Appreciating the complexity of family psychology research and clinical practice rooted in a systemic epistemology, multiple types of evidence are acknowledged to contribute meaningfully to evidence-based practice (EBP), including systematic case studies and single-case designs. Family psychology is replete with case studies; however, most lack the methodological rigor essential to EBP, and the methodologically rigorous single-case design is underutilized. The article recommends a shift in family psychology clinical research to include evidence-based case studies and single-case designs that are both methodologically rigorous and feasible within clinical practice. Implications for lessening the research—practice gap and informing EBP are discussed.

Keywords: case study methods, single-case design, evidence-based practice, couple and family psychology, systemic research

The specialty of family psychology is rooted in a systemic epistemology that recognizes human behavior to be a function of a complex contextual matrix of individual, relational, and environmental factors that are in dynamic reciprocity over time (Stanton, 2009). This systemic paradigm provides the framework for conceptualizing, assessing, treating, and researching processes and outcomes at the individual, relational, family, and larger organizational levels. The paradigm is a significant departure from the Cartesian scientific method that not only dominates Western thinking but also the research paradigms and practice of psychology. The tension between these paradigms is clearly reflected in the chasm between research and practice in family psychology.

Whereas family therapists have sought to develop interventions deemed effective with complex systems based on their clinical experience and to disseminate these via the clinical case study, family researchers in academic settings have been pressed to engage in increasingly scientifically rigorous group experimental designs, that is, randomized clinical trials (RCTs), to determine empirically supported treatments (ESTs). The chasm between family researchers and practitioners is particular evident in the EST debate (Black & Lebow, 2009). Academic researchers criticize clinicians for practicing treatments that lack empirical support, and clinicians criticize family researchers for reductionist thinking that cannot account for the diversity of clinical practice. Evidence-based practice in psychology (EBPP) promises a reconciliation of these diverging perspectives and methodologically rigorous case study a vehicle.

Demand for evidence-based practice (EBP) is ubiquitous in health care services as a means to improve quality, cost-effectiveness, and accountability. The American Psychological Association in 2005 endorsed EBPP, “to promote effective psychological practice and enhance public health by applying empirically supported
principles of psychology assessment, case formulation, therapeutic relationship, and intervention” (APA, 2006, p. 271). The APA defines EBPP as, “the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences” (APA, 2006, p. 273). Best available research evidence refers to scientific results related to intervention strategies, assessment, clinical problems, and patient populations obtained in both laboratory and field settings.

Efforts to determine the best available evidence can be seen within the specialty of family psychology. The Task Force on Evidence-Based Couple and Family Therapy has determined guidelines for evaluating evidence-based couple and family practice, and a number of evidence-based family psychology treatments have been identified (http://www.division43.apa.org/science/evidence-based-in-family-psychology). The Task Force further articulated three levels of evidence that provide an increasing index of confidence for the practitioner that an intervention/treatment model works (Sexton et al., 2011): Pre-evidence interventions lack evidence; Level I treatments demonstrate limited or weak evidence; Level II interventions have methodologically sound preliminary evidence but lack replication; Level III, the “gold standard” requires a minimum of two outcome studies that produce gains in the target population/problem beyond normal development. Level III evidence is further categorized to demonstrate (a) relative efficacy, (b) contextual efficacy, and (c) theoretically expected mechanisms of change. The distinction between EST and EBPP is relevant (APA, 2006). Whereas ESTs start with the treatment and ask whether it works for a certain disorder or problem under specified circumstances, EBPP starts with the patient and asks what research evidence will assist the psychologist in achieving the best outcome.

Appreciating the complexity of clinical practice, the APA (2006) endorsed multiple types of research evidence as contributing meaningfully to EBPP including meta-analysis, randomized clinical trials (RCTs), effectiveness studies, process-outcome studies, epidemiological studies, systematic case studies and single-case experimental designs. Although Sexton et al. (2011) indicate that numerous types of evidence may be considered, evidence-based treatments in family psychology have largely been informed by RCTs, effectiveness research, and process-based research studies (http://www.division43.apa.org/science/evidence-based-in-family-psychology). Whereas both single-subject case designs and case study methods have been proposed as promising methods for family therapy research (e.g., Dickey, 1996; Moon & Trepper, 1996), these are used infrequently to inform EBPP in family psychology. Systematic case studies, including single-case designs, are research methods that are feasible for family practitioners, and when conducted with methodological rigor, have the potential to add meaningful evidence to support EBPP in family psychology and reduce the gap between research and practice.

