomes possible to compare a student's actual progress to their trend before and after the intervention (Riley-Tillman & Burns, 2009). In comparing trends, the same principles apply as with visual analysis; effective interventions will be represented with intervention data that trend in the desired direction and are distinct from the trend produced from pre-intervention baseline data.

There are no available guidelines for determining when an observed change in slope from one phase to the other is significant. School counselors must use their professional judgment and base determinations of effectiveness on individual cases. Weekly or daily growth rates can be used to evaluate treatment progress and to assist in evaluating the feasibility of selected treatment goals (e.g., if the student would be required to increase homework rate in all classes from a baseline of 50% to 100% within two weeks, the goal may be shown to be unrealistic as indicated by the slope and trend).

The strengths of trend analysis are that it: a) is relatively easy to use with computer graphing programs; b) allows us to account for trends that were in place before the intervention was implemented; c) can be implemented with single-phase or multiple-phase intervention models; and d) it is a powerful statistical method. The limitations of this approach are that: a) regression lines are less robust with highly variable data (Brossart, Parker, & Castillo, 2011); b) OLS requires a relatively large number of data points of 25 or more for a robust outcome; and c) there are no guidelines for quantifying the magnitude of differences between phase trends. Despite these limitations, OLS can be used with caution with as few as three data points per baseline and treatment phases (although a minimum of five is recommended; Daly et al., 2010).

Percentage of Non-overlapping Data

Percentage of Non-overlapping Data (PND) is a popular method that is utilized to assess the effects of changes between baseline and treatment phases in time-series data. To calculate PND, the number of data points in the intervention phase that exceed the highest (or lowest if intervention is designed to decrease behavior) point in the baseline phase are divided by the total number of data points in the intervention phase. The resulting quotient is interpreted as a percentage. Using conventional guidelines (e.g., Scruggs, Mastropieri, & Casto, 1987), A PND greater than or equal to 80 is indicative of a strong intervention effect, 60-79 is moderate, and percentages below 60 are interpreted as evidence of no significant intervention effect.

PND is easy to calculate and to interpret and complements visual analysis methods. However, the PND statistic is significantly limited in the presence of outliers and floor effects. For example, if the goal for a school counselor is to implement an intervention that results in increases in a desired behavior, and the rate of the behavior is at or close to zero during baseline, any improvement will almost certainly result in a PND of 100. Therefore, it is recommended that alternative evaluative methods be utilized when such floor or ceiling effects are present.

Over the past decade more sophisticated methods for estimating overlap have been proposed, including the Percentage of All Non-overlapping Data (PAND; Parker, Hagan-Burke & Vannest, 2007) and Non-overlap of All Pairs (NAP; Parker & Vannest, 2009). Whereas these methods may be more statistically sophisticated, we believe that the original method has the most utility for practitioners because of its simplicity of calculation.

Reliable Change Index

Reliable Change Index (RCI; Jacobson & Truax, 1991) is a method for evaluating the difference between baseline and treatment scores from standardized assessment measures. It can be utilized as a measure of effect by school counselors who administer rating scale measures for progress monitoring to assess the effectiveness of a variety of interventions ranging from individual counseling to direct academic interventions (e.g., study skills). RCI is calculated by subtracting a pre-test score from the post-test score and dividing by the standard error of the outcome measurement. Standard error is calculated from the reliability of the measure, e.g., internal consistency, test-retest reliability, therefore RCI may be best suited to evaluations that involve standardized measures (e.g., BESS; Kamphaus & Reynolds, 2007, BPM; Achenbach, et al., 2011), as opposed to counselor-created rating scales or other forms of school-based data as reliability coefficients for these measures are commonly unknown.

