

# Using Psychodynamic, Cognitive Behavioral, and Control Mastery Prototypes to Predict Change: A New Look at an Old Paradigm for Long-Term Single-Case Research

Nnamdi Pole  
University of Michigan

J. Stuart Ablon  
Massachusetts General Hospital and Harvard Medical School

Lynn E. O'Connor  
The Wright Institute and University of California, Berkeley

This article illustrates a method of testing models of change in individual long-term psychotherapy cases. A depressed client was treated with 208 sessions of control mastery therapy (CMT), an unmanualized approach that integrates elements of psychodynamic therapy (PDT) and cognitive behavioral therapy (CBT). Panels of experts developed prototypes of ideal PDT, CBT, and CMT process using the Psychotherapy Process Q-set (PQS; J. S. Ablon & E. E. Jones, 1999; E. E. Jones, L. A. Parke, & S. Pulos, 1992; E. E. Jones & S. M. Pulos, 1993). Independent observers rated every 4th session ( $N = 53$ ) with the PQS. Using correlations between ideal and actual PQS ratings followed by paired  $t$  tests, the authors compared adherence to the CMT prototype with adherence to plausible alternative models advocated by the PDT and CBT experts. Bivariate time series analyses determined whether prototype adherence predicted an estimated index of symptom change. Results showed that the therapist's behavior was most consistent with the CMT prototype and that this aspect of the CMT prototype along with particular aspects of the other prototypes influenced estimated symptom change. The results, which replicate and extend earlier findings, support the validity of this approach to studying long-term therapies but also highlight its limitations.

*Keywords:* prototypes, single-case studies, control mastery theory, psychotherapy process research, time series analysis

*Supplemental materials:* <http://dx.doi.org/10.1037/0022-0167.55.2.221.supp>

Long-term therapy has been relatively neglected in the psychotherapy research literature. One reason for this paucity may be the practical challenges of scrutinizing long-term treatments with recommended randomized clinical trial methodologies (Chambless & Hollon, 1998). For instance, the expense in-

involved in a randomized trial of therapies lasting a few months (Elkin, 1994) could become prohibitively expensive if extrapolated to treatments lasting a few years. Furthermore, specifying the treatment model in advance in a therapy manual may become untenable in lengthy open-ended treatments that lack a fixed focus and that are intended to change course in response to unanticipated events (Westen, Novotny, & Thompson-Brenner, 2004). Without a detailed therapy manual, the conventional methods of assessing whether the treatment model was consistently applied and whether it accounted for observed therapeutic changes would be inappropriate. Yet, given evidence that long-term therapies may be more effective than short-term therapies (Howard, Kopta, Krause, & Orlinsky, 1986; Orlinsky, Grawe, & Parks, 1994; Seligman, 1995), it seems worthwhile to develop cost-effective, alternative ways to study the relationship between prescribed therapeutic processes and outcomes in these therapies. In this article, we offer one approach using an intensive quantitative single-case design involving time series analyses.

## Studying the Relationship Between Adherence and Outcome in Group Designs

In studies comparing groups of people treated by different forms of therapy, the contribution of the specific theoretical

---

Nnamdi Pole, Department of Psychology, University of Michigan; J. Stuart Ablon, Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School; Lynn O'Connor, The Wright Institute, Berkeley, CA and Department of Psychology, University of California, Berkeley.

We take this opportunity to remember our mentors Enrico E. Jones and Joseph Weiss. Enrico E. Jones died in March 2003; Joseph Weiss died in November 2004. Their brilliant vision inspired this work and lives on in the authors. We also thank Jennifer F. P. Dyer, Joel T. Nigg, William K. Lamb, Tania Mara Miller, Ms. M, and the Berkeley Psychotherapy Research Group for their roles in gathering the original data on which this study is based. Special thanks to Polly Bloomberg-Fretter for her clinical acumen and thoughtful critique of an earlier draft of this manuscript and Robert W. Levenson for his modifications of the bivariate time series analysis software. Finally, we thank Delores Carter Pole for her special contributions.

Correspondence concerning this article should be addressed to Nnamdi Pole, Department of Psychology, University of Michigan, 530 Church Street, 2260 East Hall, Ann Arbor, MI 48109-1043. E-mail: [npole@email.smith.edu](mailto:npole@email.smith.edu)

model to outcome is shown through measures of adherence. Traditionally, these measures mirror the therapy manuals that define the treatments (e.g., Hill, O'Grady, & Elkin, 1992; Shapiro & Startup, 1992). Trained judges rate how closely actual therapy sessions conform to the prescriptions of their associated manuals (e.g., Barber et al., 2006; Hill et al., 1992; Shapiro & Startup, 1992). On occasion, these therapies are also assessed for whether and to what extent they conform to therapeutic processes prescribed by other therapy manuals (e.g., Hill et al., 1992). The case for specific effects of a particular theoretical model is made by examining the correlation between adherence to that model and relevant outcomes (e.g., Barber et al., 2006; DeRubeis & Feeley, 1990). Studies in this tradition have found that (a) different forms of therapy can be distinguished on the basis of adherence measures (e.g., therapists conducting cognitive behavioral therapy [CBT] have acted in closer accord with the CBT manual than with an interpersonal therapy (IPT) manual (Hill et al., 1992); (b) specific forms of therapy can contain significant elements belonging to other forms of therapy (e.g., CBT can contain ingredients of IPT; Hill et al., 1992); and (c) greater adherence to the therapists' ascribed theoretical model can predict better outcomes (e.g., greater adherence to CBT techniques has predicted greater symptom reduction following CBT for depression; DeRubeis & Feeley, 1990). Yet, other such studies have found (a) no relationships between adherence and outcome (Barber, Crits-Christoph, & Luborsky, 1996; Elkin, 1988); (b) symptom change influencing adherence (e.g., early improvement predicting greater adherence to psychodynamic technique; Barber et al., 1996); and (c) adherence interacting with factors such as therapist competence and therapeutic alliance (e.g., adherence to a drug counseling protocol only predicting better outcomes when the therapeutic alliance was poor; Barber et al., 2006; Shaw et al., 1999).

A limitation of these adherence studies is that they depend on preexisting therapy manuals. Many therapies conducted in private practice settings and elsewhere lack manuals but nonetheless have a coherent model of ideal therapy process (Bohart, 2000). Ablon and Jones (1998, 2002) showed that adherence to theory-specific models of ideal therapy process in unmanualized or manualized therapies can be assessed using *prototype methodology*. Archived brief psychodynamic (PDT), CBT, and IPT therapies were assessed using a pantheoretical and multidimensional psychotherapy process measure (The Psychotherapy Process Q-set [PQS]; Ablon & Jones, 1999; Jones, Parke, & Pulos, 1992; Jones & Pulos, 1993). Experts on each type of therapy independently used the PQS to rate a hypothetical "ideal psychotherapy session" according to their therapy orientation. These ideal ratings were then aggregated within experts of the same theoretical orientation to yield a generic "formula" (or prototype) for each therapy type. Adherence scores were derived by correlating each orientation's prototype ratings with the process ratings given to the archived therapy sessions. Results showed that (a) different forms of therapy can generate statistically distinct prototypes (e.g., the PDT and CBT experts loaded on different factors; Ablon & Jones, 1998); (b) specific forms of therapy can contain more ingredients belonging to other forms of therapy than ingredients belonging to their own form of therapy (e.g., IPT was found to adhere more closely to

ideal CBT process than to ideal IPT process; Ablon & Jones, 2002); (c) greater adherence to the therapist's ascribed model can predict better outcomes (e.g., greater adherence to ideal PDT process predicted better outcomes in PDT; Ablon & Jones, 1998, 2002). Interestingly, these investigators also found that sometimes greater adherence to other therapy models may be even more predictive of better outcomes than adherence to the therapists' ascribed model. In one archival data set, outcome following CBT was more consistently associated with adherence to the PDT prototype than adherence to the CBT prototype (Ablon & Jones, 1998).