The purpose of this article, therefore, is to highlight the methodological advances and applications in these research designs that enhance their applicability to family psychology. For the purposes of the article, the following definitions and distinctions are made. Single-case designs refer to experimental designs that monitor the effect of change in the environment, that is, the effect of a particular intervention on a specified target, an individual, couple, or family. As with group experimental designs, a goal of the single case design is control, that is, the assurance that change can accurately be attributed to the intervention. Single-case design may be viewed as the downward extension of group experimental methodology to the single unit or small N level. In contrast, the clinical case study may be defined as a detailed analysis of individual, couples or family therapy that includes verbatim clinical case material and is instructive regarding the treatment, the problem, or population. Rather than seeking to minimize or ignore contextual variables, the case study tradition gives voice to the complexity of clinical practice. As such it is an ideal method for the complexity of systemic approaches to couple and family treatment, and it is not surprising that three generations of family therapists have published case studies (Dickey, 1996). The systematic case study is a more methodologically rigorous case study that follows a prescribed format that includes verbatim clinical case material, extends the aim of the case to study to research, and includes commonly accepted quantitative or qualitative measurement. The evidence-based case study is the integration of verbatim clinical case material with standardized measures of process and outcome evaluated at different times across treatment and with attention to
clinical significance methodology (http://www.apa.org/pubs/journals/pst/evidence-based-case-study.aspx). From this perspective, the single-case design is considered a type of evidence-based case study. The methodological rigor that is often lacking in the traditional clinical case study is enhanced in both the systematic case study and the evidence-based case study permitting these, individually and in aggregate, to contribute to evidence-based family psychology practice.

The aim of this article is educative given the limited use of evidence-based case study in family psychology. We first provide a brief history of single-case clinical research to place the limited use of case study methodology in context. The evidence-based case study and single-case design are next discussed with particular attention paid to methodological advances and application to family psychology. We endeavor to demonstrate that objective scientific data can be integrated with clinical intuitive experience and with a systemic epistemological paradigm to advance broadly the specialty of family psychology.

A Brief History of Single-Case Clinical Research

The study of the single case in psychology can be traced to the late 19th century with the work of many of the founders of experimental psychology. Fechner’s research on just noticeable differences, Wundt’s introspection, Ebbinghaus’s study of memory using nonsense-syllables, and Pavlov’s landmark studies in classical conditioning all focused on the study of single cases (Barlow, Nock, & Hersen, 2009). Dissatisfaction with the yield and the lack of experimental control in single-case research (Long & Hollin, 1995), along with the development of group statistical techniques (e.g., Fisher), led to group experimental designs becoming mainstream in psychological experimental research.

As today, the research and clinical communities were largely separate from one another in late 19th century psychology, with the case study, which lacked agreed upon guidelines, the sole methodology of clinical research (Long & Hollin, 1995). In case studies clinicians would document information about their client, the therapeutic procedures being used, and speculate about the variables responsible for successful behavior change (Kazdin, 2011). Case studies allowed for communication among colleagues about treatment successes and ultimately lead to the formation of various schools of psychotherapy. Some early case studies were well controlled, accurately described, and even used methodology similar to modern day single-case research; however, most were considered largely subjective and lacked the use of data-driven procedures, leading to criticism from the research community (Barlow et al., 2009).

The development of the scientist-practitioner model at the 1949 Boulder Conference led to a shift in focus to outcome criteria, or ways to objectively measure gains made during the therapeutic process (Barlow et al., 2009). Along with this model came an emphasis on group designs for clinical research, with researchers focusing on controlled studies that examined which techniques were effective for the average client under average circumstances (Long & Hollin, 1995). Although this marked the onset of controlled clinical research, many clinicians were unsatisfied with the use of group designs because of ethical concerns about the use of control groups and a dissatisfaction with reporting group means when much outcome data varied by individual (Barlow et al., 2009).