The robustness of treatment gains is moderated by the reliability of the outcome measure that is utilized. That is, greater change is required to demonstrate effectiveness via the RCI if an unreliable measure is used. RCI values that exceed 1.96 are considered to be statistically significant. This critical value corresponds to a two standard deviation improvement from the baseline score. Busse and colleagues (Busse, 2005; Busse, Elliott, & Kratochwill, 2010; Busse & Yi, 2013) suggested that RCI can be interpreted as an effect size and proposed the following guidelines: RCIs >1.8 are indicative of a strong, positive change, RCIs from .7 to 1.7 are indicative of moderate change, -.6 to .6 are indicative of no behavioral change, -.7 to -1.7 are indicative of a moderate negative effect, and RCIs ≤ -1.8 indicate that a behavior problem has significantly worsened.

The strengths of the RCI are that it provides counselors with a quantitative measure of effect from rating scale data that is fairly easy to calculate. The standard error information required to compute the RCI is easily obtained from most test manuals, and online RCI calculators are available. A limitation of the RCI method is that it is dependent on the utilization of outcome measures that demonstrate adequate reliability for educational and psychological decision making. A related limitation is that many school counselors may not have had training at the clinical level for the appropriate use of rating scales, such as psychometrics and interpretation of standard scores. From our perspective, the field would benefit from a concerted effort to enhance training to include standardized assessment measures, particularly as school counseling moves toward data-based outcome decision-making. As a case in point, the training program at Chapman University offers a course for professional clinical counseling that specifically focuses on just such assessment issues.

Case Study Application

In this section, we provide a case with fictional data to illustrate the use of selected outcome assessment methods presented in this article for evaluating school counseling interventions for academic and study skills problems.

Case Study: Jose

Jose is a ninth grade student who was referred to his school counselor due to difficulties he demonstrated in math class. His math teacher assessed student performance using a formative assessment model via weekly quizzes to review material that was covered in class and through homework assignments. Jose achieved failing grades on the weekly tests over the first few weeks of the school year, which raised concerns from his teacher. Jose's school counselor reviewed his cumulative file and found that Jose earned passing marks in math when he was in middle school. When asked about the discrepancy, Jose indicated that he was not used to taking so many tests and was having difficulty preparing for them. More in-depth interviews with Jose and his math instructor revealed that Jose rarely took notes in class during lectures and that he failed to attend weekly review sessions offered by the instructor after school. Additionally, Jose reported that his preparation was usually limited to "flipping through the textbook" the night before each exam and that he failed to utilize any systematic study strategies (e.g., notecards, review questions, group study with peers). To help Jose improve his math performance, his school counselor suggested several strategies: a) begin taking daily notes in his math class in a separate notebook that could be checked by his teacher or school counselor; b) attend each after school review session with his instructor; c) create mock questions on notecards to review in preparation for each exam; and d) establish a study group with several of his peers to supplement his own individual preparation. To help facilitate fidelity with these suggestions, the school counselor began weekly 15-20 minute check-ins at the beginning of the week to review his test performance from the week prior, review his notes and to work with him on test taking strategies. A GAS matrix was created for Jose based upon his use of the recommended test taking skills and review strategies. Outcome data were assessed with his math exam performance.

To supplement these indicators, the school counselor also administered the *School Motivation and Strategies Inventory* (SMALSI; Stroud & Reynolds, 2006a) to Jose as a pre- and posttest measure of intervention effectiveness. The SMALSI is a self-report rating scale that measures 10 areas related to success in learning, learning-related motivation, and study habits. The SMALSI teen form (ages 13-18) consists of 170 items that yield 10 standardized scores expressed as *T*- scores ($M = 50$, $SD = 10$) and takes approximately 20-30 minutes to complete. Given the intervention target, the school counselor focused particularly on the results obtained on the Study Strategies scale for pre-posttest analyses. According to the SMALSI *Manual* (Stroud & Reynolds, 2006b), the median reliability coefficient for the Study Strategies scale is .86.

Descriptive evaluation of Jose's outcome data. Graphical data of the effects of Jose's weekly intervention with the school counselor are shown in Figure 1. The outcome data include percent correct on weekly math quizzes and corresponding GAS ratings (Tables 1 and 2) of his use of test taking strategies and selected study skills. The school counselor monitored Jose's progress each week and evaluated the outcome after several weeks of intervention.

