

CLINICAL CASE APPLICATIONS

The Effectiveness of Collaborative/Therapeutic Assessment for Psychotherapy Consultation: A Pragmatic Replicated Single-Case Study

JUSTIN D. SMITH,¹ WENDY C. EICHLER,² KAILA R. NORMAN,² AND STEVEN R. SMITH²

¹*Department of Psychology and Neuroscience, Baylor University*

²*Department of Counseling, Clinical, and School Psychology, University of California, Santa Barbara*

This pragmatic study evaluated the effectiveness of a collaborative assessment intervention as an approach to midtherapy consultation, which has yet to be empirically tested. Ten adult participants in ongoing psychotherapy with a variety of presenting concerns, primarily consisting of general mood and adjustment issues, received a brief assessment-based intervention based on Finn's (2007) Therapeutic Assessment model. Following the collection of assessment questions and the administration of a multimethod assessment battery, clients and therapists participated in a joint feedback session with the assessor. Clients were then followed as their psychotherapy continued. The results of idiographic and aggregate analytic approaches revealed significant reductions in client-reported symptomatic distress, as evidenced by a medium effect size ($d = -.50$) and a significant change in the trajectory of distress. Client reports of the process of psychotherapy revealed a significant increase in the clients' ratings of the working alliance. The findings suggest that a midtherapy consultation using collaborative/therapeutic assessment methods is beneficial but that further rigorous investigation is needed.

Psychological assessment is at something of a crossroads. Although practicing psychologists report that assessment comprises a nontrivial portion of their professional activities (Norcross, Karpiak, & Santoro, 2005), reported declines in graduate training in assessment (e.g., Belter & Piotrowski, 2001; Curry & Hanson, 2010) and altered reimbursement from managed care (e.g., Eisman et al., 2000) have undeniably affected assessment practices. However, a recent shift in psychological assessment, toward models that are collaborative and intended to enhance or produce therapeutic outcomes, has breathed new life into the debate surrounding the clinical and treatment utility of psychological assessment. The procedures and models emanating from this paradigm shift can be classified as collaborative/therapeutic assessment (C/TA; see Finn, Fischer, & Handler, 2012). In our view, which is shared by others (e.g., Finn & Tonsager, 1997; Meyer et al., 2001), psychological assessment is fundamental to formulating a comprehensive case conceptualization that directly informs the delivery of further treatment. This contention, however, remains essentially unconfirmed (Hayes, Nelson, & Jarrett, 1987; Maruish, 2004).

Psychologists are increasingly providing less psychotherapy and are being replaced in this endeavor by clinicians who do not routinely conduct psychological assessment (e.g., social workers, licensed professional counselors). With this trend

comes an opportunity for assessment psychologists to fill a niche—assessment consultation at pre and midtherapy (Finn, 2011). Professional consultation among practicing psychologists is included in the professional code of ethics of the American Psychological Association (2010) and is considered a core clinical competency (Kaslow, 2004); however, it is unknown how often assessment by another professional is sought to inform psychotherapy and to what extent psychological assessment could be helpful in this situation. Similarly, there is a lack of empirical studies examining the effectiveness of assessment in the context of professional consultation. The topic scarcely appears in the literature, with the majority related to assessors' collaborating with personnel in schools (e.g., Noell et al., 2005) and in medical settings (e.g., Pace, Chaney, Mullins, & Olson, 1995; J. D. Smith, Finn, Swain, & Handler, 2010).

Assessment consultation, particularly midtherapy, is arguably applicable to any case, but might be best suited for cases in which progress has ceased or treatment failure is looming, the origins of which are generally multifaceted (Shimokawa, Lambert, & Smart, 2010) and often linked to the therapist's case conceptualization (e.g., Clark, 1999; Lambert, 2010). C/TA models were designed to aid case conceptualization, reduce client resistance, enhance the therapeutic alliance, and improve various clinical outcomes (Finn & Tonsager, 1997; Meyer et al., 2001). Consultation with an assessment psychologist can occur at the beginning of treatment to conceptualize the client's difficulties, and perhaps clarify the reasons for the referral, in service of determining an appropriate treatment plan. Midtherapy consultation is typically sought to (a) identify a new course of action; (b) elucidate factors contributing to ceased treatment progress despite continued need for care; and (c) help the client understand an

Received February 15, 2014; Revised July 25, 2014.

Kaila R. Norman is now in the Department of Psychiatry at Yale University School of Medicine. Steven R. Smith is now in the Pacific Graduate School of Psychology at Palo Alto University.

Address correspondence to Justin D. Smith, Department of Psychology and Neuroscience, Baylor University, One Bear Place #97334, Waco, TX 76798; Email: jd_smith@baylor.edu

aspect of himself or herself, which, if introduced by the therapist, could result in a significant impasse and jeopardize the treatment (Aschieri & McCarthy, 2014). The consultative relationship between professionals is complex and presents challenges requiring careful navigation, including the potential for triangulation, transference, and shame that contribute to professional vulnerability for both parties. The core values of the Therapeutic Assessment model (collaboration, respect, humility, compassion, and openness and curiosity), as described by Finn (2009), provide a useful framework for working with clients and referring professionals alike.

Existing research on C/TA supports the examination of its use in a consultative context. Initial randomized trials demonstrated the effectiveness of Therapeutic Assessment in reducing symptomatic distress and increasing self-esteem with college students on counseling center waitlists (Finn & Tonsager, 1992; Newman & Greenway, 1997). The results of a meta-analysis of 17 randomized trial reports with 1,496 adult and adolescent participants comparing the effects of individualized feedback, a hallmark of C/TA, to various control conditions (e.g., no treatment, traditional assessment) indicated a significant overall effect (Cohen's $d = .42$; Poston & Hanson, 2010). A reanalysis of the data (14 studies, 1,375 participants) resulted in a significant overall effect of $d = .40$ (Hanson & Poston, 2011). Further, Poston and Hanson found different effects for symptom ($d = .37$; e.g., self-esteem, depressive symptomatology) and process ($d = .55$; e.g., working alliance, session depth) outcome variables. To this point, a randomized clinical trial comparing Finn's (2007) Therapeutic Assessment to a structured pretreatment intervention for adults with personality disorders found greater readiness for subsequent treatment on salient indicators (i.e., outcome expectancies, perception of progress toward treatment, satisfaction with services) for the group receiving Therapeutic Assessment (De Saeger et al., 2014). Hilsenroth, Peters, and Ackerman (2004) also found that participating in a C/TA intervention resulted in a stronger therapeutic alliance to the assessor and to the assessor turned psychotherapist, compared to participants receiving a traditional assessment. A small number of quasi-experimental studies using a single-case time-series design provide additional evidence of the effectiveness of these models with adult clients (Aschieri & Smith, 2012; J. D. Smith & George, 2012; Tarocchi, Aschieri, Fantini, & Smith, 2013). Studies testing the effects of collaborative and Therapeutic Assessment with children, adolescents, and families show effects on family functioning and child symptom severity (Tharinger et al., 2009), reduced family distress and behavioral problems of preadolescent boys with oppositional defiant disorder (J. D. Smith, Handler, & Nash, 2010), and improved engagement in community-based services following an emergency room visit for serious self-inflicted injury (Ougrin, Ng, & Low, 2008).