In the late 1960s and early 1970s, several researchers developed single-case methodological approaches that aimed to bridge research and practice by isolating mechanisms of change in the therapeutic process. These single-case methodologies emphasized the direct observation of behavior, concrete definition of variables, and allowed for the study of individual behavior change over time instead of comparing groups of individuals collectively and statically (Scruggs & Mastropieri, 1998). With the introduction of these methodologies, single-case research became more accepted in the psychological research community, as well as within the practice communities of counseling, social work, pediatrics, and education. In 1968, the Journal of Applied Behavior Analysis, devoted to the study of the single case, was first published (Richards, Taylor, Ramasamy, & Richards, 1999). Recent estimates report that at least 45 professional journals include articles that are based on single-case designs (APA, 2002).
The 1980s, as part of the broader movement to include qualitative methods in research, witnessed the beginning of the shift from use of the case study solely as a teaching tool to appreciation of the case study as a research method particularly useful for the investigation of complex organizational phenomena including psychological investigations (Yin, 2009). The qualitative methods approach to the family therapy case study was described by Moon and Trepper (1996). The time consuming nature of qualitative methods, however, discouraged their use by practitioners.

Given momentum by the EBP movement, the recent decade has seen a growth of interest in systematic and rigorous case studies. In 2007 the journal Clinical Case Studies, “the only journal devoted entirely to case studies...involving individual, couples, and family therapy” was first published (http://www.sagepub.com/journals). This journal requires a uniform format for case presentation, as well as assessment of progress. Drawing substantially from the postmodern tradition, but arguing for reconciliation between the postmodern and positivist traditions (Fishman, 2001), in 2005 Fishman and associates set up the online, open-access journal, Pragmatic Case Studies in Psychotherapy, to generate a database of systematic, peer-reviewed case studies across a variety of theoretical approaches, including couple and family therapy (http://pcsp.libraries.rutgers.edu). This journal also requires a standard framework for case reports that ensures comprehensive, detailed, and consistent case information. Published case studies are accompanied by two commentaries, and a rejoinder by the author to assure a critical perspective. The methodological rigor of case study research has been enhanced with the development of criteria for analyzing clinically significant and reliable change by Jacobson and colleagues (Jacobson, Roberts, Berns, & McGlinchey, 1999; Jacobson et al., 1984; Jacobson & Truax, 1991; McGlinchey, Atkins, & Jacobson, 2002). Currently several APA journals include sections devoted to the publication of evidence-based case studies, for example, Psychotherapy, Journal of Clinical Child and Adolescent Psychology, Couples and Family Psychology: Research and Practice.

Evidence-Based Case Study

The case study as a research strategy is defined by Yin (2009) as: “An empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident” (p. 18). Yin adds

The case study inquiry (a) copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result (b) relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result, (c) benefits from the prior development of theoretical propositions to guide data collection and analysis (Yin, 2009, p. 18).

Case study research includes both single- and multiple-case studies with the later permitting cross-case conclusions to be drawn. The unique contribution of the case study research method is the ability to represent complexity and generate pragmatic knowledge.

The clinical case study has a long tradition in medicine and psychotherapy. In medicine it has been used to provide important clinical lessons, an opportunity to learn from errors, or highlight novel treatments, unusual presentations of illness, new or rare diseases, unusual disease associations, or unexpected outcomes. In counseling and psychotherapy, the case study may: (a) document, evaluate and disseminate existing or new approaches to therapy or new applications to clinical problems; (b) contribute to the development and refinement of theory; (c) draw attention to critical issues and areas of practice; (d) extend and enhance the interpretability of group studies or meta-analyses; (e) contribute to training and public understanding (McLeod, 2010). Yin (2009) identifies at least four different applications of the case study in research: (a) to explain the presumed causal links in real-life interventions that are too complex for experimental strategies, (b) to describe an intervention and the real-life context in which it occurred, (c) to illustrate certain special topics, and (d) to enlighten those situations in which an intervention has no clear, single set of outcomes.

Carefully planned or systematic case studies can “approximate” the results of RCTs (Kazdin, 2008). Strategies that enhance the methodolog-
Clinical rigor of case studies include using psychologically reliable and valid independent or “objective” measures of change, benchmarking against established interpretive criteria, research reflexivity, and use of multiple researchers (McLeod, 2010). Two rules are essential: (a) go beyond the anecdotal report and use strategies that allow for an objective measurement of change, and (b) establish stability of the problem such that any shifts that occur once therapy has begun are likely to be due to the treatment (McLeod, 2010).

Guidelines for the evidence-based case study in psychology are readily accessible (see http://www.apa.org/publ/journals/pst/evidence-based-case-study.aspx). These guidelines recommend that evidence-based case studies at the minimum should include (a) a standardized measure of the target symptom, (b) a standardized measure of global functioning completed by the patient or a rater who is not the therapist, and (c) one process measure, for example, therapeutic alliance, evaluated on at least three separate occasions during treatment. Verbatim clinical vignettes with several patient and therapist turns highlighting key interventions and mechanisms of change should be provided. Detailed information about client characteristics, context, culture, and the therapeutic relationship are recommended for inclusion. Appropriate informed consent must be obtained. Specific outcome data should be presented using standardized mean differences (i.e., effect sizes) and use clinical significance methodology that taken together permit determination of reliable treatment outcome.