### Studying the Relationship Between Adherence and Session Outcomes in Single Cases

Because the prototype approach does not require preexisting manuals, it may be especially well suited for examining links between adherence and session outcomes in unmanualized single cases. Pole, Ablon, O'Connor, and Weiss (2002) applied a simplified prototype approach to study this relationship in a brief *control mastery therapy* (CMT) (Weiss, 1993). CMT is a coherent but unmanualized system of psychotherapy that integrates psychodynamic and cognitive behavioral theories with other unique ideas (Pole & Bloomberg-Fretter, 2006). CMT is psychodynamic in the sense that it emphasizes unconscious mental processes and posits that early childhood experiences contribute importantly to the patient's presenting complaints. CMT is cognitive in that it considers irrational beliefs to be a major "pathogen" and challenging them to be an important focus of treatment. In addition, CMT takes the view that therapeutic progress fundamentally depends on the client's appraisals of danger and safety during therapy rather than any particular techniques (Rappoport, 1997; Sampson, 1990), and thus many CMT theorists emphasize case specificity and eschew manualization (Fretter, Bucci, Broitman, Silberschatz, & Curtis, 1994; Silberschatz, Fretter, & Curtis, 1986). CMT advises therapists to carefully monitor the therapy process for clues about the client's danger level and to choose interventions accordingly. CMT theorists stipulate that feelings of danger arise because of irrational *unconscious pathogenic beliefs* that warn of negative consequences of pursuing normal developmental goals (e.g., seeking career advancement). Pathogenic beliefs are often associated with inappropriate *guilt* (O'Connor, Berry, & Weiss, 1999), which may not be readily apparent to the client but that is usually a focus of treatment. CMT anticipates that the client's primary strategy for modifying her or his pathogenic beliefs will be to *test* them in the context of the therapy relationship (Silberschatz & Curtis, 1993), sometimes in provocative ways. If the therapist "passes" the client's tests, then the client will ostensibly have a corrective emotional experience that undermines her or his pathogenic beliefs and increases her or his safety. Conversely, failed tests reinforce pathogenic beliefs and increase client danger (see O'Connor, 2002; Pole & Bloomberg-Fretter, 2006; or see Silberschatz, 2005, for more details).

Though CMT has received considerable empirical support from other sources (Caspar et al., 2000; Foreman, Gibbins, Grienberger, & Berry, 2000; Fretter et al., 1994; Messer, Tishby, Spillman, 1992; Norville, Sampson, & Weiss, 1996;

Pole & Jones, 1998; Silberschatz, 2005; Silberschatz & Curtis, 1993; Silberschatz et al., 1986). Pole et al. (2002) were the first to ask whether adherence to a CMT prototype was associated with session outcomes in a single case. A case-specific CMT prototype was created from ratings of a hypothetical ideal session for the specific patient. The prototype ratings, which emphasized therapist interventions focused on guilt, reassurance, and supportiveness, were compared with process ratings of the actual sessions. The results showed that greater adherence to the case-specific CMT prototype was associated with less client negative affect, stronger therapeutic alliance, and better session outcomes (Pole et al., 2002). Though supportive of CMT and the value of the prototype method for examining single cases, this study was limited by the case-specific prototype, which could not be readily generalized to other cases. To maximize the scientific value of single-case research, it is important that studies are designed to permit replication (Gottman, 1973; Hilliard, 1993). A generic CMT prototype similar to Ablon and Jones' (1998) psychodynamic and cognitive behavioral prototypes would be desirable for this purpose because it could be applied to multiple cases. Furthermore, because CMT contains psychodynamic and cognitive behavioral elements, it would be useful to know whether these elements contribute to outcomes in CMT treatments. To address these prior limitations and to extend this work to the study of long-term single cases, we undertook the present study.

### The Present Study

The present study had the following three aims: (a) to develop and establish the psychometric properties of a quantitative prototype of ideal CMT; (b) to determine whether a long-term control mastery treatment would adhere more closely to the CMT prototype than to the PDT or CBT prototypes; and (c) to determine whether adherence to the CMT prototype uniquely predicted symptom change. We selected an archival long-term CMT treatment that has been previously studied both qualitatively (Bloomberg-Fretter, 2005; Fretter, 1995) and quantitatively (Jones, Ghannam, Nigg, & Dyer, 1993; Pole & Jones, 1998), thereby forming a foundation on which to build the present work. For example, it is already known that (a) the therapist was active and didactic and that the treatment focused on cognitive themes (Jones et al., 1993); (b) the patient showed clinically significant improvement and maintained her gains for at least 2 years posttreatment (Jones et al., 1993); (c) case-specific CMT process measures predicted symptom change (Pole & Jones, 1998); and (d) a case-specific PDT process measure did not directly predict symptom change (Jones et al., 1993) but was indirectly predictive of change (Pole & Jones, 1998). On the basis of the previous findings on this case and other previous literature, we hypothesized that (a) this treatment would show greater adherence to the CMT prototype than to either the CBT or PDT prototypes, and (b) adherence to the CMT prototype would predict symptom change, whereas adherence to the PDT prototype would not. There was insufficient basis to predict the relative level of CBT process in the treatment and whether it would predict symptom change. However, because cognitive interventions are included in CMT and be-

cause CBT has been shown to effectively treat depression (e.g., Elkin, 1994), it is reasonable to expect that substantial CBT process may be present in the CMT treatment and predictive of symptom change.

## Method

### Case

The case is part of the Berkeley Psychotherapy Research Project archives and has been described in detail elsewhere (see Bloomberg-Fretter, 2005; Fretter, 1995; Jones et al., 1993; Pole & Jones, 1998, for further information). Briefly, at the onset of therapy, the patient, Ms. M, was a 35-year-old middle-class Caucasian who met *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (*DSM-IV*; American Psychiatric Association, 1994) criteria for recurrent major depressive disorder but no other Axis I or Axis II disorders. She was referred for long-term therapy because previous shorter therapies never seemed to get to the root of the problem. It was anticipated that longer term therapy would allow time to clarify connections between her childhood adversities and her recurring depressions and permit repeated corrective emotional experiences with her therapist. The client provided written informed consent to participate in this Internal Review Board-approved study. At intake, she showed elevated psychiatric distress on many measures, including the Beck Depression Inventory (BDI; Beck, Steer, & Garbin, 1988) ( $BDI = 24$ ) and the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967) ( $HRSD = 22$ ). She was treated with CMT for 208, twice-weekly, 50-min sessions over a 2.5-year period in the University of California, Berkeley psychological clinic. Her therapist was a psychologist in full-time private practice with a decade of experience researching and practicing CMT. The outcome was successful, as evidenced by clinically significant reductions (Jacobson & Truax, 1991) in depression at termination (e.g., her termination scores were  $BDI = 1$  and  $HRSD = 1$ ); continued remission through the 6-, 12-, and 24-month posttreatment evaluations (Jones et al., 1993); and broader improvements in her self-concept, relationships, and understanding of the factors that triggered and maintained her symptoms (Bloomberg-Fretter, 2005; Fretter, 1995; Pole & Jones, 1998).

### Measures

*The Psychotherapy Process Q-set* (Jones, Hall, & Parke, 1991). The PQS is an observer-based psychotherapy process measure consisting of 100 cards describing (a) therapist behaviors ( $n = 41$ ), (b) client behaviors ( $n = 40$ ), and (c) therapist-client interactions ( $n = 19$ ) that might occur in an individual adult psychotherapy session. The card items were selected to be applicable to a wide range of theoretical orientations. After inspecting a therapy session, independent judges sorted the PQS cards into a nine-category normal distribution. Category 1 comprises the five items that are most extremely uncharacteristic of the session. Higher categories are used for items that are increasingly characteristic of the session, with Category 9 re-

served for the five items that are most extremely characteristic of the session.

*The Symptom Checklist 90-Revised (SCL-90-R; Deorgatis & Savitz, 2000).* The SCL-90-R is a widely used self-report psychotherapy change measure assessing 90 psychiatric symptoms on a 5-point Likert-type scale ranging from 0 (*not at all*) to 4 (*extremely*), indicating how much the client had "been distressed" by the symptom within the past 7 days. The SCL-90-R has shown internal consistency ranging from .79 to .90, test-retest reliability ranging from .70 to .90, and adequate convergent and discriminant validity. The Global Severity Index (GSI; Derogatis & Savitz, 2000), which is the mean rating across all 90 items, summarizes the client's general psychiatric symptom severity.<sup>1</sup> The client showed clinically significant reduction in GSI by the end of therapy (Jones et al., 1993).

### Procedures

A team of eight independent judges (clinical psychology graduate students and research-oriented clinicians) were trained to achieve at least .70 interrater reliability using the PQS. Training involved reading the PQS manual, discussing each item with the measure's author (E. E. Jones), and discussing ratings of multiple videotaped therapy sessions. The judges were then randomly assigned in pairs to rate videotapes of every fourth session of Ms. M's treatment ( $n = 53$ ) in a random order. Judges met weekly to discuss their ratings, examine reliability, and prevent rater drift. When the correlation between two judges' ratings was below .50, a third judge was added. The judges' PQS ratings of each session were averaged together to form a composite measure of the "actual process" used in subsequent analyses. Because these mean ratings were used to operationalize the actual process, interjudge reliability was assessed using the intraclass correlation coefficient (ICC) estimate of the reliability of the mean of the  $k$  judges' ratings ( $ICC_{3,k}$ ; Shrout & Fleiss, 1979). The average PQS interjudge reliability across sessions was .81 (range = .66–.92).