THIS STUDY

This study addresses a critical aspect of the treatment utility of psychological assessment: consultation with psychotherapists. We conducted a replicated single-case experiment with 10 participants to examine the effectiveness of C/TA in reducing clients' symptomatic distress and improving the processes and outcomes of ongoing psychotherapy. Participants were

identified and referred to the study by psychotherapists in the community. We hypothesized that participation in a midtherapy C/TA would (a) reduce client's self-reported symptomatic distress, collected using an Internet-based reporting system, and (b) improve therapy process variables that are instrumental to client change, such as the working alliance and other aspects of the psychotherapist–client relationship. Previous studies suggest that symptom improvement, as reported by clients, should follow the onset of C/TA (e.g., Aschieri & Smith, 2012; J. D. Smith & George, 2012) and that the psychotherapist–client dyad will report an improvement in therapeutic process variables, such as the working alliance (Ackerman, Hilsenroth, Baity, & Blagys, 2000). We took a pragmatic approach to the conduct of this study, exemplified by broad participant eligibility, a tailored intervention protocol, a client-centered outcome assessment strategy, and a design that is compatible with the demands of real-world professional practice. Pragmatic studies are intended to test the real-world effectiveness of an evidence-based intervention; they value generalizability over internal validity; and they facilitate research translation (Glasgow, 2013).

METHODS

Participant Recruitment

The study was advertised to licensed psychologists through e-mails and presentations given to the Santa Barbara, CA, area chapter of the state psychological association. Psychologists were asked to identify a client who might benefit from participating in an assessment as an adjunct to psychotherapy. The inclusion criteria for participation in the study included ongoing psychotherapy with the same therapist for at least 10 sessions, the therapist was a licensed doctoral-level psychologist, and both therapist and client agreed that an assessment might be useful for informing treatment. Initially, recruitment targeted clients who had become "stuck" or "stalled" in psychotherapy; however, this approach proved unsuccessful and the inclusion criteria were broadened shortly after enrollment began to allow for greater judgment by the therapist to determine the appropriateness of a referral for assessment consultation. Additionally, clients with ongoing legal involvement (e.g., a child custody case) or a medical or psychological condition that would impair their ability to consent to participation in a research study were ineligible. Seventeen therapists and their clients contacted the research team and were screened for eligibility. Three dyads meeting eligibility for the study terminated psychotherapy prior to beginning the study and five dyads opted not to enroll after screening for the following reasons: One client expressed concerns about confidentiality, one client only wanted a female assessor and openings were only available with a male at the time, and three dyads decided not to participate for unspecified reasons. The client and the therapist each received \$200 for their participation. The institutional review board of the University of California at Santa Barbara approved the research.

Participants

Eleven dyads completed the intake session and received C/TA. After enrollment, one participant was found to have

substantial cognitive issues that prohibited completion of the assessments and was thus excluded from all analyses and reported results. The remaining 10 clients were mostly female (70%), had an average age of 33.9 years ($SD = 12.1$, range = 20–50), and comprised the following racial and ethnic backgrounds: White (7), Latina (2), and multiracial (1). Eight were never married, 1 was married, and 1 was divorced. Eight licensed psychologists from the local community (3 male, 5 female) participated in the study. Formal diagnostic impressions were not gathered from the referring therapist. Examination of the presenting concerns and idiographic daily indexes derived from the referral issue (see “Measures” section) indicates that the primary concerns were cooccurring depressive symptomatology and anxiety (Participants 1, 3, 4, 5); depression (2, 7, 8); anxiety (9, 10); and anxiety, substance abuse, inattention, and sleep disturbance (6). The “Measures” section provides an indication of the specific symptomatic and behavioral issues associated with these primary concerns.

Study Design and Assessment Schedule

The study followed a replicated single-case design with three phases: baseline, intervention, and follow-up. Following enrollment in the study, the client was scheduled for an initial intake with a research assistant who explained the study procedures, gained informed consent, gathered relevant background information, and explored the client’s treatment goals in psychotherapy. Based on this interview, the research assistant, in collaboration with the principal investigator (Steven R. Smith), designed brief, idiographic outcome indexes that the client was then asked to complete on a daily basis beginning the day of the intake. The clients also completed ratings of each psychotherapy session throughout the three phases of the study. All outcome measures were administered via Survey-monkey.com. After a minimum 2-week baseline phase, which provides the necessary number of daily reports for valid and reliable analysis, clients met with the assessment clinician to begin the intervention. Clients continued to complete the daily ratings throughout the C/TA and for approximately 2 months afterward. Psychotherapy continued throughout the study procedures and the C/TA.

Intervention

The C/TA intervention consisted of an initial interview, test administration, and a joint feedback session with the client and psychotherapist. Delivery largely followed the procedures of the Therapeutic Model of Assessment described by Hilsenroth and colleagues (Hilsenroth et al., 2004) with adaptation to incorporate recent developments from the Therapeutic Assessment (Finn, 2007) model. Further, the joint feedback session was delivered according to Finn’s (2007, 2011) approach to using assessment in consultation. The assessors were a licensed clinical psychologist (male) or an advanced graduate student (female) in a clinical psychology doctoral program. The licensed psychologist (Steven R. Smith), who oversaw the clinical and research activities, has extensive training in C/TA and served as a clinician in the studies conducted by Ackerman et al. (2000) and Hilsenroth et al. (2004) using the C/TA model. The graduate student was supervised throughout the study and had completed course work in C/TA

and provided assessment services using the model before the study began.

The research assistants provided the assessors with the clients’ background and demographic information collected during the intake appointment. However, the assessors remained blind to the clients’ daily rating items so as to reduce potential bias. In the initial interviews, the clients’ symptoms, presenting problems, course of psychotherapy, impressions of therapy progress, and goals for the assessment were discussed. Following the procedures of the model, the client was asked to pose two or three questions that they hoped the assessment process could address (Finn, 2007). The clients’ psychotherapists were then interviewed by telephone regarding their therapeutic work with the client and their current conceptualization of the case. Therapists also posed two to three questions for the assessment.

Following the interview, clients were scheduled for one or more testing sessions. The administration of a multimethod personality assessment battery, consisting of the Personality Assessment Inventory (Morey, 1991), the Rorschach Inkblot Test (Rorschach, 1921/1942), and Thematic Apperception Test (Murray, 1943), required one to three 60- to 90-min testing sessions to complete. Six clients completed the battery in one session and two and three sessions were required for two clients in each, resulting in an average total time of about 5 hr with the assessor in the C/TA intervention (including the joint feedback session). The Rorschach was scored using the Comprehensive System (Exner, 2003).

Shortly after the final assessment meeting, a joint feedback session among the assessment clinician, client, and referring therapist was held in the therapist’s office during the client’s regularly scheduled therapy time. Although somewhat flexible, the sessions had the following procedural components. First, the assessment clinician thanked both the client and therapist for participation and the opportunity to enter their therapeutic endeavor. Second, the therapist was instructed to observe the process of their client participating in the feedback and serve as a “translator” of the findings—adding nuance to the feedback by relating the information back to particular instances in the client’s life or the therapy. Defining the therapist’s role as predominantly supportive was deemed valuable to providing a structure for this potentially complicated interpersonal exchange. The focus of the joint feedback session was on addressing the assessment questions posed by the client and therapist at the beginning of the assessment. The feedback followed the evidence-based procedures described by Finn and colleagues (Finn, 2007; J. D. Smith & Finn, 2014; Tharinger et al., 2008). At the end, the therapist and client were presented with a technical assessment report and a personalized letter, respectively (see Finn, 2007).