**Clinical Significance Change Methods**

Clinical significance refers to changes due to treatment that are meaningful for the client (McGlinchey et al., 2002). Kazdin (2011) defines clinically significant change as meeting any of the following criteria: (a) a return to functioning within the normative range, (b) no longer meets diagnostic criteria, (c) a large change, or (d) improvement on a measure of high social impact, for example, truancy. The clinical significant change criteria proposed by Kazdin appear commonly followed in systematic case studies. Guidelines for the evidence-based case study in psychology, however, propose use of the clinical significance analysis method proposed by Jacobson et al. (1984), subsequently summarized by Jacobson and Truax (1991), and henceforth referred to as the “JT method.” The JT method proposes a two-step criterion: (a) by the end of therapy, clients should change from a dysfunctional range to a functional range; and (b) the magnitude of the change should be statistically reliable (McGlinchey et al., 2002).

**Cutoff points.** First, there should be a cutoff point established for each client that must be crossed in moving from the dysfunctional to the functional distribution. Jacobson and Truax identified three possible cutoffs.

The first, **Cutoff A**, was defined as the point two standard deviations beyond the range of the pretherapy mean. The second, **Cutoff B**, was defined as the point two standard deviations within a recognized functional range. The third, **Cutoff C**, was a weighted midpoint between the means of a functional and dysfunctional distribution (McGlinchey et al., 2002, p. 530).

When a standardized measure is used to measure outcomes, and norms are available, **Cutoff B** or **C** is preferable to **A**, and **Cutoff C** is preferable to **B** when the functional and dysfunctional populations overlap (Jacobson et al., 1999).

**Reliable change index.** The second step in the JT method is the determination of reliable pre- to posttest change. To determine whether the magnitude of the change is statistically reliable, Jacobson et al. (1984) proposed a reliable change index (RCI). The RCI shows whether client(s) changed sufficiently such that the change exceeds measurement error. The RCI is the change in a client’s pretreatment and posttreatment score divided by the standard error of the difference for the test(s) being used. Thus, the RCI value calculation includes the initial standard deviation of the outcome measure, the measure’s established reliability, as well as the pre- and posttreatment scores of the client. If the RCI is 1.96 or greater, the difference or change in scores as a result of treatment is statistically significant (1.96 equates to the 95% confidence interval). If the RCI is less than 1.96, the difference is not significant, which indicates that, according to this criterion, the client did not display reliable change. The formula for the reliable change index, which was amended by Christensen and Mendoza (1986) is as follows: 

\[ \text{RCI} = \frac{x_2 - x_1}{s_{\text{diff}}} \]

where \( x_1 \),
represents a subject’s pretest score, $x_2$ represents that same subject’s posttest score, and $S_{\text{diff}}$ is the standard error of difference between the two test scores; it describes the spread of the distribution of change scores that would be expected if no actual change had occurred. The $S_{\text{diff}}$ can easily be computed by hand using the following formula: $S_{\text{diff}} = \sqrt{2(S_0)^2}$. The RCI can also be computed with reputable online calculators (e.g., http://www.psyctc.org/stats/rcsc1.htm) by submitting the standard deviation and reliability of the outcome measure used and the client’s prepost scores.

Based on the two criteria, the JT method classifies individuals: recovered, if they pass both criteria; improved, if they pass the RCI but not the cutoff criteria; unchanged, if they pass neither criteria; and deteriorated, if the RCI criteria is passed but the direction of change is toward dysfunction. Jacobson et al. (1999) provide an illustration of their method in which the pretreatment score on a common marital measure was 85, indicating great lack of adjustment, and the posttreatment score was 114, which is almost the normative mean. Because the posttreatment score passed the cutoff value ($A$) of 105.2, and the reliable change score was greater than 1.96, the individual was classified recovered.

The JT method for determining clinically significant change is among the most frequently used by clinical researchers; however, the preferred method of determining clinically significant change has been the subject of considerable controversy. In comparison of five methods, McGlinchey et al. (2002) concluded that all methods produced good predictive validity thus lending support to the retention of the JT method as a “null” method, given its prevalence, until rejected by an alternative method demonstrating superior performance.