The client completed SCL-90-R ratings at the beginning of therapy and every 16 sessions thereafter, resulting in 14 SCL-90-R GSI (symptom) assessments. The therapist was kept blind as to the client's symptom scores. Unfortunately, to accomplish the time series analysis (described in the *Data Analysis* section), a contemporaneous symptom score was required for each of the 53 PQS process assessments. The missing symptom scores were estimated using the same SPSS linear interpolation (LINT) procedure followed by Jones et al. (1993) and Pole and Jones (1998). In this procedure, the average of two consecutive observed symptom scores were used to estimate the missing midpoint between them. The estimated midpoint was then averaged with the observed score before and after it to estimate additional missing data points between them. It is unknown how well these estimated data points represent what the patient's symptom scores would have been if they were directly assessed. Consequently, we refer to these scores as "estimated symptom scores." We took additional steps to evaluate the validity of our estimated scores (described in Footnote 2).

Prototypes were developed for the present study by combining new expert PQS ratings of an ideal CMT session ( $n = 9$ ) with previously obtained expert PQS ratings of ideal PDT ( $n = 11$ ) and ideal CBT ( $n = 10$ ) sessions (Ablon & Jones, 1998). All of the experts were internationally recognized leading authorities on their

respective theories and highly experienced as practitioners, scholars, and teachers of their school of therapy. Most had published extensively about their approach to psychotherapy, some as the progenitor of their theory. None of the experts were involved in rating the present case. Each expert was given an opportunity to indicate whether important aspects of their ideal therapy process were missing from the 100 PQS items. None indicated important omissions. The level of interrater reliability was high within PDT ( $\alpha = .94$ ), CBT ( $\alpha = .95$ ), and CMT ( $\alpha = .90$ ) experts. A Q-mode factor analysis (Fiedler, 1951; Stephenson, 1953) was applied, which transposed the data so that each of the 30 experts was treated as a variable and each PQS item was treated as a case. The resulting data were then subjected to a principal components analysis followed by varimax rotation yielding three factors with eigenvalues above 1.0, which together explained 71% of the variation among experts. Experts of the same orientation loaded primarily on the same factor, and experts of different orientations loaded primarily on different factors. The average primary factor loading was .77 (range = .66–.88) for the psychodynamic therapists, .77 (range = .64–.86) for the cognitive behavioral therapists, and .70 (range = .59–.84) for the control mastery therapists. To determine the contribution of each PQS item to each orientation's prototype, factor scores were created using linear regression. Higher factor scores indicate that the PQS item descriptor is more characteristic of a particular psychotherapy system. The 100 factor scores within each theoretical orientation comprise the quantitative prototype of the orientation (see Appendix A, which is available as supplemental material online). The PDT and CBT prototypes have shown good face validity, content validity, and criterion validity in previous work (Ablon & Jones, 1998).

Adherence scores for each session were initially calculated in accordance with Ablon and Jones' (1998) methods. The factor scores associated with the set of 100 PQS items for each prototype were correlated with the corresponding composite observer PQS ratings for each session, yielding one adherence score for each prototype per session. We refer to these as *full* adherence scores reflecting adherence to ideal therapy process rather than only ideal therapist technique. Criterion validity of the full CBT and PDT adherence scores is supported by findings of higher PDT scores in archived PDT and higher CBT scores in archived CBT (Ablon & Jones, 1998). Though archived CMT cases were unavailable to support the validity of the full CMT adherence scores, we found such support in the correlation between these scores and a factor of PQS items shown by Pole and Jones (1998) to reflect case-specific CMT-prescribed process,  $r(51) = .60, p < .001$ . To provide a

<sup>1</sup> Consistent with previous research on this case (e.g., Jones et al., 1993; Pole & Jones, 1998), the GSI of the SCL-90-R was chosen to operationalize symptom change. We considered using a measure that directly targeted the client's major depression diagnosis (e.g., the BDI). However, we found that such measures showed a floor effect within the first few months of treatment (see Jones et al., 1993) in accord with evidence that depression symptoms can substantially resolve within 16–20 psychotherapy sessions (e.g., Elkin, 1994). Yet, such benefits of brief therapy can also be highly susceptible to relapse (Shea et al., 1992). The SCL-90-R GSI scores, however, showed greater variability over the course of the 2.5-year treatment, which we believed captured the broad array of psychological symptoms that resolved more slowly in response to long-term psychotherapy and may have contributed to the client's maintenance of therapeutic benefits for more than 2 years posttreatment.

means of determining which PQS items contributed to the “full” adherence score findings, we extended the Ablon and Jones (1998, 2002) method by calculating component adherence scores for subsets of PQS items addressing therapist behavior ( $n = 41$ ), (b) client behavior ( $n = 40$ ), and (c) the therapist–client interaction ( $n = 19$ ). Therapist behavior and client behavior items clearly referenced the actor (e.g., “Therapist is tactless” or “Patient is anxious or tense”). Items classified as therapist–client interaction items were not specific as to the actor and often referenced the topic of discussion (e.g., “Love or romantic relationships are discussed”) (see Appendix A for complete information on category assignment). Three component adherence scores were calculated for each theoretical orientation and for each session, yielding nine component adherence scores per session. All adherence correlations were  $z$ -transformed to increase normality prior to statistical analysis but are presented in terms of Pearson’s  $r$  for ease of interpretation.

### Data Analysis

To determine whether Ms. M’s therapy process adhered more closely to the CMT prototype than to the other therapy prototypes, we compared the full and component adherence scores using paired  $t$  tests. These analyses and the creation of the prototypes were accomplished with SPSS 14.0. To test whether adherence to the prototypes predicted estimated symptom scores, we applied Gottman and Ringland’s (1981) bivariate time series analysis using BIVAR software (Williams & Gottman, 1982). This procedure has been used in the past to model important social relationships such as mother–infant (Gottman & Ringland, 1981) and husband–wife (Levenson & Gottman, 1985) interactions. Both Jones et al. (1993) and Pole and Jones (1998) showed that this approach can be applied to assess therapist–patient interactions. Though this method lacks the credibility of causal inference afforded by random assignment in group designs, it can make predictive inferences by capitalizing on unique properties of time series data. Multiple observations of the same process and session outcome variables unfolding over time provide the investigator with information about *when* outcome changes occurred in relation to process changes. Thus, the direction of influence between process and session outcomes can be implied by precedence. Moreover, though the Gottman and Ringland (1981) procedure cannot rule out the causal influence of unmeasured variables, it can make a stronger case for the directionality of effects than mere correlation.

The Gottman and Ringland (1981) procedure determines whether the *cross-regression* between the process and session outcome variables (e.g., the capacity of *past adherence* values to predict *present symptom* values) exceeds the *autoregression* within the symptom variable (e.g., the capacity of *past symptom* values to predict *present symptom* values). If so, then the process variable contributes to the prediction of the session outcome variable. The bivariate time series analysis also checks the possibility that session outcome values influence process scores by switching the order of predictor and criterion variables in the analysis. Bidirectional or reciprocal influence is implied if both sets of analyses show predictive influences. As applied to the present data, pairs of adherence and estimated symptom scores were conceptualized as occurring contemporaneously and at regular time intervals (i.e.,

every fourth session). The only assumption of the Gottman and Ringland procedure is that both time series are stationary (i.e., show similar mean and variability over time) or can be made stationary through transformation. Because the stationarity assumption was not met in these data (e.g., psychiatric symptoms decreased over time), a difference transformation was applied to each time series (see Appendix B, which is available as supplemental material online). Each data point in the series was subtracted from the point that followed it, resulting in a new slightly shorter time series (i.e.,  $n = 52$ ).

## Results

### Description of CBT, PDT, and CMT Prototypes

Table 1 presents the 10 PQS items that were judged by the experts to be most characteristic of ideal PDT, CBT, and CMT. Each item is accompanied by its factor score and a notation indicating whether it describes therapist behavior, client behavior, or the therapist–client interaction. The absence of a particular item in Table 1 does not mean that the experts judged the item to be uncharacteristic of their school of therapy. It only means that the experts did not consider that item among *the 10 most important* features of an ideal session of their type of therapy. Each prototype actually includes all 100 PQS items but with different factor score weightings reflecting the different emphases given by each orientation (see Appendix A for all 100 items).