Table 1
Jose's Goal Attainment Scaling (GAS) Matrix

GAS Rating	Goal Description
+2	Self-study, notes, and attend review session.
+1	Self-study, notes, problem notecards.
0	Self-study with notes
-1	Self-study with no notes
-2	No self-study

Table 2
Goal Attainment Scaling (GAS) Outcomes

Week	GAS Rating
4	0
5	+1
6	+1
7	+2
8	+1
9	+2
10	+1

As shown in Figure 1, visual analysis of the effect on Jose's percent correct on math quizzes indicates a change from baseline to treatment with a relatively large level change. The trend analysis indicates that the baseline data were trending slightly in the same direction as the treatment phase, therefore one cannot conclude that the observed increase in test performance was due entirely to the intervention. The PND for these data is 88% (a total of 8 data points were collected, and 7 of these were above the baseline levels; 7/8 = .88), indicating a strong positive effect. A VAR of +1 was assigned because, although there was clearly improvement, there was not enough to describe these results as a "strong positive effect." Weekly GAS ratings yielded an average rating of +1.14 for the treatment phase, indicating a moderate positive effect.

Jose obtained a pre-intervention Study Strategies score of 38 on the SMALSI which falls within the low *average* range at the 13th percentile when compared to his same aged peers. When administered the same measure at the conclusion of seven weeks of check-ins with the school counselor, his Study Strategies score improved to 51 which fell within the *average* range (53rd percentile). The pre-postest difference of 13 points resulted in an RCI value of +2.46, well above the threshold (1.96) for clinical significance recommended by Jacobson and Truax (1991). Taken together, these data led the teacher and school counselor to decide that Jose was evidencing improvement and to continue with the weekly check in and to do a follow-up evaluation in one month to determine if Jose's performance continues to improve or if further intervention is warranted.

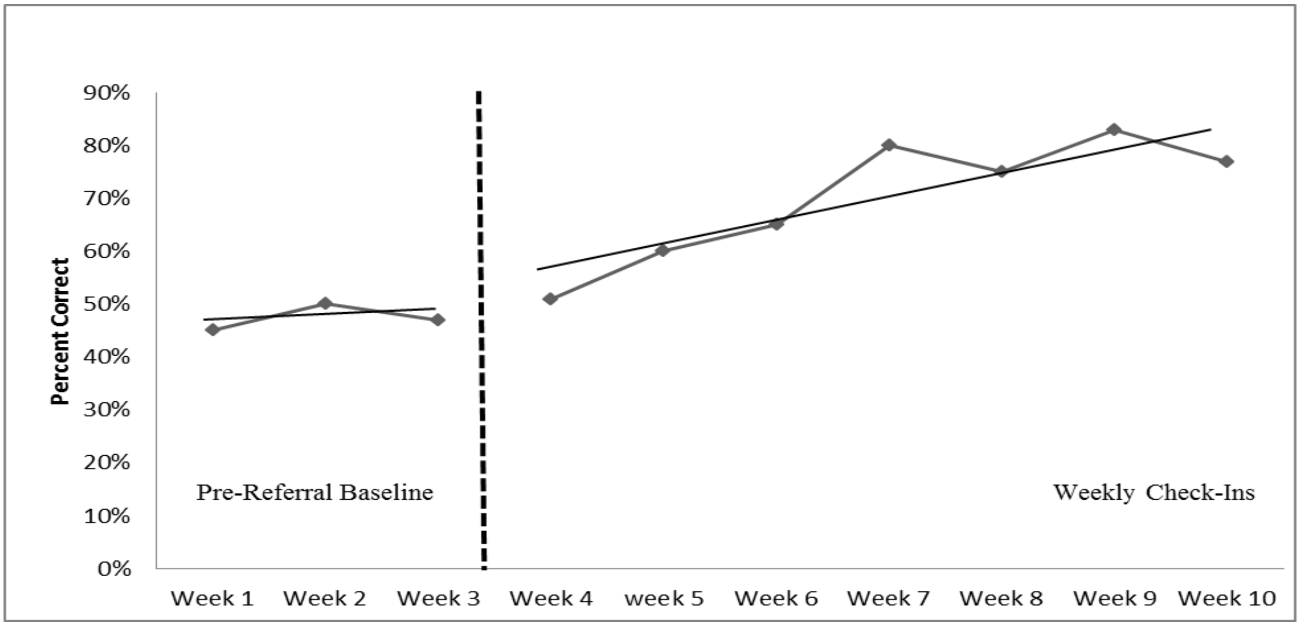

Figure 1. Single-Case Graphic Array of Jose's Weekly Quiz Data.