Measures

Daily ratings. In collaboration with the research assistants, each client identified idiographic indexes to be answered daily throughout the study. All participants reported on their general mood each day. With the exception of Participant 1, who had only this item, all other clients reported an additional two to four idiographic items, which are listed in Table 1, for a total of three to five items rated each day by each participant. The items were selected specifically for each client as

TABLE 1.—Descriptives of individualized daily measures of symptomatic distress and results of idiographic level- and slope-change analyses.

| Client | Baseline | | | | Intervention | | | | Follow-Up | | | | AR | Level Change | Slope Change |
|--------|----------|------|------|-----------|--------------|------|------|-----------|-----------|------|------|-----------|-------|--------------|--------------|
| | N | M | SD | % Missing | N | M | SD | % Missing | N | M | SD | % Missing | | | |
| 1 | 22 | 5.76 | 1.18 | 18.5 | 22 | 6.11 | 1.23 | 18.2 | 50 | 5.08 | 1.01 | 25.5 | .189 | .14 | -.14 |
| 2 | 28 | 2.59 | 2.52 | 12.5 | 37 | 3.57 | 1.87 | 41.7 | 77 | 3.57 | 2.18 | 64.1 | .185 | | |
| 3 | 8 | 3.89 | 0.27 | 0.0 | 44 | 3.74 | 0.30 | 0.0 | 82 | 3.52 | 0.33 | 6.0 | .253 | -.22* | .41*** |
| 4 | 21 | 4.22 | 0.83 | 0.0 | 33 | 3.76 | 1.03 | 15.6 | 103 | 4.21 | 0.88 | 31.9 | .272 | -.07 | .03 |
| 5 | 15 | 4.48 | 1.03 | 7.7 | 63 | 2.57 | 0.90 | 11.3 | 53 | 2.83 | 1.03 | 13.2 | .310 | -.50*** | .08 |
| 6 | 21 | 4.81 | 0.52 | 40.0 | 43 | 3.82 | 1.00 | 31.0 | 17 | 3.36 | 0.87 | 34.7 | .044 | -.44** | .32* |
| 7 | 14 | 4.97 | 0.62 | 35.3 | 22 | 5.35 | 0.64 | 40.9 | 100 | 3.48 | 0.62 | 35.4 | -.422 | -.40* | .68** |
| 8 | 19 | 4.53 | 1.18 | 13.3 | 79 | 3.98 | 1.27 | 12.8 | 74 | 3.58 | 1.06 | 21.9 | .113 | -.18* | .19* |
| 9 | 17 | 3.43 | 0.77 | 0.0 | 41 | 3.02 | 0.53 | 40.0 | 66 | 3.58 | 0.58 | 37.9 | .223 | -.05 | -.26 |
| 10 | 19 | 5.20 | 1.04 | 0.0 | 61 | 4.42 | 1.22 | 37.3 | 67 | 4.12 | 1.43 | 38.2 | .186 | -.27* | .28* |
| M | 18.40 | 4.39 | 1.00 | 12.73 | 44.50 | 4.03 | 1.00 | 24.88 | 68.90 | 3.73 | 1.00 | 30.88 | .220 | | |

Note. The *N* refers to the total duration of days in each phase. AR = Lag-1 autocorrelation, calculated for the entire data stream after replacing missing data using the expectation-maximization algorithm. Level and slope change results are presented as correlation coefficients (Pearson's *r*). Negative level change indicates reduced symptom distress. Positive slope change indicates a correlation to a vector that is flat during the baseline phase and declining during the intervention phase. Participant 2 was excluded from the analysis due to the degree of missing data. Individual item content: All participants reported on general mood; additional individual items were: 2 (hope, energy), 3 (anxious, depressed), 4 (productivity, anxiety, mindful of eating), 5 (confidence, anxiety, energy, comfort with self), 6 (attitude toward the future, restful sleep, focus at work, felt in-the-moment), 7 (pain, self-confidence, professional competence, self-worth), 8 (depressed, energy, motivation at work), 9 (self-worth, depressed, focus), 10 (hopeful, inspired, sense of purpose, fulfilled).

p* < .05. *p* < .01. ****p* < .001.

measurable indicators of improvement based on the presenting concerns and treatment goals. The indexes also had to be assessable on a daily basis. The items were rated on a response scale with a range from 1 to 7 and pertained to the degree to which the client experienced specific affective and symptomatic distress, such as "I feel hopeful" (1 = *not at all*, 7 = *very*) and "My anxiety level is" (1 = *none at all*, 7 = *very severe*). Items were recoded prior to analysis such that decreased ratings were indicative of desired improvement. An electronic reporting platform (SurveyMonkey.com) was selected because compliance rates have been found to be higher compared to daily paper-and-pencil measures (Palermo, Valenzuela, & Stork, 2004) and electronic time stamping prohibits retrospective reporting. Clients were given a brief overview of the website by research assistants in the intake interview. After the initial intake and the selection of daily measure items, research assistants activated the client's online account and daily reporting began. Each client was provided with a unique user name and password to ensure security and confidentiality. These procedures are modeled after the methods described by Borckardt and colleagues (2008), which have been used in studies of the effectiveness of assessment-based interventions (e.g., J. D. Smith, Handler, et al., 2010). Although the sample size of this study prohibits estimating the reliability of the daily ratings, published daily report studies of general mood symptoms (Cranford et al., 2006), for example, reported high internal consistency ($\alpha > .80$).

Psychotherapy process. Salient indicators of therapy processes were assessed after each psychotherapy session. Positive and negative aspects of the psychotherapy process were rated by the client using the 40-item revised version of the Vanderbilt Psychotherapy Process Scale (VPPS) patient form (S. R. Smith, Hilsenroth, Baity, & Knowles, 2003), which assesses six aspects of patients' attitudes and behaviors regarding their most recent treatment session, including therapist warmth and friendliness, therapist exploration, patient exploration, negative therapist-patient relationship, patient psychic distress, and patient dependency. Items pertaining to in-session behaviors consisted of descriptions of common

therapeutic interactions (e.g., "actively participated" and "seemed motivated"). The following stem preceded descriptors of attitude (e.g., tense, withdrawn, hostile), which were rated on a 5-point Likert scale: "Describe your demeanor during this hour." The revised patient form of the VPPS has been found to have adequate internal reliability with an overall alpha of .77 for the client report version (S. R. Smith et al., 2003). The Working Alliance Inventory-Short Form Revised (WAI-SR; Hatcher & Gillaspay, 2006) was used to assess the client-psychotherapist relationship. The WAI-SR assesses three core aspects of the therapeutic alliance: agreement on the tasks of therapy, agreement on the goals of therapy, and the development of the therapeutic bond. The three subscales have demonstrated high internal consistency ($\alpha > .80$). The therapist versions of these two measures were also administered; however, low return rates prohibited analysis of these data.

Data Analysis

To determine intervention effects on symptomatic distress, the daily report data were analyzed using three complementary approaches. First, we evaluated idiographic effects of the C/TA using Simulation Modeling Analysis (SMA; Borckardt et al., 2008). Next, we applied multilevel modeling (Shadish, Kyse, & Rindskopf, 2013; Van den Noortgate & Onghena, 2003) and then the *d*-statistic (Hedges, Pustejovsky, & Shadish, 2013) to evaluate the aggregate effect.