According to the cognitive behavioral experts, ideal CBT is most strongly characterized by a focus on such topics as homework, belief systems, treatment goals, and the patient’s recent (rather than early) life circumstances. The prototypical CBT therapist actively structures the interaction, is supportive, asks for elaboration, gives explicit advice or guidance, and encourages the patient to try new behaviors outside of the session. According to the psychodynamic experts, ideal PDT is most strongly characterized by a focus on the patient’s dreams, fantasies, and/or the therapy relationship. The patient should ideally achieve new insights as the therapist clearly interprets defenses, transferences, and unconscious mental processes and maintains empathy, neutrality, and nonjudgmental acceptance. For control mastery experts, “accurately perceiving the therapy process” is most important in order to respond appropriately to potentially provocative client “tests.” Beyond this, ideal CMT gives high priority to clarifying the client’s goals by discussing aspirations and ambitions; making unconscious barriers explicit by discussing dreams, fantasies, and unrecognized guilt; challenging pathogenic beliefs by actively distinguishing reality from fantasy and suggesting alternative meanings of others’ behavior; and providing direct reassurance, support, and clear communication to facilitate the client’s sense of safety.

### Adherence to Prototypes in the Case of Ms. M

Descriptive statistics summarizing the extent to which Ms. M’s actual therapy process conformed to the three full prototypes and their components are presented in Table 2. On average, the sessions moderately resembled the full CBT ( $r = .39$ ) and full CMT ( $r = .27$ ) prototypes but did not resemble the full PDT ( $r = .01$ ) prototype. Paired  $t$  tests comparing the full adherence scores revealed that the sessions were significantly closer to ideal CBT process than ideal CMT process,  $t(52) = 7.82, p < .001, d = 1.00$ ,

Table 1  
*Ten Most Characteristic Psychotherapy Process Q-Set (PQS) Items in Cognitive Behavioral, Psychodynamic, and Control Mastery Prototypes Ranked by Factor Score*

Most characteristic cognitive behavioral items	Factor score	Most characteristic psychodynamic items	Factor score	Most characteristic control mastery items	Factor score
There is discussion of specific activities or tasks for the P to attempt outside of session (PQS#38). <sup>c</sup>	1.93	P's dreams or fantasies are discussed (PQS#90). <sup>c</sup>	1.71	T accurately perceives therapy process (PQS#28). <sup>a</sup>	2.11
Discussion centers on cognitive themes (e.g., ideas or belief systems) (PQS#30). <sup>c</sup>	1.68	T is neutral (PQS#93). <sup>a</sup>	1.57	T focuses on guilt (PQS#22). <sup>a</sup>	1.91
P's treatment goals are discussed (PQS#4). <sup>c</sup>	1.51	T points out P's use of defenses (PQS#36). <sup>a</sup>	1.53	T is directly reassuring (PQS#66). <sup>a</sup>	1.69
T encourages P to try new ways of behaving with others (PQS#85). <sup>a</sup>	1.49	T interprets transference (PQS#100). <sup>a</sup>	1.47	T communicates in a clear coherent style (PQS#46). <sup>a</sup>	1.61
T actively exerts control over the interaction (e.g., structuring, introducing new topics) (PQS#17). <sup>a</sup>	1.45	T is sensitive to P's feelings, attuned, empathic (PQS#6). <sup>a</sup>	1.46	P is provocative, tests the limits of the therapy relationship (PQS#20). <sup>b</sup>	1.55
T adopts a supportive stance (PQS#45). <sup>a</sup>	1.43	T interprets unconscious wishes, feelings, or ideas (PQS#67). <sup>a</sup>	1.43	P's dreams or fantasies are discussed (PQS#90). <sup>c</sup>	1.45
Dialogue has a specific focus (PQS#23). <sup>c</sup>	1.38	T conveys nonjudgmental acceptance (PQS#18). <sup>a</sup>	1.38	T suggests the meaning of other's behavior (PQS#43). <sup>a</sup>	1.41
T asks for more information or elaboration (PQS#31). <sup>a</sup>	1.37	P achieves a new understanding or insight (PQS#32). <sup>b</sup>	1.32	T adopts a supportive stance (PQS#45). <sup>a</sup>	1.33
P's recent life situation is emphasized in discussion (PQS#69). <sup>c</sup>	1.35	The therapy relationship is discussed (PQS#98). <sup>c</sup>	1.28	P's aspirations/ambitions are discussed (PQS#41). <sup>c</sup>	1.32
T gives explicit advice and guidance (PQS#27). <sup>a</sup>	1.32	T communicates in a clear, coherent style (PQS#46). <sup>a</sup>	1.24	T actively distinguishes reality from fantasy (PQS#68). <sup>a</sup>	1.26

*Note.* PQS item numbers are given in parentheses following PQS item content. Each prototype is actually composed of all 100 PQS items but with different factor scores reflecting the different emphasis given by the experts. The absence of a particular process in a given prototype in this table should not be taken to mean that the experts judged that process to be unimportant. It only means that the experts did not consider it among the 10 most important features of the therapy.

T = Therapist. P = Patient.

<sup>a</sup> PQS therapist behavior item. <sup>b</sup> PQS client behavior item. <sup>c</sup> PQS therapist–client interaction item.

or ideal PDT process,  $t(52) = 14.08$ ,  $p < .001$ ,  $d = 2.30$ , and significantly closer to ideal CMT process than to ideal PDT process,  $t(52) = 12.06$ ,  $p < .001$ ,  $d = 1.58$ .

Examination of the component adherence scores revealed a more nuanced picture. Paired  $t$  tests showed that the therapist's adherence to ideal CMT behaviors ( $r = .40$ ) was significantly higher than her adherence to either ideal CBT behaviors ( $r = .35$ ),  $t(52) = 2.58$ ,  $p < .01$ ,  $d = 0.34$ , or ideal PDT behaviors ( $r = -.06$ ),  $t(52) = 14.57$ ,  $p < .001$ ,  $d = 1.94$ , and her adherence to ideal CBT behaviors was significantly higher than her adherence to ideal PDT behaviors,  $t(52) = 10.51$ ,  $p < .001$ ,  $d = 1.75$ . However, the client's behavior most closely approximated ideal CBT behaviors ( $r = .46$ ) as compared with either ideal CMT ( $r = .04$ ),  $t(52) = 23.24$ ,  $p < .001$ ,  $d = 2.27$ , or ideal PDT behaviors ( $r = .27$ ),  $t(52) = 11.41$ ,  $p < .001$ ,  $d = 1.08$ . In fact, the client's behavior was significantly closer to ideal PDT than to ideal CMT,  $t(52) = 9.21$ ,  $p < .001$ ,  $d = 1.35$ . Finally, with regard to therapist–client interaction items, adherence to both ideal CMT ( $r = .31$ ),  $t(52) = 20.31$ ,  $p < .001$ ,  $d = 2.17$ , and ideal CBT ( $r = .24$ ),  $t(52) = 7.14$ ,  $p < .001$ ,  $d = 1.82$ , significantly exceeded adherence to ideal PDT ( $r = -.18$ ), but adherence to ideal

CMT was not significantly different from adherence to ideal CBT,  $t(52) = 1.47$ ,  $ns$ ,  $d = 0.33$ .

#### *Did Adherence to the Full or Component Prototypes Predict Estimated Symptom Change?*

The results of bivariate time series analyses examining the relationships between the full prototype adherence scores and estimated symptom change scores are presented in Table 3. The middle three columns (A, B, and  $SSE_{GSI-EST}$ ) of the table describe and compare four types of regression models that were constructed to test whether the adherence scores predicted change in estimated symptom severity. The right three columns (C, D, and  $SSE_{CBT}$ ) of the table check for bidirectional effects by constructing models to test whether estimated symptom severity scores predicted change in the adherence scores. The four types of models comprising each bivariate time series analysis differ in their number of autoregressive terms (enumerated under Columns A and C) and cross-regressive terms (enumerated under Columns B and D). Each term corresponds to an observation (or "lag") into the past. In this study, each lag corresponds to four

Table 2  
Descriptive Statistics of Cognitive Behavioral, Psychodynamic, and Control Mastery Therapy Prototype Adherence Scores in the Case of Ms. M (N = 53)