Although fictional, the case study illustrates how multiple evaluative indicators can be used in concert to systemically appraise the integrity of applied interventions in school counseling practice. As illustrated above, an SCD framework was utilized to examine the effect of the school counselor-led intervention on Jose's math performance, GAS was used as a supplement to provide outcome data of a different dimension—Jose's use of suggested study and test taking strategies, and the RCI metric was utilized to evaluate the significance of pre-post ratings on an outcome measure related to the target behavior(s). Whereas, all outcome indicators were in relative alignment indicating a moderately positive intervention effect, we suggest that the value of utilizing multiple indicators is accentuated when there is a lack of convergence. For example, if Jose's exam performance improved over the course of the intervention and GAS ratings were negative it would support that Jose's improvement was not due to the use of the strategies recommended by the school counselor. While not sufficient in and of itself to question the relative value of the intervention that was recommended, it would be important information for a practitioner to consider when faced with a similar referral problem from a future client.

Discussion

The methods we described in this article have potential practical implications. Most importantly, school counselors can utilize these methods to monitor children's improvement. Secondly, the methods can be used to demonstrate the effectiveness of school counseling services and/or interventions to relevant stakeholders within educational settings. Such exhibitions are consistent with established evidence-based practices and may help to ensure that school counseling practices align with calls for greater accountability within educational settings and as potential safeguard against the reification of clinical practices that lack requisite evidentiary support (Kratochwill, 2007; Lilienfeld, Amirrati, & David, 2012; Norcross & Wampold, 2011).

We have been deliberate in our selection of effect sizes to highlight in this article. We chose methods that we believe may be more readily understood and utilized in school-based practice. In particular, goal attainment scaling provides a readily accessible method for monitoring interventions that we believe should be a part of every school counselor's toolkit. Although some researchers have called for the use of more sophisticated methods (Kratochwill et al., 2010), we believe that the typical practitioner will not have the time, technical training, or resources to successfully engage in their use. The methods we presented are easily calculated either by hand or with available online resources. It is, of course, incumbent on the practitioner to learn more about what may be useful practice and to apply those methods with which they have had appropriate training (Kratochwill, 2010; Kratochwill & Levin, 2010).

Whereas trend analysis and RCI are useful tools, these methods may require additional training on the part of school counselors to use as intended. We included trend analysis due to its widespread use in response to intervention (RTI) models and, as such, although school counselors may not be comfortable using the method, they should at the very least be aware of its applications as they interact with those (e.g., school psychologists) who may. In sum, we suggest that indicators that are easy to calculate and interpret such as visual analysis and PND will have the greatest utility for school counselors who evaluate interventions in applied practice.

Evaluation of Multiple Outcome Indicators

In keeping with best practice, we have outlined an evaluative model in which school counselors can appraise intervention outcomes through simultaneous evaluation of multiple indicators. However, concerns have been raised about how to best synthesize multiple outcome indicators in order to make reliable and valid judgments of intervention effects (Parker et al., 2005). Although more objective methods for converging evidence from multiple sources (e.g., Busse, Elliott, & Kratochwill, 2010; Busse, McGill, & Kennedy, 2015) have been proposed within the professional literature, it remains to be seen whether these methods provide practitioners with the necessary protections against the sources of error (i.e., Type 1 and Type II error) that are endemic in clinical decision-making (Watkins, 2009).