SMA was designed for the analysis of time-series intervention studies with the type of short baseline periods typical in applied clinical research. We conducted a level change analysis, which compares the degree of change in the mean level of symptomatic distress from baseline to the intervention phase of the study (including the follow-up phase) and a slope change analysis, which compares the observed trajectory of the data to an a priori model; in this case, a flat slope during the baseline phase followed by a linear decrease during the intervention phase. Pearson's *r* correlation coefficients are produced to indicate the magnitude of level change and the degree of concordance with the a priori slope change model.

SMA then applies a Monte Carlo simulation of 5,000 data streams sharing the same characteristics as the observed data (e.g., length of the two phases, degree of autocorrelation) to determine the actual probability of obtaining the observed intervention effects (Borckardt et al., 2008).

Multilevel modeling of single-case time-series data can provide evidence of a treatment effect in two forms: (a) an immediate change when the intervention begins, and (b) a change in the slope of symptom ratings occurring with the onset of the intervention (Van den Noortgate & Onghena, 2008). Recent developments in the specification of multilevel models demonstrate applicability to small ($N \leq 10$) sample sizes (Ferron, Bell, Hess, Rendina-Gobioff, & Hibbard, 2009; Shadish, Kyse, et al., 2013). Consistent with previous research demonstrating that the effects of C/TA might be dependent on the individualized feedback session (Poston & Hanson, 2010), and that effects continue to grow in the months following the assessment (J. D. Smith, Handler, et al., 2010), we hypothesized a significant change in slope but no significant immediate reduction in symptomatic distress at the onset of C/TA. Because multilevel modeling procedures do not produce a standardized effect size, we then calculated a d statistic, which is equivalent in magnitude and interpretation to Cohen's d but is corrected for small sample bias, which tends to inflate estimates of intervention effects (Hedges et al., 2013). This metric was found to be least affected by autocorrelation, or the serial dependence, of time-series data streams (Manolov & Solanas, 2008). A significant change in the level and trajectory of symptomatic distress, as evidenced by the results of SMA and multilevel modeling, and a meaningful effect size (d statistic) provides evidence of effectiveness.

The psychotherapy process data were analyzed with a repeated measures analysis of variance to examine change between successive phases in the study. The results of the therapy process variable analyses in this study are preliminary due to the small sample size and its effect on statistical power. Further, clients contributed differential proportions of data in each phase because of attending differing numbers of psychotherapy sessions and noncompliance. For these reasons, the data were aggregated within the three phases for each participant.

RESULTS

Preliminary Analyses

The burden placed on participants by daily measurement strategies necessitated methods for accounting for missing observations that are typically unavoidable in this type of research (Shiffman, Stone, & Hufford, 2008). The issue of rate of missing observations is analogous to compliance. The average proportion of missing data in this sample was 21%, which is consistent with previous research using similar daily diary methods (e.g., J. D. Smith, Handler, et al., 2010). Simulation studies of missing data in time-series data streams indicate that inferential precision is acceptable when maximum likelihood estimation techniques are applied and autocorrelation estimates are estimated to be less than .80 (J. D. Smith, Borckardt, & Nash, 2012). One participant¹ had unacceptable

rates of compliance with the daily measures—only responding on half of the days during the study period—and was thus excluded from analyses of daily report data. The expectation-maximization algorithm (Dempster, Laird, & Rubin, 1977) was applied to the data to impute missing values prior to conducting analyses in SMA and calculating the d statistic. We used the default restricted maximum likelihood estimator for analyses conducted in SAS Proc Mixed, which is considered the gold standard for longitudinal methods in a multilevel modeling framework (Ibrahim, Chen, Lipsitz, & Herring, 2005). Maximum likelihood estimation techniques provide accurate estimates of missing values only when data are missing completely at random (MCAR). Using all available daily ratings and relevant client variables (age, gender, ethnicity), as well as therapist, Little's (1988) MCAR test was nonsignificant for the entire data stream, $\chi^2(3404) = 3276.73, p = .940$. The test was also nonsignificant when applied to the baseline data, $\chi^2(353) = 95.18, p = 1.00$, and the intervention plus follow-up data, $\chi^2(2186) = 2094.97, p = .917$. Thus, missing data did not introduce bias into the analyses.

Although high-density measurement strategies such as daily report reduce the potential for retrospective bias by collecting real-time information and increase the validity and reliability of ratings by reducing measurement error (Bolger, Davis, & Rafaeli, 2003; Shiffman et al., 2008), subjective reports on multiple related indexes are prone to high intercorrelations (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). This might be particularly true when indexes are indicators or correlates of a single clinical syndrome. For this reason, each participant's daily measures were assessed for daily intercorrelation, referred to as a lag-0 cross-correlation analysis. The results indicated medium to strong intercorrelations among each participant's daily ratings with a range of Pearson's r from .29 to .85 ($M = 0.55, SD = 0.18$). Thus, creating a daily mean score to capture general *symptomatic distress* on a daily basis was deemed appropriate. The individual items reported by each participant are included in the note in Table 1. Figure 1 provides a graphical representation of the composite daily ratings data. This data reduction step reduces Type I error rates. Last, we examined the autocorrelation estimates of the individual data streams after applying the expectation-maximization algorithm to impute missing values. Lag-1 autocorrelations were generally small for the composite daily measure with an average of .220 ($SD = .10$). Descriptive statistics of the participants' daily mean scores and proportion of missing data are presented in Table 1. Examination of the mean differences between phases at the individual level suggests that response to the intervention was quite variable. The sample mean shows a modest decrease from baseline levels ($M = 4.39$) to the intervention phase ($M = 4.03$) and into the follow-up ($M = 3.73$). The average duration of the phases were found to be 18.4 days in the baseline, 44.5 during the C/TA, and 68.9 during follow-up.

Intervention Effects

Idiographic. The results of the level and slope change analysis conducted in SMA are presented in Table 1. Both analyses compared the daily ratings during the baseline phase to those from the intervention and follow-up phases combined. To summarize, the level change analyses indicated that 6 of

¹Participant 2 in Table 1.