**Application to Couple and Family Therapy**

An excellent example of a family treatment evidence-based case study is a report on the effectiveness of the Family Check-Up (FCU) with a family with a toddler at risk for behavior problems (Gill, Hyde, Shaw, Dishion, & Wilson, 2008). Strengths of this evidence-based case study are numerous: the intervention and methodology were clearly described; a multifaceted, multilevel, multimethod assessment was completed and data are provided; measures used were reliable and valid; outcome data were reported against clinical cut-offs consistent with Kazdin’s (2011) recommendations for determining clinically significant change; multiple clinical vignettes illustrated the impact of key interventions on the change process; extended follow-up data were collected; the therapy and outcomes were discussed in reference to existing theory and research. With consideration to systemic assessment recommendations (e.g., Carlson, Krumholz, & Snyder, in press), data were collected at the individual, parent-child, and marital relationship levels and included both self-report and observational methods. The only limitation noted in this evidence-based case study was reliance upon movement from the clinical to nonclinical range as evidence of successful clinical outcomes without consideration of the magnitude of the change that would be provided with the calculation of a reliability change index or similar method. There also was no indication that a therapy process measure as a moderator was included in the evaluation.

The FCU case study (Gill et al., 2008) illustrates the possibilities as well as some of the challenges presented with the evidence-based case study evaluation of couple or family treatment. Whereas the amount of data that are recommended to collect in the systemic assessment of a couple or family (see Carlson et al., in press) can seem overwhelming, this case study demonstrates how such data can be clearly presented. Systemic data, however, also present challenges that were not addressed by the case study researchers. For example, although most outcome measures in the FCU case study improved from the clinical or at-risk range in pretreatment to the normative range posttreatment and through follow-up, some measures showed deterioration or no change. There do not currently exist clear guidelines as to how to interpret whether or not a treatment is efficacious when some, but not all, measures show reliable clinical change, which is likely to be the case more often than not (Kazdin, 2008). This problem may be compounded with multifaceted data in multimember systems such as the couple or family. Of course, this and other research
challenges deserve empirical study and are raised only to illustrate the methodological research potential of the evidence-based case study.

**Single-Case Research**

One of the most important approaches within the case study research tradition has been the single-case design (also known as the single-subject, n = 1, n-of-1, or small-N design) (McLeod, 2010). Broadly speaking, single-case designs are employed to monitor one person or unit’s response to a change in the environment (Barlow et al., 2009). Traditionally, single-case design methodology, based in behavior theory, relied exclusively on behavior change and focused on a change in the individual client or identified patient (IP). However, the usefulness of single-case designs to family therapy research has been strongly argued (e.g., Cross, 1984; Dickey, 2006; Stern & Reid, 1999).

Single-case designs address the efficacy question of “is this therapy effective?” Single-case research is based on a number of methodological principles: (a) reliable and valid measurement of outcome variables, (b) accurate description of the intervention, (c) time-series analysis of patterns of change, and (d) the logic of replication (McLeod, 2010). Single-case research involves (a) a design to follow in systematically gathering evidence, (b) visual analysis of the data, and (c) more recently, determination of effect size. A brief description of each of these key components to single-case research is provided below.

**Design Types**

The six most common single-case designs are: AB-designs, Reversal designs, Extended ABA designs, Alternating Treatments designs, Multiple Baseline designs, and Changing Criterion designs (Richards et al., 1999).

**ABA designs.** During an ABA design (also commonly called a Withdrawal design) the typical AB design is completed, followed by the initial baseline phase being reinstated (Long & Hollin, 1995). Intervention effectiveness is determined by examining the degree to which an individual’s behavior returns to the original baseline level when the intervention is removed (Morgan & Morgan, 2009). To further increase experimental control, additional A and B phases can be introduced past the additional ABA phases, resulting in an extended ABA design (e.g., ABABAB).

**Multiple baseline designs.** Multiple baseline designs involve the combination of several AB designs (with A representing a baseline phase and B representing an intervention phase), either across multiple participants, multiple behaviors within a single participant, or multiple settings within a single participant (Morgan & Morgan, 2009). The unique component of multiple baseline designs is that each AB design has a baseline of varying lengths. Intervention effectiveness is determined if there is a desired change from the baseline to intervention phase for all participants, regardless of length of baseline condition. Multiple baseline designs can be useful when it may be undesirable or unethical to withdraw an intervention already in place and to avoid potential intervention carry-over effects (i.e., the risk that the effects of one phase will carry over to another phase). Both Cross (1984) and Dickey (1996) propose the use of multiple baseline designs in family therapy research where each family member or various outcomes are represented by the different baselines.