Conclusion and Implications for Professional Practice

The purpose of this article is to present methods for school counselors to evaluate the effectiveness of their interventions. According to Meier (2008) it is important that we monitor treatment outcomes in counseling in a systematic format. Otherwise, how do we know that even well-designed counseling interventions have been evaluated appropriately? Our primary goal was to outline methods that may be useful for school counselors and to provide a format for interpretation and decision making processes regarding individual academic and emotional/behavioral interventions. We hope this framework will be of benefit as the field continues to move forward toward a data-based service delivery model.

References

Achenbach, T., McConaughy, S., Ivanova, M., & Rescorla, L. (2011). *Manual for the ASEBA Brief Problem Monitor (PBM).* Burlington, VT: University of Vermont

American School Counselor Association (2012). *The ASCA national model: A framework for school counseling programs* (3rd ed.). Alexandria, VA: Author.

Baer, D. M. (1977). "Perhaps it would be better not to know everything." *Journal of Applied Behavior Analysis, 10,* 167-172. doi: 10.1901/jaba.1977.10-167

Barlow, D. H., Nock, M. K., & Hersen, M. (2008). *Single-case experimental designs: Strategies for studying behavior change* (3rd ed.). Boston: Allyn & Bacon.

Brady, J., Busse, R. T., & Lopez, C. (2014). Monitoring school consultation intervention outcomes for data-based decision making: An application of the goal attainment scaling method. *Counseling Outcome Research and Evaluation, 5,* 63-69.

Brossart, D. F., Parker, R. I., & Castillow, L. G. (2011). Robust regression for single-case data analysis: How can it help? *Behavior Research Methods, 43,* 710-719.

Brown, R., Steege, M., & Bickford, R. (2014). Best practices in evaluating the effectiveness of interventions using single-case methods. In A. Thomas & P. Harrison (Eds.), *Best practices in school psychology* (4th ed., Vol. 4, pp. 371-382). Bethesda, MD: National Association of School Psychologists.

Bulté, I. & Onghena, P. (2012). When the truth hits you between the eyes: A software tool for the visual analysis of single-case experimental data. *Methodology, 8,* 104-114. doi: 10.1027/1614-2241/a000042

Bulté, I. & Onghena, P. (2013). The single-case data analysis package: Analysing single-case experiments with R software. *Journal of Modern Applied Statistical Methods, 12*(2), 450-478. Retrieved from http://digitalcommons.wayne.edu/jmasm/vol12/iss2/28

Busse, R. T. (2005, March). *Methods for assessing response to intervention*. Workshop presented at the annual convention of the National Association of School Psychologists, Atlanta, GA.

Busse, R. T., Elliott, S. N., & Kratochwill, T. R. (2010). Convergent evidence scaling for multiple assessment indicators: Conceptual issues, applications, and technical challenges. *Journal of Applied School Psychology, 26,* 149-161.

Busse, R. T., McGill, R. J., & Kennedy, K. S. (2015). Methods for assessing single-case school based intervention outcomes. *Contemporary School Psychology, 19,* 136-144. doi: 10.1007/s40688-014-0025-7

Busse, R. T., & Yi, M. (2013). Behavioral and academic rating scale applications within the problem-solving model. In R. Brown-Chidsey & A. Andren (Eds.), *Assessment for intervention: A problem-solving approach (2nd ed.)*(pp. 180-198). New York: Guilford.

Coffee, G., & Ray-Subramanian, C. E. (2009). Goal attainment scaling: A progress-monitoring tool for behavioral interventionists. *School Psychology Forum, 3,* 1-12.

Daly, E. J. III, Barnett, D. W., Kupzyk, S., Hofstadter, K. L., & Barkley, E. (2010). Summarizing, evaluating, and drawing inferences from intervention data. In G. Gimpel Peacock, R. A. Ervin, E. J. Daly III, & K. W. Merrell (Eds.), *Practical handbook of school psychology: Effective practices for the 21ST century* (pp. 497-512). New York: Guilford.

Dimmitt, C., Carey, J. C., & Hatch, T. (2007). *Evidence-based school counseling: Making a difference with data-driven practices*. Thousand Oaks, CA: Corwin Press.