Type of therapy	Full prototype				Therapist behaviors				Client behaviors				Therapist–client interaction			
	Min	Max	M	SD	Min	Max	M	SD	Min	Max	M	SD	Min	Max	M	SD
CBT	.07	.59	.39 <sub>a</sub>	.12	.11	.61	.35 <sub>a</sub>	.14	-.25	.80	.46 <sub>a</sub>	.19	-.34	.67	.24 <sub>a</sub>	.22
PDT	-.39	.46	.01 <sub>b</sub>	.20	-.57	.55	-.06 <sub>b</sub>	.30	-.08	.51	.27 <sub>b</sub>	.16	-.65	.26	-.18 <sub>b</sub>	.24
CMT	.01	.53	.27 <sub>c</sub>	.12	.06	.67	.40 <sub>c</sub>	.15	-.44	.49	.04 <sub>c</sub>	.18	-.19	.67	.31 <sub>a</sub>	.21

Note. Adherence scores are Pearson correlations (*r*) between composite observer Psychotherapy Process Q-set (PQS) ratings of each session and prototypes derived from expert PQS ratings of a hypothetical ideal session from the perspectives of cognitive behavior therapy (CBT), psychodynamic therapy (PDT), and control mastery therapy (CMT). Full prototype refers to adherence scores involving all 100 PQS items. Therapist behaviors refers to adherence scores involving PQS items describing therapist behaviors (*n* = 41). Client behaviors refers to adherence scores involving PQS items describing client behaviors (*n* = 40). Therapist–client interaction refers to adherence scores involving PQS items describing qualities of the therapist–client interaction, including topics of discussion (*n* = 19). See Appendix A, which is available as supplemental material online, for a complete listing of PQS items contributing to these three categories. Means in the same column that do not share subscripts significantly differ on the basis of paired *t* tests at *p* < .05 or less. Min = minimum value in sampled sessions. Max = maximum value in sampled sessions. *M* = mean value across sampled sessions. *SD* = standard deviation across sampled sessions.

Table 3  
Summary of Bivariate Time Series Analyses Predicting Estimated Symptom Change (GSI-EST) With Cognitive Behavior Therapy (CBT), Psychodynamic Therapy (PDT), and Control Mastery Therapy (CMT) Full Prototype Adherence Scores

Regression model	Regression models predicting EST with CBT			Regression models predicting CBT with EST		
	A	B	SSE <sub>GSI-EST</sub>	C	D	SSE <sub>CBT</sub>
1: Oversized AR + CR	10	10	.07	10	10	.52
2: Optimal AR + CR	6	8	.08	5	0	.71
3: Optimal AR	6	0	.13	5	0	.71
4: Oversized AR	10	0	.10	10	0	.64
Model 1 vs. Model 2	<i>Q</i> (6) = 5.87, <i>ns</i> , <i>z</i> = -0.04			<i>Q</i> (15) = 13.3, <i>ns</i> , <i>z</i> = 0.32		
Model 2 vs. Model 3	<i>Q</i> (8) = 20.4, <i>p</i> < .01, <i>z</i> = 3.10			<i>Q</i> (0) = 0, <i>ns</i> , <i>z</i> = 0		
Model 3 vs. Model 4	<i>Q</i> (4) = 8.79, <i>ns</i> , <i>z</i> = 1.70			<i>Q</i> (5) = 4.35, <i>ns</i> , <i>z</i> = -0.21		
Conclusions	CBT predicts GSI-EST.			GSI-EST does not predict CBT.		
Regression model	Regression models predicting EST with PDT			Regression models predicting PDT with EST		
	A	B	SSE <sub>GSI-EST</sub>	C	D	SSE <sub>PDT</sub>
1: Oversized AR + CR	10	10	.07	10	10	1.38
2: Optimal AR + CR	9	4	.09	3	0	1.96
3: Optimal AR	9	0	.10	3	0	1.96
4: Oversized AR	10	0	.10	10	0	1.71
Model 1 vs. Model 2	<i>Q</i> (7) = 10.28, <i>ns</i> , <i>z</i> = 0.88			<i>Q</i> (17) = 14.6, <i>ns</i> , <i>z</i> = -0.41		
Model 2 vs. Model 3	<i>Q</i> (4) = 7.06, <i>ns</i> , <i>z</i> = 1.08			<i>Q</i> (0) = 0, <i>ns</i> , <i>z</i> = 0.00		
Model 3 vs. Model 4	<i>Q</i> (1) = 0.01, <i>ns</i> , <i>z</i> = -0.70			<i>Q</i> (7) = 5.83, <i>ns</i> , <i>z</i> = -0.31		
Conclusions	PDT does not predict GSI-EST.			GSI-EST does not predict PDT.		
Regression model	Regression models predicting EST with CMT			Regression models predicting CMT with EST		
	A	B	SSE <sub>GSI-EST</sub>	C	D	SSE <sub>CMT</sub>
1: Oversized AR + CR	10	10	.07	10	10	.52
2: Optimal AR + CR	5	8	.09	3	1	.70
3: Optimal AR	5	0	.13	3	0	.76
4: Oversized AR	10	0	.10	10	0	.65
Model 1 vs. Model 2	<i>Q</i> (7) = 11.26, <i>ns</i> , <i>z</i> = 1.14			<i>Q</i> (16) = 12.8, <i>ns</i> , <i>z</i> = -0.56		
Model 2 vs. Model 3	<i>Q</i> (8) = 17.72, <i>p</i> < .05, <i>z</i> = 2.43			<i>Q</i> (1) = 2.94, <i>ns</i> , <i>z</i> = 1.37		
Model 3 vs. Model 4	<i>Q</i> (5) = 9.67, <i>ns</i> , <i>z</i> = 1.45			<i>Q</i> (7) = 6.41, <i>ns</i> , <i>z</i> = -0.16		
Conclusions	CMT predicts GSI-EST.			GSI-EST does not predict CMT.		

Note. Values under Columns A and C represent the number of autoregressive (AR) terms in each model. Values under Columns B and D represent the number of cross-regressive (CR) terms in each model. Models 1 and 4 are oversized models with an arbitrarily large number of AR and CR terms (conventionally set at 10). Model 2 is the optimal model combining both AR and CR terms. Model 3 is the optimal model containing only AR terms. SSE = Sum of squares for error (unexplained error variance when a given model is applied).

sessions in the past because observations were made every fourth session. The number of autoregressive and cross-regressive terms in each model indicates the number of lags used to explain the data. The amount of variance left unexplained by each model is represented by values in the columns labeled *SSE* (*sum of squares error*). Model 1 is an “oversized” model combining more autoregressive and cross-regressive terms than should be necessary to explain the data. Conventionally, 10 lags into the past are used to construct the oversized model. The time series is thus effectively shortened by 10 observations (i.e., the 11th observation in the time series is the first observation predicted by each model). Model 2 is constructed by successively removing autoregressive and cross-regressive terms from the oversized model until an “optimal” combined model is achieved that minimizes the number of autoregressive and cross-regressive terms needed to predict the series without significantly changing the amount of explained variance. Model 3 is constructed by dropping the cross-regressive terms from the model to achieve an “optimal” purely autoregressive model (one with the minimal number of terms). Finally, Model 4 determines how much variance would be explained by an oversized purely autoregressive model (one with 10 terms). As a validity condition for the main analysis, the optimal models must not significantly differ from the oversized models in terms of unexplained variance. Pairwise comparisons of the variance left unexplained by each model are made using likelihood ratio tests, each yielding a  $Q$  statistic. This statistic has a chi-square distribution and is evaluated at the  $p < .05$  significance level (two-tailed). The analysis also yields a  $z$ -score, which can be used as an indicator of effect size. A nonsignificant difference between Models 1 and 2 indicates that the smallest combined model has equivalent explanatory power to the oversized combined model. Thus, the simpler (smaller) model is preferred (Model 2). A nonsignificant difference between Models 3 and 4 indicates that the smallest purely autoregressive model has equivalent explanatory power to the oversized autoregressive model. Thus, the simpler model is preferred (Model 3). The heart of the analysis is the comparison between Model 2 (the optimal combined model) and Model 3 (the optimal autoregressive model). If there is a significant difference between these models, then one series is said to contribute to the prediction of the other because significantly more variance is explained by considering the other series than would be explained without it (see Gottman & Ringland, 1981, and Appendix B for further details).