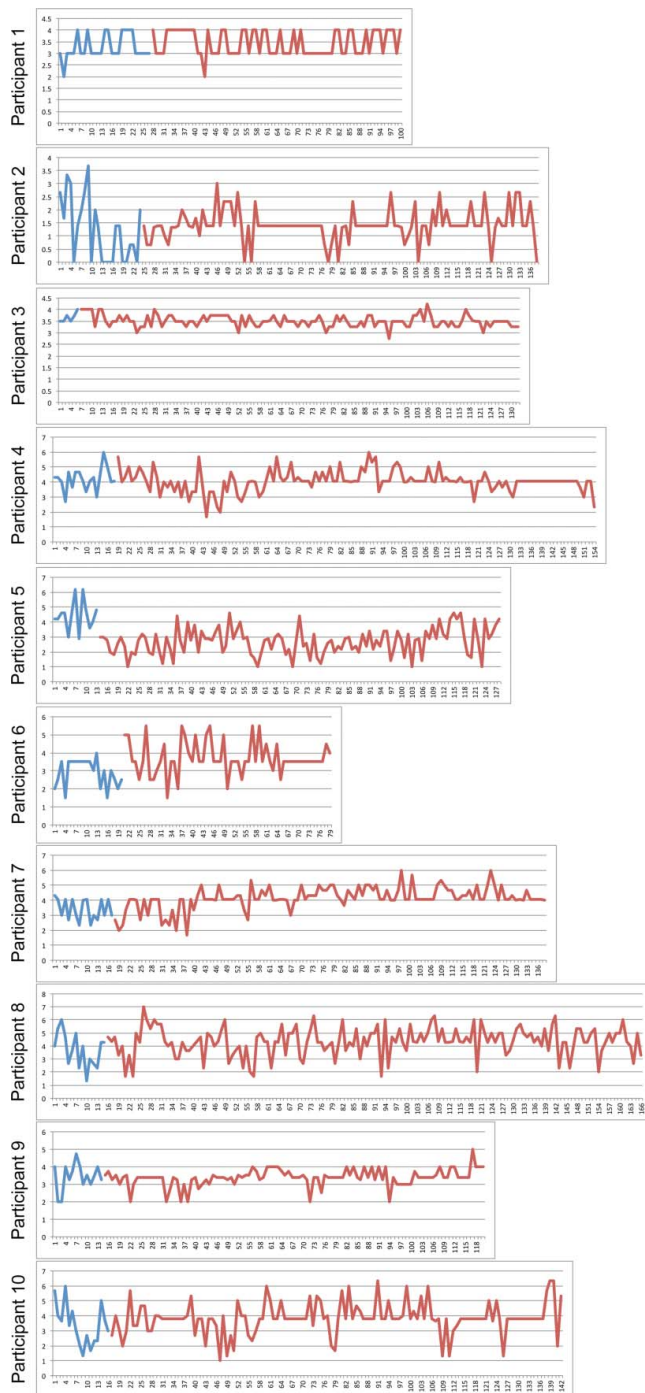


FIGURE 1.—Mean score graphs for daily reports of individualized symptomatic distress. *Note.* Blue line is baseline. Red line is the intervention (collaborative/therapeutic assessment) and follow-up phases combined.

the 9 participants experienced statistically significant ($p < .05$) reductions in the mean level of symptomatic distress between phases. Correlation coefficients of significant results ranged from $r = -.18$ to $-.50$. The results of 1 participant indicated an increase in symptomatic distress, but the change was not statistically significant ($r = -.17$). Results of the slope change analyses indicated significant correlation to the a priori model (flat during baseline followed by a linear decrease) for 5 of the

participants. Correlation coefficients of significant slope change analyses were small to large (range = $.19$ – $.68$).

Synthesis. We conducted a multilevel modeling analysis for multiple baseline designs in SAS Proc Mixed (SAS Institute, 2008) following the procedures described by Shadish et al. (2014) and Van den Noortgate and Onghena (2003). As it is unlikely that the baseline level and treatment effect were the same across all participants, the case-specific coefficients were allowed to vary across participants, and we modeled heterogeneous within-case variances in both phases to address potentially biasing assumptions of multilevel modeling (Ferron et al., 2009). We first fit an unconditional two-level growth model by testing linear, quadratic, cubic, and polynomial terms to account for possible trends in the data, which could bias estimates of intervention effects if found to be significant. Best fit was determined by the lowest Bayesian information criterion (BIC) value. A model with linear and quadratic terms provided the best fit to the data. Each of the parameters in the best fitting model had significant variance: intercept ($B = 4.55$, $SE = .09$, $p < .001$), linear slope ($B = -.022$, $SE = .007$, $p < .01$), quadratic slope ($B = 0.0002$, $SE = .00004$, $p < .001$). The positive quadratic term indicates that the rate of downward change slows with time.

We proceeded by adding a treatment condition variable to indicate onset of the intervention for each participant. We also added a parameter to model the autocorrelation of the ratings and included pertinent potential covariates, including client age and gender, as well as assessor. Following the backward trimming method described by Singer and Willett (2003), as covariates were entered into the model, one at a time, those that significantly related to one of the model's parameters were retained. None of the covariates tested (client age, gender, ethnicity; assessor) were significantly associated and thus were not retained in the final model. The autocorrelation parameter indicated a nonsignificant lag-1 estimate of $.211$. Inclusion of the autocorrelation parameter did not influence the immediate intervention effect or the slope change results and thus the autocorrelation parameter was not included in the final model. The final model contains two parameters of interest. First, immediate intervention effects, defined as a significant difference at the onset of the intervention, were in the desired direction but were not significant, $B = -.15$, $t(832) = -.67$, $p = .30$. The second parameter pertains to an effect of the intervention on the trend or slope of symptomatic distress. Estimates of the slope of the baseline period indicated a slight downward trajectory that was not significantly different from zero ($B = -.007$, $SE = .01$, $p = .36$) and indicates stability. Our final model revealed a significant increase in the downward rate of linear change of symptomatic distress ratings coinciding with the onset of the intervention, $B = -.06$, $t(830) = -2.98$, $p < .01$.

Based on the evidence from the idiographic and multilevel analyses indicating that onset of treatment had an effect on the level and trajectory of symptomatic distress, we calculated the d statistic to garner the overall magnitude of effect for the intervention (Hedges, Pustejovsky, & Shadish, 2012; Shadish et al., 2014). We used the DHPS macro (Shadish, Pustejovsky, & Hedges, 2013) in SPSS Statistics (2012) for this analysis. Based on Cohen's (1988) interpretive guidelines, the results indicated a medium intervention effect ($d = -.50$,

TABLE 2.—Descriptives of client-reported psychotherapy process variables by study phase.

| Variable | Baseline | | Intervention | | Follow-Up | |
|-----------------------------------|----------|-----------|--------------|-----------|-----------|-----------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| Working alliance (Total score) | 44.90 | 5.38 | 47.69 | 5.92 | 49.12 | 6.29 |
| Working Alliance: Goals | 14.15 | 2.18 | 14.57 | 2.50 | 15.30 | 2.84 |
| Working Alliance: Task | 12.93 | 3.61 | 14.51 | 3.66 | 14.98 | 3.79 |
| Working Alliance: Bond | 17.82 | 1.12 | 18.57 | 1.44 | 18.84 | 1.06 |
| Negative relationship | 1.26 | 0.28 | 1.34 | 0.22 | 1.20 | 0.22 |
| Patient psychic distress | 2.55 | 0.63 | 2.28 | 0.59 | 2.02 | 0.76 |
| Therapist exploration | 4.09 | 0.57 | 4.20 | 0.65 | 4.29 | 0.65 |
| Therapist warmth and friendliness | 4.37 | 0.44 | 4.40 | 0.38 | 4.63 | 0.31 |
| Patient participation | 3.85 | 0.55 | 3.84 | 0.45 | 3.92 | 0.42 |
| Patient dependency | 2.83 | 0.67 | 2.47 | 0.51 | 2.58 | 0.32 |

Note. Within-participant means were calculated and then an overall mean score was derived.

variance = .013) in reducing symptomatic distress, comparing the baseline phase to the remainder of the study period after the intervention began (i.e., intervention and follow-up phases combined).