Dixon, M. R., Jackson, J. W., Small, S. L., Horner-King, M. J., Lik, N. M., Garcia, Y., & Rosales, R. (2009). Creating single-subject design graphs in Microsoft Excel 2007. *Journal of Applied Behavior Analysis, 42,* 277-293. doi: 10.1901/jaba.2009.42-277.

Dollarhide, C. T., & Lemberger, M. E. (2006).No Child Left Behind: Implications for school counselors. *Professional School Counseling, 9,* 295–304. Retrieved from http//:www.schoolcounselor.org

Dollarhide, C. T., & Saginak, K. A. (2012). *Comprehensive school counseling programs: K-12 delivery systems in action* (2nd ed.). Upper Saddle River, NJ: Pearson.

Elliott, S. N., & Busse, R. T. (2004). Assessment and evaluation of students' behavior and intervention outcomes: The utility of rating scale methods. In R. B. Rutherford, M. M. Quinn, & S. R. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders (pp. 123-142).* New York: Guilford.

Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis, 36,* 387-406. doi: 10.1901/jaba.2003.36-387

Foster, L., Watson, S. Meeks, C., & Young, J. S. (2002). Single-subject research design for school counselors: Becoming an applied researcher. *Professional School Counseling, 6,* 146-154. Retrieved from http://search.proquest.com

Hatch, T. (2013). *The use of data in school counseling: Hatching results for students, programs, and the profession*. Thousand Oaks: CA. Corwin Press.

Isaacs, M. L. (2003). Data-driven decision making: The engine of accountability. *Professional School Counseling, 5,* 235-248. Retrieved from http://search.proquest.com

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12-19. doi: 10.1037/0022-006X.59.1.12

Kamphaus, R., & Reynolds, C. (2007). *Behavioral and emotional screening system manual*. Minneapolis, MN: Pearson.

Kiresuk, T. J., & Sherman, R. E. (1968). Goal attainment scaling: A general method for evaluating community mental health programs. *Community Mental Health Journal, 4,* 443-453. doi: 10.1007/BF01530764

Kratochwill, T. R. (2007). Preparing psychologists for evidence-based school practice: Lessons learned and challenges ahead. *American Psychologist, 62,* 826-843. doi: 10.1037/0003-066X.62.8.826

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods, 15,* 124-144. doi: 10.1037/a0017736

Kratochwill, T. R., Levin, J. R., Horner, R. H., & Swoboda, C. M. (2014). Visual analysis of single-case intervention research: Conceptual and methodological considerations. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 91-125). Washington, DC: American Psychological Association

Lilienfeld, S. O., Ammirati, R., & David, M. (2012). Distinguishing between science pseudoscience in school psychology: Science and scientific thinking as safeguards against human error. *Journal of School Psychology, 50,* 7-36. doi: 10.1016/j.jsp.2011.09.006

Lundervold, D. A., & Belwood, M. F. (2000). The best kept secret in counseling: Single-case (*N* = 1) experimental designs. *Journal of Counseling and Development, 78,* 92-102. doi: 10.1002/j.1556-6676.2000.tb02565.x

Manassis, K. (2012). *Problem solving in child and adolescent psychotherapy: A skills-based collaborative approach*. New York: Guilford Press.

McGill, R. J. (2016). Single-case design and evaluation in R: An introduction and tutorial for school psychologists. *International Journal of School & Educational Psychology*. Advance online publication. doi: 10.1080/21683603.2016.1173610

McGill, R. J., & Busse, R. T. (2014). An evaluation of multiple single-case outcome indicators using convergent evidence scaling. *Contemporary School Psychology, 18,* 13-23.

Meier, S. T. (2008). *Measuring change in counseling and psychotherapy*. New York: Guilford Press.

Merrell, K. (2007). *Strong kids: A social and emotional learning curriculum, grades 6-8.* Baltimore, MD: Paul H. Brookes Publishing.

Murphy, J. (2008). *Solution-focused brief therapy in schools*. Alexandria, VA: American Counseling Association.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: United States Department of Education.