After applying these methods, we found that both adherence to the full CBT prototype,  $Q(8) = 20.4$ ,  $p < .01$ ,  $z = 3.10$ , and adherence to the full CMT prototype,  $Q(8) = 17.7$ ,  $p < .05$ ,  $z = 2.43$ , predicted estimated symptom scores, but adherence to full PDT prototype did not,  $Q(4) = 7.06$ ,  $ns$ ,  $z = 1.08$ . When the analyses were repeated to check for bidirectional effects, the results showed that estimated symptom scores did not predict adherence to the full CBT,  $Q(0) = 0$ ,  $ns$ ,  $z = 0.00$ ; CMT,  $Q(1) = 2.94$ ,  $ns$ ,  $z = 1.37$ ; or PDT,  $Q(0) = 0$ ,  $ns$ ,  $z = 0.00$ , prototypes<sup>2</sup> (see Table 3). We repeated this procedure with the component adherence scores. To conserve space, we only present the comparisons between Models 2 and 3 here. It should be noted that all of the validity tests yielded the expected results (i.e., the differences between Models 1 and 2 and between Models 3 and 4 were not significant). We found that among the CMT prototype components, estimated symptom change was significantly predicted by adherence in therapist behaviors,  $Q(5) = 18.98$ ,  $p < .01$ ,  $z = 4.42$ , and client behaviors,  $Q(5) = 18.37$ ,  $p < .01$ ,  $z = 4.23$ , and

exhibited a trend toward being predicted by adherence to the therapist–client interaction items,  $Q(1) = 4.75$ ,  $p = .06$ ,  $z = 2.66$ . Within the CBT prototype, estimated symptom change was significantly predicted by adherence to therapist–client interaction items,  $Q(5) = 15.07$ ,  $p < .05$ ,  $z = 3.19$ , showed a trend toward being predicted by adherence to therapist behavior items,  $Q(5) = 12.49$ ,  $p = .06$ ,  $z = 2.37$ , and not significantly predicted by adherence to client behavior items,  $Q(5) = 10.63$ ,  $ns$ ,  $z = 1.78$ . For the PDT prototype, estimated symptom change was significantly predicted by adherence to both client behavior items,  $Q(9) = 27.70$ ,  $p < .01$ ,  $z = 4.41$ , and therapist–client interaction items,  $Q(10) = 27.33$ ,  $p < .01$ ,  $z = 3.88$ , but not significantly predicted by adherence to therapist behavior items,  $Q(7) = 13.59$ ,  $ns$ ,  $z = 1.76$ . When the analyses were reversed, estimated symptoms scores did not significantly predict any of the component adherence scores.<sup>3</sup> In summary, the time series analyses indicated that Ms. M's symptom change was significantly driven by the therapist's adherence to ideal CMT behaviors, the client's adherence to ideal CMT and PDT behaviors, and therapist–client interaction consistent with ideal CBT and ideal PDT.

## Discussion

The control mastery prototype introduced in this article was accompanied by good psychometric properties. The very high interrater reliability among CMT experts shows that, despite the case specificity emphasized by CMT researchers (Fretter et al.,

<sup>2</sup> We checked the validity of our SCL-90-R GSI score interpolation procedure in two ways. First, we identified a measure in the data set that was highly correlated with the 14 SCL-90-R GSI data points prior to imputation, conceptually related to the SCL-90-R GSI measure, and available for all 53 observations. The composite ratings of PQS Item 94 “Patient is sad or depressed” satisfied these criteria ( $r = .80$ ,  $p < .001$ ). After excluding this item from the calculation of the full adherence scores, we recomputed the time series analyses and found that greater adherence to the revised CBT prototype predicted the PQS Item 94 rating,  $Q(8) = 22.55$ ,  $p < .01$ ,  $z = 3.64$ , and greater adherence to the revised PDT prototype did not,  $Q(0) = 0.00$ ,  $ns$ ,  $z = 0.00$ . However, contrary to our original results, adherence to the revised CMT prototype did not predict the item rating,  $Q(7) = 11.0$ ,  $ns$ ,  $z = 1.07$ . As a second check, we repeated the full prototype time series analyses using only the 14 pairs of process and outcome measures that were actually observed (i.e., discarding 39 actual adherence score observations and the 39 imputed symptom scores). After shortening the number of start-up observations from 10 to 2 (to accommodate the shorter time series), we replicated our original findings. SCL-90-R GSI scores were predicted by the full CBT,  $Q(2) = 9.88$ ,  $p < .01$ ,  $z = 3.94$ , and full CMT,  $Q(2) = 8.82$ ,  $p < .05$ ,  $z = 3.41$ , adherence scores but not by the full PDT adherence scores,  $Q(2) = 5.54$ ,  $ns$ ,  $z = 1.77$ . Taken together, these follow-up analyses increase confidence in the linear interpolation procedure but cannot guarantee that our results were unaffected by our data estimation procedure.

<sup>3</sup> It may be surprising to some that none of our time series analyses yielded evidence of reciprocal effects (i.e., symptom changes influencing adherence to particular prototypes). This result runs counter to clinical experiences and some empirical evidence (e.g., Barber et al., 1996) that therapists may modify their techniques in response to changes in the client's symptoms (e.g., becoming more structured if the client's symptoms worsen). We speculate that the absence of such findings in the present study may have been due to the sampling interval of our study (i.e., increments of every fourth session). It is possible that future studies investigating consecutive sessions would detect more reciprocal effects.

1994; Silberschatz et al., 1986), experienced CMT therapists were able to reach consensus about what constitutes ideal “generic” CMT process. Second, despite the theoretical and empirical similarities among CMT, CBT, and PDT, their ideal descriptions were distinct enough to generate separate statistical factors. Finally, the overlap between the most characteristic items in our CMT prototype and both theoretical descriptions of CMT (e.g., Weiss, 1993) and a previously published case-specific CMT prototype (Pole et al., 2002) supports the validity of the generic prototype.

The prototype adherence scores featured in this article were not intended to measure adherence to a specific therapy manual. Rather, the scores indexed the extent to which the actual therapy process conformed to theoretical ideals articulated by each group of experts. We found that adherence to the CMT prototype was in the moderate range, which is in accord with therapist’s account that she flexibly applied CMT principles to this case (Bloomberg-Fretter, 2005; Fretter, 1995). Some have argued that such moderate adherence should be the rule when any therapy model is applied in practice because theoretical ideals must be adjusted to fit specifics of the case (Barber et al., 2006). Unlike conventional adherence scores (e.g., Hill et al., 1992), the full prototype adherence scores were not only limited to therapist behaviors but also assessed client behaviors and the therapist–client interaction. Component adherence scores, novel to the present study, clarified the full prototype results. For example, contrary to our hypothesis, the full prototype adherence scores suggested that Ms. M’s therapy most closely resembled ideal CBT. However, the component adherence scores revealed that this result was largely driven by the client’s behavior, which was most consistent with ideal CBT. According to data presented in Appendix A, this suggests that the client demonstrated an understanding of (PQS#72) and commitment to (PQS#73) therapy and did not seem controlling (PQS#87), passive (PQS#15), or wary (PQS#44). Similarly, the greater adherence of the therapist–client interaction items to ideal CMT and CBT than to ideal PDT implies that the sessions focused on the client’s treatment goals (PQS#4), personal aspirations (PQS#41), and belief system (PQS#30). Finally, the finding that the therapist’s behavior was closest to ideal CMT indicates that the therapist accurately perceived the therapy process (PQS#28), focused on guilt (PQS#22), actively differentiated reality from fantasy (PQS#68), and provided direct reassurance (PQS#66). From the conventional conceptualization of adherence, which emphasizes therapist behavior, this last result argues in support of our hypothesis by showing that the therapist most strongly adhered to ideal CMT.

The time series analyses were similarly elucidated by attention to the component scores. Our finding that adherence to the full CMT prototype predicted estimated symptom change was supplemented by evidence that both its therapist and client behavior components significantly contributed to this change. These results support our hypothesis, echo previous findings on this case on the basis of case-specific CMT measures (Pole & Jones, 1998), and are broadly congruous with other literature showing associations between recommended CMT processes and desired client changes (Fretter et al., 1994; Norville et al., 1996; Pole et al., 2002; Silberschatz & Curtis, 1993; Silberschatz et al., 1986). In fact, the component score findings might be interpreted as supporting earlier work indicating that both therapist behaviors (e.g., plan-compatible interventions; Silberschatz et al., 1986) and client behaviors (e.g., testing; Silberschatz & Curtis, 1993) contribute to

the efficacy of CMT. It is worth noting in this respect that adherence to ideal CMT client behavior predicted change even though such adherence was, on average, low in Ms. M’s treatment. This echoes a point made by Ablon and Jones (1998) that relatively rare processes can have clinically meaningful impact. Our discovery that adherence to the full CBT prototype also predicted estimated symptom change was similarly clarified by follow-up findings showing that only adherence to the CBT interaction items significantly predicted this change. These items included focusing on homework (PQS#38) and belief systems (PQS#30), both of which have been empirically associated with improvement in the cognitive behavioral treatment of depression (DeRubies & Feeley, 1990). Note also that though the client’s behavior was most consistent with ideal CBT, this component did not predict symptom change, indicating that the most conspicuous processes may not be the most effective processes. Finally, our findings that neither adherence to the full PDT prototype nor adherence to its therapist behavior component predicted estimated symptom change supported our hypothesis and mirrored Jones et al.’s (1993) finding based on a case-specific “psychodynamic technique” measure. However, adherence to ideal PDT client behaviors (e.g., the client achieving new insights) and therapist–client interactions (e.g., discussing the client’s dreams or fantasies) did predict such change. These findings also agree with previous work on this case showing that a focus on psychodynamic topics and increases in the client’s free association predicted estimated symptom change (Pole & Jones, 1998). Taken together, these results indicate that, among therapist behavior prototypes, the therapist’s adherence to ideal CMT behaviors was the driving force in this treatment. However, the client’s behavior and the therapist–client interaction also contributed to the client’s recovery through adherence to ideal CMT, CBT, and/or PDT processes. Yet, all of these time series results must be accepted with caution because of the large number of estimated symptom scores involved in this study. Though imputed scores were unavoidable given the requirements of the time series analysis and the limitations of this unique archival data set, and though the results were similar when the major analyses were repeated with the imputed values excluded, it is still possible that different results may have emerged if all symptom scores were available. This problem should be avoided in the future by obtaining symptom and process measures on the same schedule.