**Alternating treatments designs.** During an alternating treatments design, two or more conditions are alternated randomly across time (e.g., ABBABABAA or BACCBABC; Morgan & Morgan, 2009). As with the above designs, alternating treatments designs are useful when comparing the effects of an intervention and baseline condition, but can also be used to compare the effects of two or more interventions. Intervention effectiveness is determined by the degree to which behavior changes consistently across phases.

**Changing criterion designs.** A changing criterion design involves evaluating the effects of an intervention on the systematic increase or decrease of the target behavior
Visual Analysis Techniques for Single-Case Designs

Understanding the data visually provides significant advantages, particularly in work on systemic processes in therapy. When single-case design data are presented graphically, several characteristics of the data should be examined and combined in order to determine if intervention effects are present including level, latency of change, variability, and trend.

One should compare the level, or the mean or median, of the data during the baseline phase to that of the intervention phase (Morgan & Morgan, 2009). If there is a sharp difference between the levels of the baseline and intervention phases, one can make preliminary judgments that the intervention was effective. Latency of change refers to how quickly behavior changes after a change in phase (e.g., from baseline to intervention or from intervention B to intervention C). Although it is not necessary for behavior to change immediately with the onset of intervention, the more quickly a change in behavior occurs, the more confident one can be that the behavior change is due to a change in treatment phase (Riley-Tillman & Burns, 2009).

Variability refers to the amount of variation in range and/or to consistency in a set of data (Riley-Tillman & Burns, 2009). Variability within a phase can be examined visually and is sometimes expressed by giving a high-low range of scores. Variability can be examined within and between phases. If behavior is highly variable within an intervention phase there is a strong likelihood that the behavior is not related to the intervention, or there is the possibility that the behavior is neither properly defined nor being measured inaccurately (Morgan & Morgan, 2009).

Finally, trend is the tendency for the data to show systematic increases or decreases over time (Kazdin, 2011). If there is an increase in the trend of the data from baseline to intervention, one can tentatively conclude that positive results of the intervention are being displayed. However, it is important to be certain that a trend in the intervention phase is not merely a continuation of a positive trend in the baseline phase (Morgan & Morgan, 2009).

Effect size analysis. Although visual analysis can be useful, researchers have objected to using it as the sole method of single-case design data analysis due to concerns about subjectivity and modest interrater reliability (Kazdin, 2011). Alternative methods of single-case design analysis have been suggested, including calculating effect sizes. The APA (2010) suggests that all manuscripts submitted for publication include effect size calculations to aid in interpretation of intervention outcome. Effect sizes used with single-case designs (a) improve measurement precision when results are not large and obvious; (b) provide a means of comparing intervention success across single-case studies, either at a local level or in a more elaborate meta-analysis; and (c) provide an objective summary of results when visual judgments do not agree (Parker & Hagan-Burke, 2007).

Although no consensus exists in the literature regarding the best effect size calculation method, the regression-based Allison-MT (Allison & Gorman, 1993), and percentage of nonoverlap-based Tau-U (Parker, Vannest, Davis, & Sauber, 2011) methods both show promise. These effect size methods each provide an estimate of the magnitude of the effect of an intervention while controlling for trend in the baseline phase (which can artificially inflate scores). The Allison-MT method has been compared to a variety of other single-case effect size calculation techniques and has been found to be less affected by autocorrelated data than other techniques (Parker & Brossart, 2003) and most similar to results obtained from visual methods of single-case analysis (Brossart, Parker, Olson, & Mahadevan, 2006). The Tau-U method is a new technique that combines nonoverlap between phases with trend from within the intervention phase, while controlling for trend in the baseline phase (Parker et al., 2011). Although little research exists that examines the Tau-U method, it has been shown to hold up well in the
presence of autocorrelation (Parker et al., 2011). In a preliminary study that compared effect size scores obtained from the Tau-U and Allinson-MT techniques on a sample of 66 previously published single-case data sets, Tau-U appeared to be less affected by autocorrelated data, but no significant difference was present between effect size scores after lag-1 autocorrelation was removed from original data (Ross & Begeny, 2012). Future research is needed to examine effect size methods for single-case research in a variety of settings, including family therapy, to develop guidelines for determining clinically significant change.