Ninci, J., Vannest, K. J., Willson, V., & Zhang, N. (2015). Intrrater agreement between visual analysts of single-case data: A meta-analysis. *Behavior Modification, 39,* 510-541. doi: 10.1177/0145445515581327

Norcross, J. C., & Wampold, B. E. (2011). Evidence-based therapy relationships: Research conclusions and clinical practices. *Psychotherapy, 48,* 98-102. doi: 10.1037/a0022161

Parker, R. I., Brossart, D. F., Vannest, K. J., Long, J. R., De-Alba, R. G., Baugh, F. G., & Sullivan, J. R. (2005). Effect sizes in single case research: How large is large? *School Psychology Review, 34,* 116-132. Retrieved from http://www.nasponline.org

Parker, R. I, Hagan-Burke, S., & Vannest, K. J. (2007). Percentage of all non-overlapping data (PAND): An alternative to PND. *The Journal of Special Education, 40,* 194-204. doi: 10.1177/00224669070400040101

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*, 357-367. doi: 10.1016/j.beth.2008.10.006

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect sizes in single-case research: A review of nine nonoverlap indices. *Behavior Modification, 35,* 303-322. doi: 10.1177/0145445511399147

Parsonson, B. S., & Baer, D. M. (1992). The visual analysis of data and current research into the stimuli controlling it. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 15-40). Hillsdale, NJ: Erlbaum

Perdrix, S., Stauffer, S., Masdonati, J., Massoudi, K., & Rossier, J. (2012). Effectiveness of career counseling: A one-year follow up. *Journal of Vocational Behavior, 80,* 565-578. doi: 10.1016/j.jvb.2011.08.011

R Development Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria. R Foundation for Statistical Computing. Available from http://www.r-project.org

Reynolds, C. & Kamphaus, R. (2015*). Behavior assessment system for children, Third edition*. San Antonio, TX: Pearson.

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

Riley-Tillman, T. C., & Walcott, C. M. (2007). Using baseline logic to maximize the value of educational interventions. *School Psychology Forum: Research in practice, 1*(2), 87-97. Retrieved from http://www.nasponline.org

Roach, A.T., & Elliott, S.N. (2005). Goal attainment scaling: An efficient and effective approach to monitoring student progress. *Teaching Exceptional Children, 37*(4), 8-17.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1,* 199-223. doi: 10.1037/1082-989X.1.2.199

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and special education, 8,* 24-33. doi: 10.1177/074193258700800206

Sink, C. A. (2009). School counselors as accountability leaders: Another call for action. *Professional School Counseling, 13,* 68-74. Retrieved from http//:www.schoolcounselor.org

Stroud, K. C., & Reynolds, C. R. (2006a). *School Motivation and Learning Strategies Inventory*. Torrance, CA: Western Psychological Services.

Stroud, K. C., & Reynolds, C. R. (2006b). *School Motivation and Learning Strategies Inventory-manual*. Torrance, CA: Western Psychological Services.

Swaminathan, H., & Rogers, H. J. (2007). Statistical reform in school psychology research: A synthesis. *Psychology in the Schools, 44,* 543-549. doi: 10.1002/pits.20246

Vannest, K. J., Davis, J. L., & Parker, R. I. (2013). *Single-case research in schools: Practical guidelines for school-based professionals*. New York: Routledge.

Watkins, M. W. (2009). Errors in diagnostic decision making and clinical judgment. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (4th ed., pp. 210-229). Hoboken, NJ: John Wiley.

Whiston, S. C., & Sexton, T. L. (1998). A review of school counseling outcome research: Implications for practice. *Journal of Counseling and Development, 76,* 412-426. doi: 10.1002/j.1556-6676.1998.tb02700.x

Young, A., & Kaffenberger, C. (2011). The beliefs and practices of school counselors who use data to implement comprehensive school counseling programs. *Professional School Counseling, 15,* 67-75. Retrieved from http://search.proquest.com

Zaslofsky, A. F., & Volpe, R. J. (2010). Graphing single case data in Microsoft Excel. *School Psychology Forum, 4*(2), 15-24. Retrieved from http://www.nasponline.org