Psychotherapy Process

A repeated measures analysis of variance was conducted on the WAI and its three subscales as well as the six subscales of the VPPS. Descriptive statistics are presented in Table 2. One participant (8) was excluded from the analysis of psychotherapy process variables because the therapy was terminated shortly after the feedback session and no measures were completed during the follow-up phase. The remaining 9 participants provided at least one set of ratings in each phase of the study and means were calculated for each participant within phase. Client gender, age, and ethnicity, as well as assessor, were tested as covariates but were not significantly related to the linear trajectory of any of the outcome variables. Thus, for parsimony they are not included in the final results presented. Cohen's *d* values are provided, which were converted from partial eta squared. A significant effect was found for increasing the working alliance between the client and therapist after the joint feedback session, $F(2, 94) = 6.96, p = .048, d = 1.32$. Examination of the WAI subscales revealed a significant increase in the Task subscale, $F(2, 94) = 16.75, p = .015, d = 2.04$, but not the Goals, $F(2, 94) = 1.21, p = .33, d = .55$, or Bond, $F(2, 94) = 2.64, p = .18, d = .81$. The positive indicators of the VPPS increased over time but were not significant: patient participation, $F(2, 94) = .52, p = .51, d = .36$; therapist exploration, $F(2, 94) = 1.62, p = .27, d = .64$; and therapist's warmth and friendliness, $F(2, 94) = 1.39, p = .30, d = .59$. The negative indicators of the VPPS were also found to be declining over time but no statistically significant reductions were found: negative relationship, $F(2, 94) = -.76, p = .43, d = -.44$; patient psychic distress, $F(2, 94) = -2.861, p = .17, d = -.85$; patient dependency, $F(2, 94) = -.50, p = .52, d = -.35$.

DISCUSSION

The collective body of research on collaborative and therapeutic models of assessment suggests that this intervention

approach is effective in reducing client symptoms and improving salient psychotherapy process variables. Yet, these models had not previously been evaluated in the context of a consultation with psychotherapists in ongoing treatment. Our results indicate that participation in a midtherapy consultation using C/TA coincides with a significant reduction in clients' symptomatic distress. Our findings are consistent with the magnitude of intervention effects in published studies of this type of intervention. Our overall effect of $d = -.50$ on symptomatic distress is commensurate with the findings of Poston and Hanson's (2010) meta-analysis, which reported an effect size of $d = .37$ for client outcomes only, as opposed to the overall effect that included therapeutic process variables. The effect size of this study is particularly promising, given that pragmatic studies invariably produce smaller effects due to the heterogeneity of the participants, the flexibility the clinician is afforded in delivering the intervention, and other factors typically controlled for in an explanatory (i.e., efficacy) intervention trial (Ware & Hamel, 2011). In the context of a within-subject design, our results indicate that clients experienced a reduction in symptomatic distress of 0.5 *SD* after beginning the C/TA in comparison to baseline levels. Further, outcomes were not a function of the assessor delivering the C/TA in our study. The aggregate results are supported by the findings from an idiographic analytic approach, which suggested that 6 of the 9 clients experienced significant reductions. Causality cannot be determined from this study design and replication of the effect in other samples with randomized designs and controls for possible confounding process are needed. For example, the effects found in this study could plausibly be a result of the ongoing psychotherapy.

A crucial question that is unanswerable from this study design, but begs for discussion, is the precise mechanism of action in the application of C/TA for psychotherapy consultation. There is more than one possible component of the C/TA that could be responsible for observed changes. First, the literature supports the assertion that the receipt of collaborative feedback alone is sufficient to improve client outcomes (Poston & Hanson, 2010). Second, an implicit mechanism of assessment consultation is that the findings aid the therapist in the clinical endeavor moving forward—likely through revising the case conceptualization. Both hypotheses are plausible and supported by empirical evidence in the literature but cannot be explicitly tested with the design used in this study. With the confines of the study in mind, we believe that the feedback jointly provided to both the therapist and the client is responsible for subsequent changes—feeling understood by the assessor reduced the client's distress, the therapist's empathy for the client increased, and the therapist adjusted the therapeutic approach according to a more accurate and compassionate conceptualization of the client's difficulties.

When interpreting the findings, it is important to keep in mind that the baseline period of this study was not a true "pretreatment" baseline, but rather a measure of symptomatic distress during psychotherapy prior to beginning the assessment. Presumably after 10 or more sessions of psychotherapy, which was required for enrollment, alleviation of acute distress symptoms has likely already occurred (e.g., Kopta, Howard, Lowry, & Beutler, 1994), meaning our results are representative of improvements in chronic symptomatic

distress that are likely attributable to the addition of the C/TA intervention. Similarly, previous research suggests that the working alliance is less variable after approximately the ninth session of psychotherapy and the aggregate alliance score between sessions 3 and 9 accounts for the most variance in outcomes (Crits-Christoph, Gibbons, Hamilton, Ring-Kurtz, & Gallop, 2011). Despite these previous findings, we found statistically significant increases in the therapist–client working alliance in this study—mainly in the area of agreement on the tasks of therapy—indicating some malleability. This finding is consistent with previous research (e.g., Hilsenroth et al., 2004). It seems plausible that the C/TA, and specifically the joint feedback session, clarified the case conceptualization and helped to identify the therapeutic tasks required to address the client’s issues. This needs to be tested more explicitly and our results in this domain should be considered preliminary due to the small sample, nonrandomized design, and variability in participants’ response rates. We should also caution against interpretation of the Cohen’s *d* effect sizes in these analyses, as low-powered within-subjects studies are known to produce inflated effects (Lakens, 2013). The lack of statistical significance despite seemingly large effects attests to this point.

Limitations and Future Directions

There are some limitations to this study that affect the generalizability and interpretation of the results. First, the sample size would be considered small for an intervention study but it meets the standard for a well-designed replicated single-case or multiple baseline trial in terms of the number of participants (Kratochwill et al., 2010). Second, although the intervention protocol was semistructured, there was some variation in its delivery among the participants, which sacrifices internal validity for external validity and generalizability. Third, there are benefits and drawbacks of the individualized indexes used for the daily report: Identifying measures that are relevant to the client increases the clinical significance and validity of observed change (Borckardt et al., 2008), but also results in a somewhat broad metric of symptomatic distress. Ratings were also aggregated due to high intercorrelations. Subsequent studies using daily report methods could consider including unvarying indexes aligning with the intervention targets, as suggested by J. D. Smith (2012), including additional norm-based assessment of outcomes and observer ratings consistent with a multimethod approach. One potential limitation to generalizability of the findings was our requirement that the referring clinician be a psychologist. Replication with master’s-level therapists is a potential future direction. Another issue pertaining to generalizability was the criteria for referral to the study. As is likely consistent with typical community-based care, we provided minimal guidance to referring professionals, thus allowing them to determine appropriateness based on their own clinical judgment. The composition of the participants having concerns predominantly related to anxiety and depression limits the generalizability of the findings to other clinical syndromes and referral issues. Future studies should consider more systematic referral criteria. Last, there was no explicit assessment of fidelity to C/TA. The individually tailored nature of the protocol affords adaptation and flexibility, which is consistent with a pragmatic trial, but demonstrated

fidelity to the core components of C/TA would increase confidence that the observed intervention effects were due to the model and not extraneous factors. Last, multilevel modeling has received significant attention from single-case researchers and methodologists. Yet, there remains a need to further verify that the assumptions of multilevel modeling can be readily applied to single-case data of varying lengths, distributions, and structures (Shadish, 2014).