Application to Couple and Family Therapy

Dickey (1996) provided a comprehensive and compelling argument for the use of single-case design methods in family therapy, as well as guidance as to the research questions, sampling and selection, data collection, data analyses, reporting, and alternative designs appropriate to the evaluation of family therapy. Dickey further identified numerous advantages that single-case methods provide (over traditional case studies): (a) they employ checks for validity that permit the clinician-researcher to be relatively sure that obtained results are due to treatment and not to investigator subjectivity, (b) they are relatively easy and inexpensive to undertake, (c) new techniques can be developed and tested quickly, (d) objective feedback on performance can have a beneficial impact on clients, (e) treatment must be well-specified and employable by other clinicians, (f) theories regarding reciprocal influence and second-order change can be tested with designs that incorporate baselines for each family member, and finally (g) the ability to document treatment effects is consistent with ethical and accountable professional practice.

In an application of single-case design methods to family treatment, Stern and Reid (1999) measured the complex relationship between family system change and IP symptom/problem change using a multifaceted measurement strategy (termed multitracking to highlight the multiple measurement levels) with a simple time series AB design with follow-up (C) and booster (D). The multitracking design involved collecting data on the presenting problem, one or more family systems variables, and a therapy process variable. This measurement strategy allowed examination of not only treatment effectiveness, but also predicted relationships between specified system functioning and the presenting problem. The authors provided options for both visual data analysis and statistical data analysis. A limitation of this single-case design illustration was the failure of the authors to complete statistical analyses of their data, including effect sizes. Despite this limitation, Stern and Reid provide an excellent illustration of the compatibility between single-case design methods and systemic epistemology.

Systemic Measurement

A key component of evidence-based case study methods, including single-case designs, is reliable and valid measurement. Guidelines for these research methods include reliable and valid methods for measuring change, the use of multiple methods of measurement, continuous assessment of both the target outcome variables and the therapy process, and consideration of the unique cultural and idiosyncratic nature of the phenomena being measured. Both evidence-based case studies and single-case experimental designs employ specific methods for determining the reliability of observed change and recommend replication of the pattern of results over multiple cases.

There is considerable overlap in the recommendations for measurement in the evidence-based case study with recommendations for couple and family assessment (see Carlson et al., in press); however, it is commonly acknowledged that couple and family assessment is considerably more complex than assessment of the individual patient. Widespread acceptance of the systems theory premise that complex systems are organized hierarchically, has lead to consensus among family scholars that a comprehensive assessment includes measurement at multiple system levels, and research has confirmed that different levels of family functioning, although clearly related, contribute unique information about the family (Hayden et al., 1998). Although this adds complexity to the evidence-based case study, methodological innovations are encouraging.

There are numerous couple and family assessment options available to family clinical researchers, however, we wish to highlight the STIC system developed by Pinsof and colleagues (Pinsof & Chambers, 2009) because
this measurement system closely aligns with systemic epistemology and measurement guidelines for the evidence-based case study. Briefly, the STIC system consists of two client-report questionnaires that measure various dimensions of family functioning; one (INITIAL STIC) is completed at pretreatment with the shorter version (INTERSESSION STIC) completed prior to every therapy session. Three versions of measurement of the therapeutic alliance have been developed, one each for individual, couple, or family therapy. Finally a STIC feedback system provides a multisystemic and multidimensional assessment of the couple/family and monitors change over the course of therapy on measured dimensions. The ongoing assessment of the therapeutic alliance allows for the measurement of a key moderator of change. In short, the STIC system provides the practitioner with a feasible and continuous quantitative measurement system that is both consistent with systemic epistemology and the measurement demands of the evidence-based case and single-case design.