The validity of our symptom score estimation does, however, gain further support from the consistency of our results with several of Ablon and Jones’ findings relating full adherence scores to a broad range of unmodified outcome measures. For example, our finding that Ms. M’s CMT most closely adhered to the full CBT prototype runs parallel to Ablon and Jones’ (2002) finding that interpersonal psychotherapies also most closely adhered to the full CBT prototype. Similarly, our finding that adherence to the full CBT prototype predicted estimated symptom change in this CMT treatment is reminiscent of Ablon and Jones’ (1998) finding that adherence to the full PDT prototype was associated with favorable outcomes in CBTs. From a perspective that broadly defines adherence in terms of therapist behavior, client behavior, and their interaction, our results reinforce Ablon and Jones’ (1998, 2002) point that names of therapies can be misleading. Even when a therapist is using techniques consistent with her or his orientation, it is possible that unexpected processes belonging to other types of therapy may be wittingly or unwittingly present and

influencing outcomes. Yet, from a narrower perspective that sees adherence only in terms of the therapist's behavior, our results offer an alternative explanation of Ablon and Jones' earlier work and suggest that an examination of components might have yielded different conclusions.

This brings us to a broader discussion of the limitations associated with the procedures outlined in the present article. First, as the contrasts between the full and component adherence scores illustrate, consideration of only the full prototypes may obscure meaningful distinctions between its components. The prototypes differed in whether therapist behavior, client behavior, or therapist–client interaction items were rated as most characteristic. For example, whereas the CMT experts rated therapist behavior items as most important, the CBT experts selected therapist–client interaction items as most important. This could be why the full CMT adherence results were consistent with its therapist component and the full CBT adherence results were consistent with its therapist–client interaction component. Thus, we recommend following full prototype analyses with analyses addressing meaningful components. Second, the Q-sort methodology forces greater discrimination in ratings than might occur without it. Thus, there is a potential that ratings can be exaggerated by this method and might be more valid if straightforward Likert scales were used. Third, it is reasonable to raise the criticism that the prototypes may be too general to capture specific applications of theoretical models to particular clinical problems (e.g., CBT for depression vs. CBT for panic disorder) or for particular stages of therapy (e.g., PDT in the working vs. termination stage). This problem could be addressed by developing even more specialized prototypes addressing these variations. Fourth, given the emphasis that CMT has traditionally placed on case specificity and its deemphasis on specific techniques, the omission of case-specific elements and emphasis on specific therapist behaviors in our CMT prototype may be disconcerting to some. It is possible that the CMT prototype failed to adequately capture key aspects of CMT process (especially processes that would vary from case to case) or gave undue weight to technique over other subtle processes. If so, then this might further explain why some of the CMT adherence scores were low (especially those associated with client behaviors). We are somewhat reassured by the high agreement among CMT experts about ideal CMT process and by the fact that none of them indicated that important items were omitted from the CMT prototype, but this issue warrants further examination, perhaps by comparing the predictive power of generic and case-specific CMT prototypes within the same case. Fifth, the method does not incorporate a way of determining whether the therapist's competence (or lack thereof) might explain the adherence results (Barber et al., 2006; Shaw et al., 1999). In the present example, the therapist was highly experienced in practicing CMT, but one could legitimately ask whether she delivered CBT and PDT interventions competently and whether her level of competence with these techniques contributed to the results. Sixth, there are limitations in the Gottman and Ringland (1981) bivariate time series analysis itself. For example, this method cannot accommodate more than two variables at a time and, therefore, cannot simultaneously test between competing predictors. Furthermore, the procedure has not been widely used or thoroughly studied for potential statistical biases and shortcomings. Thus, it may well be that other time series approaches may prove more powerful for future applications. A

final basic limitation is one inherent to single-case research; namely, that the results cannot be assumed to generalize to other cases. For example, our results do not suggest that all or even most CMT treatments contain significant CBT elements. We merely show that a CMT treatment conducted by an experienced CMT practitioner can contain significant CBT elements and that these CBT elements can predict symptom change. We consider this to be an interesting result in its own right. However, to test its generalizability, "N of one research" must be replicated to achieve what Gottman (1973) called, "N-of-one-at-a-time research." The prototypes described can be applied to unlimited new cases to test the repeatability of our findings.

Limitations notwithstanding, we believe that this study makes several contributions. First, it replicates and extends previous single-case studies supporting CMT (Jones et al., 1993; Messer et al., 1992; Pole et al., 2002; Silberschatz & Curtis, 1993; Silberschatz et al., 1986) and previous studies of this case (Jones et al., 1993; Pole & Jones, 1998). Such replications are vital for a developing science and particularly important for single-case research (Gottman, 1973; Hilliard, 1993). Second, through expert prototypes and adherence scores, it offers new ways of assessing the integrity of widely used therapies (i.e., PDT and CBT) in situations in which specific therapy manuals are not available, and it also provides a means of making a lesser known but empirically "supported" therapy (CMT) more accessible to the field. The prototypes show that although some elements of all theoretical approaches are case specific, there are also general elements that may contribute importantly to understanding symptom change. Third, the bivariate time series analyses provide a sophisticated method of identifying mechanisms of change in long-term therapies, a topic that is virtually absent in the literature. In combination with the adherence scores, which assess the multivariate match between actual process and the optimal blending of therapy processes proposed by experts, a priori hypotheses may be tested. The time series analysis may also be used to test case-specific or post hoc research questions (e.g., Jones et al., 1993; Pole & Jones, 1998). Finally, though the methods outlined in this article may be applied to many theoretical orientations, they offer an especially important alternative to the "events paradigm" methods (Rice & Greenberg, 1984) typically used to study CMT. Rather than developing individualized formulations for every case (Curtis, Silberschatz, Sampson, & Weiss, 1994), the CMT prototype provides a generic "formulation" that can be used with any case. Rather than rating every therapist intervention and every client response for its consistency with CMT (e.g., Silberschatz et al., 1986) (a cumbersome task in long-term therapies) an entire session can be rated with the PQS in less than 2 hrs.

In conclusion, it has long been argued that there may be a gap between what therapists say they do and what they actually do in their consulting rooms (Buckley, Karasu, Charles, & Stein, 1979). The prototype method offers a window into the match between a therapist's professed theoretical model and his or her actual psychotherapy process, including a way of assessing whether the client's behavior and the therapist–client interaction conforms to theoretical ideals. When combined with statistical methods that are designed to assess predictive relationships between variables unfolding repeatedly over time, a further opportunity is opened to study process and outcome relationships in long-term psychotherapy. Furthermore, these questions can be addressed at the level of

the single case, thereby avoiding the expense and difficulty of the randomized clinical trial. Though such designs are not a substitute for randomized trials, they can make valuable contributions to a very sparse literature and address questions that otherwise might never be asked.