This study also has a number of strengths. To our knowledge, this is the first study to empirically test the effectiveness of C/TA in the context of psychotherapy consultation, which is one of the primary ways assessors practicing this model use it in the community (Finn, 2007; Finn et al., 2012). Second, we employed a methodologically strong research design with state-of-the-science methods for single-case research, which has traditionally relied on visual analysis despite varying accuracy in terms of intervention effects (Brossart, Parker, Olson, & Mahadevan, 2006; Harrington, 2013) and inadequate adherence to systematic visual analysis procedures (J. D. Smith, 2012). Last, due to the pragmatic nature of this study, the results are quite likely to be representative of expected effects with typical clients in psychotherapy. Inclusion of stalled or struggling therapist–client dyads could yield larger intervention effects, but this hypothesis will have to wait for a future trial.

CONCLUSIONS

The empirical study of treatment utility is critical to maintaining the relevance of psychological assessment in clinical psychology. Our findings indicate that collaborative and therapeutic models of assessment are a promising approach in the context of consultation with community psychotherapists. Assessment psychologists practicing C/TA are uniquely poised to provide an evidence-based perspective to other professionals, simultaneously affecting symptomatic change and the therapeutic relationship. Although interest in psychological assessment is waning in some professional circles, the advent of C/TA has reinstated its clinical relevance and opened new doors to addressing questions of treatment and clinical utility.

ACKNOWLEDGMENTS

The authors wish to thank the clients and therapists who generously participated in this study; Mark Van Ryzin and David Morso for assistance with the analytic methods; Filippo Aschieri and Katherine Thomas for their comments on drafts of this article; and members of the Society for Personality Assessment (SPA) who generously contributed to the SPA Foundation to support this research.

FUNDING

This research was supported by a research grant from the Society for Personality Assessment Foundation, awarded to Steven R. Smith. Justin D. Smith received support from the Society for Personality Assessment Foundation as the 2013 John E. Exner Scholar.

REFERENCES

- Ackerman, S. J., Hilsenroth, M. J., Baity, M. R., & Blagys, M. D. (2000). Interaction of therapeutic process and alliance during psychological assessment. *Journal of Personality Assessment*, 75, 82–109. doi:10.1207/S15327752JPA7501_7
- American Psychological Association. (2010). 2010 Amendments to the 2002 "Ethical principles of psychologists and code of conduct." *American Psychologist*, 65, 493.
- Aschieri, F., & McCarthy, M. G. (2014, March). *The secrets of successful assessment practice: Developing supporting relationships with referring professionals*. Paper presented at the annual meeting of the Society for Personality Assessment, Washington, DC.
- Aschieri, F., & Smith, J. D. (2012). The effectiveness of an adult Therapeutic Assessment: A single-case time-series experiment. *Journal of Personality Assessment*, 94, 1–11. doi:10.1080/00223891.2011.627964
- Belter, R. W., & Piotrowski, C. (2001). Current status of doctoral-level training in psychological testing. *Journal of Clinical Psychology*, 57, 717–726. doi:10.1002/jclp.1044
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, 54, 579–616. doi:10.1146/annurev.psych.54.101601.145030
- Borckardt, J. J., Nash, M. R., Murphy, M. D., Moore, M., Shaw, D., & O'Neil, P. (2008). Clinical practice as natural laboratory for psychotherapy research. *American Psychologist*, 63, 1–19. doi:10.1037/0003-066X.63.2.77
- Brossart, D. F., Parker, R. I., Olson, E. A., & Mahadevan, L. (2006). The relationship between visual analysis and five statistical analyses in a simple AB single-case research design. *Behavior Modification*, 30, 531–563. doi:10.1177/0145445503261167
- Clark, D. A. (1999). Case conceptualization and treatment failure: A commentary. *Journal of Cognitive Psychotherapy*, 13, 331–337.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, 32, 917–929. doi:10.1177/0146167206287721
- Crits-Christoph, P., Gibbons, M. B. C., Hamilton, J., Ring-Kurtz, S., & Gallop, R. (2011). The dependability of alliance assessments: The alliance–outcome correlation is larger than you might think. *Journal of Consulting and Clinical Psychology*, 79, 267–278. doi:10.1037/a0023668
- Curry, K. T., & Hanson, W. E. (2010). National survey of psychologists' test feedback training, supervision, and practice: A mixed methods study. *Journal of Personality Assessment*, 92, 327–336. doi:10.1080/00223891.2010.482006
- Dempster, A., Laird, N., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39, 1–38.
- De Saeger, H., Kamphuis, J. H., Finn, S. E., Smith, J. D., Verhuel, R., van Busschbach, J. J. V., . . . Horn, E. (2014). Therapeutic Assessment promotes treatment readiness but does not affect symptom change in patients with personality disorders: Findings from a randomized clinical trial. *Psychological Assessment*, 26, 474–483. doi:10.1037/a0035667
- Eisman, E. J., Dies, R. R., Finn, S. E., Eyde, L. D., Kay, G. G., Kubiszyn, T. W., . . . Moreland, K. L. (2000). Problems and limitations in using psychological assessment in the contemporary health care delivery system. *Professional Psychology: Research and Practice*, 31, 131–140. doi:10.1037/0893-3200.31.2.131
- Exner, J. E., Jr. (2003). *The Rorschach: A comprehensive system* (4th ed.). New York, NY: Wiley.
- Ferron, J., Bell, B. A., Hess, M. R., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences from multiple-baseline data: The utility of multilevel modeling approaches. *Behavior Research Methods*, 41, 372–384. doi:10.3758/BRM.41.2.372
- Finn, S. E. (2007). *In our client's shoes: Theory and techniques of Therapeutic Assessment*. Mahwah, NJ: Erlbaum.
- Finn, S. E. (2009). Core values in Therapeutic Assessment. Retrieved from <http://www.therapeuticassessment.com>
- Finn, S. E. (2011). Journeys through the valley of death: Multimethod psychological assessment and personality transformation in long-term psychotherapy. *Journal of Personality Assessment*, 93, 123–141. doi:10.1080/00223891.2010.542533
- Finn, S. E., Fischer, C. T., & Handler, L. (Eds.). (2012). *Collaborative/therapeutic assessment: A casebook and guide*. Hoboken, NJ: Wiley.
- Finn, S. E., & Tonsager, M. E. (1992). Therapeutic effects of providing MMPI-2 test feedback to college students awaiting therapy. *Psychological Assessment*, 4, 278–287. doi:10.1037/1040-3590.4.3.278
- Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment*, 9, 374–385. doi:10.1037/1040-3590.9.4.374
- Glasgow, R. E. (2013). What does it mean to be pragmatic? Pragmatic methods, measures, and models to facilitate research translation. *Health Education & Behavior*, 40, 257–265. doi:10.1177/1090198113486805
- Hanson, W. E., & Poston, J. M. (2011). Building confidence in psychological assessment as a therapeutic intervention: An empirically based reply to Lilienfeld, Garb, and Wood (2011). *Psychological Assessment*, 23, 1056–1062. doi:10.1037/a0025656
- Harrington, M. A. (2013). Comparing visual and statistical analysis in single-subject studies. *Open Access Dissertations*, Paper 5.
- Hatcher, R. L., & Gillaspay, J. A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. *Psychotherapy Research*, 16, 12–25. doi:10.1080/10503300500352500
- Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist*, 42, 963–974. doi:10.1037/0003-066X.42.11.963
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods*, 3, 224–239. doi:10.1002/jrsm.1052
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods*, 4, 324–341.
- Hilsenroth, M. J., Peters, E. J., & Ackerman, S. J. (2004). The development of therapeutic alliance during psychological assessment: Patient and therapist perspectives across treatment. *Journal of Personality Assessment*, 83, 332–344. doi:10.1207/s15327752jpa8303_14
- Ibrahim, J. G., Chen, M.-H., Lipsitz, S. R., & Herring, A. H. (2005). Missing-data methods for generalized linear models: A comparative review. *Journal of the American Statistical Association*, 100, 332–346. doi:10.1198/016214504000001844
- Kaslow, N. J. (2004). Competencies in professional psychology. *American Psychologist*, 59, 774–781. doi:10.1037/0003-066x.59.8.774
- Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, 62, 1009–1016.
- 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
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, 4, 863.
- Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring, and feedback in clinical practice*. Washington, DC: American Psychological Association.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83, 1198–1202.
- Manolov, R., & Solanas, A. (2008). Comparing $N = 1$ effect sizes in presence of autocorrelation. *Behavior Modification*, 32, 860–875. doi:10.1177/0145445508318866
- Maruish, M. E. (Ed.). (2004). *The use of psychological testing for treatment planning and outcomes assessment*. Mahwah, NJ: Erlbaum.

- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., ... Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, *56*, 128–165. doi:10.1037//0003-066X.56.2.128
- Morey, L. C. (1991). *Personality Assessment Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Murray, H. A. (1943). *Thematic Apperception Test: Manual*. Cambridge, MA: Harvard University Press.
- Newman, M. L., & Greenway, P. (1997). Therapeutic effects of providing MMPI–2 test feedback to clients at a university counseling service: A collaborative approach. *Psychological Assessment*, *9*, 122–131. doi:10.1037/1040-3590.9.2.122
- Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., ... Duhon, G. J. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review*, *34*, 87–106.
- Norcross, J. C., Karpiak, C. P., & Santoro, S. O. (2005). Clinical psychologists across the years: The Division of Clinical Psychology from 1960 to 2003. *Journal of Clinical Psychology*, *61*, 1467–1483. doi:10.1002/jclp.20135
- Ougrin, D., Ng, A. V., & Low, L. (2008). Therapeutic Assessment based on cognitive-analytic therapy for young people presenting with self-harm: Pilot study. *Psychiatric Bulletin*, *32*, 423–426. doi:10.1192/pb.bp.107.018473
- Pace, T. M., Chaney, J. M., Mullins, L. L., & Olson, R. A. (1995). Psychological consultation with primary care physicians: Obstacles and opportunities in the medical setting. *Professional Psychology: Research and Practice*, *26*, 123.
- Palermo, T. P., Valenzuela, D., & Stork, P. P. (2004). A randomized trial of electronic versus paper pain diaries in children: Impact on compliance, accuracy, and acceptability. *Pain*, *107*, 213–219. doi:10.1016/j.pain.2003.10.005
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903. doi:10.1037/0021-9010.88.5.879
- Poston, J. M., & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment*, *22*, 203–212. doi:10.1037/A001s679
- Rorschach, H. (1942). *Psychodiagnostics* (5th ed.). Berne, Switzerland: Verlag Hans Huber. (Original work published 1921)
- SAS Institute. (2008). *The SAS system for Windows, Version 9*. Cary, NC: Author.
- Shadish, W. R. (2014). Statistical analyses of single-case designs: The shape of things to come. *Current Directions in Psychological Science*, *23*, 139–146. doi:10.1177/0963721414524773
- Shadish, W. R., Hedges, L. V., Pustejovsky, J. E., Boyajian, J. G., Sullivan, K. J., Andrade, A., & Barrientos, J. L. (2014). A d-statistic for single-case designs that is equivalent to the usual between-groups d-statistic. *Neuropsychological Rehabilitation*, *24*, 528–553. doi:10.1080/09602011.2013.819021
- Shadish, W. R., Kyse, E. N., & Rindskopf, D. M. (2013). Analyzing data from single-case designs using multilevel models: New applications and some agenda items for future research. *Psychological Methods*, *18*, 385–405. doi:10.1037/a0032964
- Shadish, W. R., Pustejovsky, J. E., & Hedges, L. V. (2013). *The DHPS SPSS macro*. Retrieved from <http://faculty.ucmerced.edu/wshadish/software/software-meta-analysis-single-case-design>
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1–32. doi:10.1146/annurev.clinpsy.3.022806.091415
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, *78*, 298–311. doi:10.1037/a0019247
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.
- Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods*, *17*, 510–550. doi:10.1037/a0029312
- Smith, J. D., Borckardt, J. J., & Nash, M. R. (2012). Inferential precision in single-case time-series datastreams: How well does the EM procedure perform when missing observations occur in autocorrelated data? *Behavior Therapy*, *43*, 679–685. doi:10.1016/j.beth.2011.10.001
- Smith, J. D., & Finn, S. E. (2014). Therapeutic presentation of multimethod assessment results: Empirically supported guiding framework and case example. In C. J. Hopwood & R. F. Bornstein (Eds.), *Multimethod clinical assessment of personality and psychopathology* (pp. 403–425). New York, NY: Guilford.
- Smith, J. D., Finn, S. E., Swain, N. F., & Handler, L. (2010). Therapeutic Assessment in pediatric and primary care psychology: A case presentation of the model's application. *Families, Systems, & Health*, *28*, 369–386.
- Smith, J. D., & George, C. (2012). Therapeutic Assessment case study: Treatment of a woman diagnosed with metastatic cancer and attachment trauma. *Journal of Personality Assessment*, *94*, 331–344. doi:10.1080/00223891.2012.656860
- Smith, J. D., Handler, L., & Nash, M. R. (2010). Therapeutic Assessment for preadolescent boys with oppositional-defiant disorder: A replicated single-case time-series design. *Psychological Assessment*, *22*, 593–602. doi:10.1037/a0019697
- Smith, S. R., Hilsenroth, M. J., Baity, M. R., & Knowles, E. S. (2003). Assessment of patient and therapist perspectives of process: A revision of the Vanderbilt Psychotherapy Process Scale. *American Journal of Psychotherapy*, *57*, 195–205.
- SPSS Statistics Version 21.0.0. (2012). Chicago, IL: SPSS.
- Tarocchi, A., Aschieri, F., Fantini, F., & Smith, J. D. (2013). Therapeutic Assessment of complex trauma: A single-case time-series study. *Clinical Case Studies*, *12*, 228–245. doi:10.1177/1534650113479442
- Tharinger, D. J., Finn, S. E., Gentry, L., Hamilton, A. M., Fowler, J. L., Matson, M., ... Walkowiak, J. (2009). Therapeutic Assessment with children: A pilot study of treatment acceptability and outcome. *Journal of Personality Assessment*, *91*, 238–244. doi:10.1080/00223890902794275
- Tharinger, D. J., Finn, S. E., Hersh, B., Wilkinson, A. D., Christopher, G., & Tran, A. (2008). Assessment feedback with parents and pre-adolescent children: A collaborative approach. *Professional Psychology: Research and Practice*, *39*, 600–609. doi:10.1080/00223890902794275
- Van den Noortgate, W., & Onghena, P. (2003). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly*, *18*, 325–346. doi:10.1521/scpq.18.3.325.22577
- Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-Based Communication Assessment and Intervention*, *2*, 142–151. doi:10.1080/17489530802505362
- Ware, J. H., & Hamel, M. B. (2011). Pragmatic trials—Guides to better patient care? *New England Journal of Medicine*, *364*, 1685–1687.