**Ethical Considerations**

Several ethical considerations are relevant in conducting and publishing evidence-based case study research. The core ethical issues include informed consent, maintaining confidentiality, and avoiding harm to participants (McLeod, 2010). Regarding informed consent, McLeod recommends that prospective and full informed consent should be obtained from the couple/family, by someone other than the therapist, at each stage of the case study inquiry cycle up to and including the final release to publish. Two options are available to protect confidentiality (APA, 2010). One option is to disguise or delete aspects of the case material that might be identifiable. The second option is to give the client the opportunity to read and comment on a draft of the case study and to stipulate the deletion or further disguising of information. The use of multiple cases or composite cases may provide an alternative means to protect privacy; however, all choices to protect confidentiality come with the risk of misleading the reader about the phenomena being described. Regarding harm to the client, it is important that the case study or single-case design process not intrude on the therapy process. Dickey (1996) noted the motivational benefit to therapy of sharing with the family visual outcome data. The practice of sharing with the family the case report, however, must be balanced against any possible negative impact (McLeod, 2010). Assuming continued informed consent and support for publication, clients should be encouraged to make a personal statement about the case report to be included in the final published version (McLeod, 2010). In addition to these patient protection considerations, the ethical challenges created with ABAB single-case designs have already been noted, and this design is not recommended for use; the ABCD design (Stern & Reid, 1999) appears to provide some of the advantages of the ABAB design without the ethical concerns. Finally, if treatment is conducted within an institutional setting, institutional review board approval should be sought.

**Discussion**

Systemic epistemology appreciates the complexity and diversity of human systems and the art, as well as the science, of facilitating their change. Whereas historically the art and the science of family therapy practice have diverged, the contemporary EBPP movement provides the opportunity for a closer relationship between the family psychology scientist and practitioner. In this article, a case has been made for use of the evidence-based case study, including single-case designs, in couple and family therapy research as a promising method for clinicians and researchers alike to contribute to EBPP.

Methodological advances have enhanced the usefulness of the case study as a research method. The publication of guidelines that standardize the case study, used in conjunction with reliable measurement and methodological advances for determining reliable clinical change, enhance both the objectivity and generalizability of the evidence-based case study while permitting in depth study of complex family system change. The addition of effect size calculations to single-case design research, the application of the multiple baseline design to the multifaceted measurement of the multimember family system, and creative applications of this design to the simultaneous monitoring of family process systems change and individual symptomatology suggest that renewed consideration of this method is appropriate.
The use of evidence-based case study and single-case design in couple and family psychology is consistent with recommendations for EBPP (APA, 2006). These methods have the potential to contribute to the Pre-Evidence, Level I, and Level II Interventions in family psychology, as well as to inform existing evidence for Level III Evidence-Based couple and family treatments (Sexton et al., 2011). As noted by Kazdin (2008), there is evidence for many treatments, but not much in the way of evidence that draws from and modifies the application of these treatments in light of clinical judgment, expertise, and the contextual considerations in practice. Fishman (2001) recommends the development of single-case databases to permit a variety of cross case and within case analyses to determine, for example, how therapeutic processes link with outcomes.

Although guidelines have been recommended for evidence-based case studies and single-case designs, there are many gaps in our knowledge of the best methods to use. These methodological issues deserve empirical study. For single-case research, effect sizes offer a useful supplement to traditional visual analysis methods; however, there is no consensus in the literature on the best way to classify effectiveness using effect sizes with single cases. Effect sizes for single cases tend to be much larger than those found in group designs; therefore using the same guidelines for classifying small, medium, and large effects are not appropriate (Parker et al., 2005). Future research should also examine using effect sizes with single-case designs other than the AB design. Although the AB design is relevant for practitioners, and avoids the ethical concerns of treatment withdrawal, this design provides the weakest evidence for treatment effect. Significant controversy also remains regarding the best method of determining clinical significance in the evidence-based case study. Extending clinical significance methodology to the evidence-based family therapy case study is largely unexamined and will hopefully spur methodological research.

Kazdin (2008) suggests three shifts in emphasis in clinical research in psychology to advance the knowledge base, improve patient care, and reduce the research-practice gap, that seem appropriate as well for family psychology. These include giving greater priority to (a) the study of mechanisms of therapeutic change, (b) the study of moderators of change that can inform clinical practice, and (c) the use of qualitative research methods to better understand the individual experience of patients and therapists, to codify treatment changes, and to do so in replicable ways. Kazdin notes that a significant barrier to the use of qualitative methods, in particular, is the lack of training provided in graduate programs such that clinicians remain uninformed as to ways they can meaningfully engage in and contribute to research outside the academic setting.

In summary, the zeitgeist in psychology is moving toward reconciliation of the research and practice divide. The demand for improved quality, cost-effectiveness, and accountability reflected in the EBPP movement, provides motivation for the development of practitioner-friendly methods of systemic measurement, training in research designs that can be used by family practitioners, and the dissemination of clinical research through evidence-based case studies that meet the rigorous methodological and reporting standards we have come to expect from group experimental research.

References


Received January 20, 2012
Revision received February 1, 2012
Accepted February 1, 2012