## References

- Ablon, J. S., & Jones, E. E. (1998). How expert clinicians' prototypes of an ideal treatment correlate with outcome in psychodynamic and cognitive-behavioral therapy. *Psychotherapy Research, 8*, 71–83.
- Ablon, J. S., & Jones, E. E. (1999). Psychotherapy process in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 67*, 64–75.
- Ablon, J. S., & Jones, E. E. (2002). Validity of controlled clinical trials of psychotherapy: Findings from the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry, 159*, 775–783.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Barber, J. P., Crits-Christoph, P., & Luborsky, L. (1996). Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *Journal of Consulting and Clinical Psychology, 64*, 619–622.
- Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M. E., Weiss, R. D., & Gibbons, M. B. C. (2006). The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: Results from the National Institute Drug Abuse Collaborative Cocaine Treatment Study. *Psychotherapy Research, 16*, 229–240.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years later. *Clinical Psychology Review, 8*, 77–100.
- Bloomberg-Fretter, P. (2005). Clinical use of the plan formulation in long-term psychotherapy. In G. Silberschatz (Ed.), *Transformative relationships* (pp. 93–109). New York: Brunner-Routledge.
- Bohart, A. C. (2000). Paradigm clash: Empirically supported treatments versus empirically supported psychotherapy practice. *Psychotherapy Research, 10*, 488–493.
- Buckley, P., Karasu, T. B., Charles, E., & Stein, S. P. (1979). Theory and practice in psychotherapy: Some contradictions in expressed belief and reported practice. *Journal of Nervous & Mental Disease, 167*, 218–223.
- Caspar, F., Pessier, J., Stuart, J., Safran, J. D., Wallner Samstag, L., & Guirguis, M. (2000). One step further in assessing how interpretations influence the process of psychotherapy. *Psychotherapy Research, 10*, 309–320.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*, 7–18.
- Curtis, J. T., Silberschatz, G., Sampson, H., & Weiss, J. (1994). The plan formulation method. *Psychotherapy Research, 4*, 197–207.
- Derogatis, L. R., & Savitz, K. L. (2000). The SCL-90-R and Brief Symptom Inventory (BSI) in primary care. In M. E. Maruish (Ed.), *Handbook of psychological assessment in primary care settings* (pp. 297–334). Mahwah, NJ: Erlbaum.
- DeRubeis, R. J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research, 14*, 469–482.
- Elkin, I. (1988, June). *The relationship of adherence to outcome*. Paper presented at the Society for Psychotherapy Research, Santa Fe, NM.
- Elkin, I. (1994). The NIMH treatment of depression collaborative research program: Where we began and where we are. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 114–139). New York: Wiley.
- Fiedler, F. E. (1951). Factor analysis of psychoanalytic, nondirective, and Adlerian therapeutic relationships. *Journal of Consulting Psychology, 15*, 32–38.
- Foreman, S. A., Gibbins, J., Grienberger, J., & Berry, J. W. (2000). Developing methods to study child psychotherapy using new scales of therapeutic alliance and progressiveness. *Psychotherapy Research, 10*, 450–461.
- Fretter, P. B. (1995). A control-mastery case formulation of a successful treatment for major depression. In *Session: Psychotherapy in Practice, 1*, 3–17.
- Fretter, P. B., Bucci, W., Broitman, J., Silberschatz, G., & Curtis, J. T. (1994). How the patient's plan relates to the concept of transference. *Psychotherapy Research, 4*, 58–72.
- Gottman, J. M. (1973). N-of-one and N-of-two research in psychotherapy. *Psychological Bulletin, 80*, 93–105.
- Gottman, J. M., & Ringland, J. (1981). The analysis of dominance and bidirectionality in social development. *Child Development, 52*, 393–412.
- Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 6*, 278–296.
- Hill, C. E., O'Grady, K. E., & Elkin, I. (1992). Applying the collaborative study psychotherapy rating scale to rate therapist adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management. *Journal of Consulting and Clinical Psychology, 60*, 73–79.
- Hilliard, R. B. (1993). Single-case methodology in psychotherapy process and outcome research. *Journal of Consulting and Clinical Psychology, 61*, 373–380.
- Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist, 41*, 159–164.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.
- Jones, E. E., Ghannam, J., Nigg, J. T., & Dyer, J. F. P. (1993). A paradigm for single-case research: The time series study of a long-term psychotherapy for depression. *Journal of Consulting and Clinical Psychology, 61*, 381–394.
- Jones, E. E., Hall, S., & Parke, L. A. (1991). The process of change: The Berkeley Psychotherapy Research Group. In L. Beutler & M. Crago (Eds.), *Psychotherapy research: An international review of programmatic studies* (pp. 98–107). Washington, DC: American Psychological Association.
- Jones, E. E., Parke, L. A., & Pulos, S. (1992). How therapy is conducted in the private consulting room: A multivariate description of brief psychodynamic treatments. *Psychotherapy Research, 2*, 16–30.
- Jones, E. E., & Pulos, S. M. (1993). Comparing the process in psychodynamic and cognitive-behavioral therapies. *Journal of Consulting and Clinical Psychology, 61*, 306–316.
- Levenson, R. W., & Gottman, J. W. (1985). Physiological and affective predictors of change in relationship satisfaction. *Journal of Personality and Social Psychology, 49*, 85–94.
- Messer, S. B., Tishby, O., & Spillman, A. (1992). Taking context seriously in psychotherapy research: Relating therapist interventions to patient progress in brief psychodynamic therapy. *Journal of Consulting and Clinical Psychology, 60*, 678–688.
- Norville, R., Sampson, H., & Weiss, J. (1996). Accurate interpretations and brief psychotherapy outcome. *Psychotherapy Research, 6*, 16–29.
- O'Connor, L. E. (2002). Pathogenic beliefs and guilt in human evolution: Implications for psychotherapy. In P. Gilbert & K. Bailey (Eds.), *Genes of the couch: Explorations in evolutionary psychology* (pp. 276–303). London: Brunner-Routledge.
- O'Connor, L. E., Berry, J. W., & Weiss, J. (1999). Interpersonal guilt, shame, and psychological problems. *Journal of Social and Clinical Psychology, 18*, 181–203.
- Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in

- psychotherapy—Noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 270–376). New York: Wiley.
- Pole, N., Ablon, J. S., O'Connor, L. E., & Weiss, J. (2002). Ideal control mastery technique correlates with change in a single case. *Psychotherapy: Theory, Research, Practice, Training*, 39, 88–96.
- Pole, N., & Bloomberg-Fretter, P. (2006). Using control mastery therapy to treat major depression and posttraumatic stress disorder. *Clinical Case Studies*, 5, 53–70.
- Pole, N., & Jones, E. E. (1998). The talking cure revisited: Content analyses of a two-year psychodynamic psychotherapy. *Psychotherapy Research*, 8, 171–189.
- Rappoport, A. (1997). The patient's search for safety: The organizing principle in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, 34, 250–261.
- Rice, L. N., & Greenberg, L. S. (1984). The new research paradigm. In L. N. Rice & L. S. Greenberg (Eds.), *Patterns of change* (pp. 7–25). New York: Guilford Press.
- Sampson, H. (1990). How the patient's sense of danger and safety influence the analytic process. *Psychoanalytic Psychology*, 7, 115–124.
- Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The Consumer Reports Study. *American Psychologist*, 50, 965–974.
- Shapiro, D. A., & Startup, M. (1992). Measuring therapist adherence in exploratory psychotherapy. *Psychotherapy Research*, 2, 193–203.
- Shaw, B. F., Elkin, I., Yamaguchi, J., Olmstead, M., Vallis, T. M., Dobson, K. S., et al. (1999). Therapist competence ratings in relation to clinical outcome in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, 67, 837–846.
- Shea, M. T., Elkin, I., Imber, S. D., Sotsky, S. M., Watkins, J. T., Collins, J. F., et al. (1992). Course of depressive symptoms over follow up: Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Archives of General Psychiatry*, 49, 782–787.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428.
- Silberschatz, G. (2005). An overview of research on control-mastery theory. In G. Silberschatz (Ed.), *Transformative relationships* (pp. xxx–xxx). New York: Brunner-Routledge.
- Silberschatz, G., & Curtis, J. T. (1993). Measuring the therapist's impact on the patient's therapeutic progress. *Journal of Consulting and Clinical Psychology*, 61, 403–411.
- Silberschatz, G., Fretter, P. B., & Curtis, J. T. (1986). How do interpretations influence the process of psychotherapy? *Journal of Consulting and Clinical Psychology*, 54, 646–652.
- Stephenson, W. (1953). *The study of behavior: Q-technique and its methodology*. Chicago: University of Chicago Press.
- Weiss, J. (1993). *How psychotherapy works*. New York: Guilford Press.
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631–663.
- Williams, E. A., & Gottman, J. M. (1982). *A user's guide to the Gottman-Williams time-series analysis computer programs for social scientists*. New York: Cambridge University Press.

Received December 11, 2006

Revision received December 3, 2007

Accepted December 3, 2007 ■

## Low Publication Prices for APA Members and Affiliates

**Keeping you up-to-date.** All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

**Essential resources.** APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association*, and on dozens of new topical books each year.

**Other benefits of membership.** Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

**More information.** Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.